Anastasia Efklides
Plousia Misailidi
Editors

Trends and Prospects in Metacognition Research

Springer

Trends and Prospects in Metacognition Research

Anastasia Efklides · Plousia Misailidi
Editors

Trends and Prospects in Metacognition Research

Springer

Editors
Anastasia Efklides
Aristotle University of Thessaloniki
School of Psychology
541 24 Thessaloniki
Greece
efklides@psy.auth.gr

Plousia Misailidi
Department of Primary Education
University of Ioannina
451 10 Ioannina
Greece
pmisaili@uoi.gr

ISBN 978-1-4419-6545-5 e-ISBN 978-1-4419-6546-2
DOI 10.1007/978-1-4419-6546-2
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010930044

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Contributors

Mary Ainley
University of Melbourne, Melbourne, VIC, Australia

Carl Martin Allwood
Department of Psychology, University of Gothenburg, BOX 500,
SE-40530 Gothenburg, Sweden
cma@psy.gu.se

Qais Almeqdad
Yarmouk University, Department of Counseling and Educational Psychology,
Irbid, Jordan

Alessandro Antonietti
Catholic University of the Sacred Heart, Milano, Italy

Elisabeth Bacon
Psychiatry Department, National Institute for Health and Medical Research
(U666 INSERM) University Hospital, BP 426, 67091 Strasbourg
Cedex, France
bacon@alsace.u-strasbg.fr

Igor Bajšanski
University of Rijeka, 51000 Rijeka, Croatia

Kathryn Bartimote-Aufflick
Institute for Teaching and Learning, University of Sydney, Carslaw F07, NSW,
2006, Australia
k.aufflick@usyd.edu.au

Michael J. Beran
Language Research Center, Georgia State University,
Atlanta, GA 30302-3965, USA
mjberan@yahoo.com

Joseph Boomer
Department of Psychology, State University at Buffalo, Atlanta, GA, USA

Angela Brew
Macquarie University, Sydney, Australia

Donna Bryce
University of Cambridge, 184 Hills Rd, Cambridge, UK

Barbara Colombo
Department of Psychology, Catholic University of the Sacred Heart, Largo Gemelli 1, 20123 Milano, Italy
barbara.colombo@unicatt.it

Justin J. Couchman
Department of Psychology, State University at Buffalo, Atlanta, GA, USA

Mariana V. C. Coutinho
Department of Psychology, State University at Buffalo, Atlanta, GA, USA

Csaba Csíkos
Institute of Education, University of Szeged, Petőfi S. sgt. 30-34, 6722 Szeged, Petőfi S. sgt. 30-34, Hungary
csikoscs@edpsy.u-szeged.hu

Demetra Demetriou
University of Cambridge, 184, Hills Rd, Cambridge, UK

Annemie Desoete
University College Arteveldehogeschool, Ghent, Belgium;
Department of Clinical Psychology, University of Ghent, Ghent, Belgium

Zoltán Dienes
Department of Psychology, Pevensey Building, University of Sussex, Falmer, BN1 9QH, UK
z.dienes@sussex.ac.uk

Simon C. Duff
University of Liverpool, Liverpool, UK

Anastasia Efklides
School of Psychology, Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece
efklides@psy.auth.gr

Anne-Mari Folkesson
Department of Human Sciences, Kalmar University, 391 82 Kalmar, Sweden, anne-mari.folkesson@hik.se

Simona Ghetti
Department of Psychology & Center for Mind and Brain, University of California, Davis, CA, USA

Valeska Grau
University of Cambridge, 184, Hills Rd, Cambridge, UK

Paola Iannello
Catholic University of the Sacred Heart, Milano, Italy

Marie Izaute
Social and Cognitive Psychology Laboratory (LAPSCO UMR CNRS 6024),
Clermont University, Blaise Pascal University, 34 avenue Carnot,
63037 Clermont-Ferrand, France
marie.izaute@univ-bpclermont.fr

Riitta Kinnunen
Centre of Learning Research, University of Turku, FIN-20014, Finland
riitta.kinnunen@utu.fi

Sabina Kleitman
School of Psychology, The University of Sydney, Sydney NSW 2006, Australia
sabinak@psych.usyd.edu.au

Svjetlana Kolić-Vehovec
Department of Psychology, University of Rijeka, Croatia,
Slavka Krautzeka b.b., 51000 Rijeka, Croatia
skolic@ffri.hr

Kristen E. Lyons
Department of Psychology & Center for Mind and Brain, University of California,
1 Shields Avenue, Davis, CA 95616, USA
kelyons@umn.edu

Plousia Misailidi
Department of Primary Education, University of Ioannina,
451 10 Ioannina, Greece
pmisaili@uoi.gr

Tanya Moscrop
School of Psychology, the University of Sydney, Sydney, NSW 2006, Australia

Elisabeth Norman
Department of Psychosocial Science, Psychology Faculty,
University of Bergen, Christies gate 12, 5015 Bergen, Norway
Elisabeth.Norman@psysp.uib.no

Mark C. Price
Department of Psychosocial Science, University of Bergen,
Christies gate 12, 5015 Bergen, Norway

Barbara Rončević Zubković
University of Rijeka, 51000 Rijeka, Croatia

Claire Sangster
University of Cambridge, 184, Hills Rd, Cambridge, UK

J. David Smith
Department of Psychology and Center for Cognitive Science,
State University at Buffalo, Buffalo, NY, USA

Ryan Scott
Department of Psychology, University of Sussex, Pevensey Building,
Falmer, BN1 9QH, UK
r.b.scott@sussex.ac.uk

János Steklács
Teacher Training Faculty, Kecskemét College, 6000 Kecskemét,
Kaszap u. 6-14, Hungary
steklacs@gmail.com

Lena Swalander
Department of Human Sciences, Kalmar University, 391 82 Kalmar, Sweden &
Department of Behavioural Sciences and Learning, Linköping University,
581 83 Linköping, Sweden
lena.swalander@hik.se

Alexandra Touroutoglou
School of Psychology, Aristotle University of Thessaloniki,
541 24 Thessaloniki, Greece

Ruth Vanderswalmen
Arteveldehogeschool, Campus Kantienberg,
Opleiding Logopedie, Voetweg 66, B-9000 Gent, Belgium
ruth.vanderswalmen@arteveldehs.be

Marja Vauras
University of Turku, FIN-20014 Turku, Finland

Joke Vrijders
Department of Clinical Psychology, University of Ghent, Ghent, Belgium

David Whitebread
University of Cambridge, 184, Hills Rd., Cambridge, UK
dgw1004@cam.ac.uk

Chapter 1
Introduction: The Present and the Future in Metacognition

Anastasia Efklides and Plousia Misailidi

It has been more than 30 years since the notion of *metacognition* was first introduced by Flavell (1976, 1979). In the ensuing years, a multitude of phenomena representative of metacognition have been studied both in basic and applied psychological research. Yet, there are still major issues that require our attention and pose challenges for future research. The conceptualization of metacognition and understanding of the mechanism(s) underlying its functioning constitute the top priority of theory and basic research. The relations of metacognition with consciousness and self-regulation are also issues at the core of research in metacognition, whereas the development of metacognition and the trainability of metacognitive skills are two of the main research areas in developmental and educational psychology. At the same time, metacognition in animals is a growing area of interest and so is metacognition in clinical populations, such as schizophrenics. All these developments promise a bright future for metacognition research, owing particularly to the development of new methodologies which will allow a deeper insight into the nature of metacognitive phenomena.

Yet, one often has the feeling that the communication between the various lines of metacognition research is limited and that there is lack of theoretical integration; overcoming this drawback is a challenge for the future. This book is an attempt, first, to bring together work in various areas of research on metacognition representing trends that may have implications for theory and future research. Second, it aims to underscore convergences and divergences that allow, on the one hand, theoretical integration and on the other discrimination between the various manifestations of metacognition. In this way the prospects of research on metacognition will be made salient.

A. Efklides (✉)
School of Psychology, Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece
e-mail: efklides@psy.auth.gr

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_1,

The book is organized into two parts, the first being more directed towards basic research and the second towards developmental and applied research in educational contexts. In what follows, first, we present a brief overview of each of the book chapters and then discuss the underlying convergences and divergences, the possible integration of ideas and findings that seem unconnected and disparate, as well as the methodological developments that allow new approaches to the study of metacognitive phenomena. Finally, the prospects of metacognition research, as revealed by the various contributions to the book, will be highlighted.

1 The Contributing Chapters

1.1 Part I: Basic Research in Metacognition

The basic questions "What is the nature of metacognition and how is metacognition functioning?" are the epicenter around which the nine chapters comprising the first part of the book are organized. All contributors to this book share the definition of metacognition as cognition of cognition (Flavell, 1979) that has two functions, namely monitoring and control of cognition (see also Nelson, 1996). However, the distinction between cognition and metacognition is often hard to be made and the diversity of metacognitive phenomena suggests that there is no single mechanism that can explain them all. Therefore, in order to go to the roots of metacognition one has to consider, first, whether non-human animals have metacognition and, if yes, what form it takes; second, whether metacognition is the product of solely conscious processes, or whether nonconscious[1] processes can also give rise to metacognition; third, what is the cognitive basis of metacognitive phenomena and how is metacognitive awareness, or the phenomenological aspects of metacognition, formed.

To open the discussion on the nature of metacognition Beran, Couchman, Coutinho, Boomer, and Smith (this volume) provide a critical overview of animal metacognition research. The authors start with the fundamental question "Is metacognition a uniquely human trait?" (see Chap. 2) and present evidence suggesting that it is not. They discuss different research paradigms and the critique that is often leveled at them, that is, whether these paradigms indeed measure metacognition rather than simply cognition. However, as the authors argue, the evidence coming from studies performed using the uncertainty monitoring paradigm clearly shows that animals monitor uncertainty and regulate their behavior accordingly.

The chapter by Scott and Dienes (this volume) addresses the passage from unconscious to conscious knowledge and the building of metacognition. Specifically, the authors focus on implicit learning (as studied by artificial grammar learning) as opposed to explicit learning. The former is learning without metacognition whereas the latter is learning with metacognition. Scott and Dienes (this volume) argue that

[1]Except for the term "nonconscious", the term "unconscious" is also used in some of the chapters of this book.

implicit (or unconscious) learning can be best described by a single updating model of learning that is similar to the ones described by connectionist theories. Single updating processes allow the representation of the structure of reality; for instance, the rule underlying strings of letters. The person, however, has no awareness of the processes through which implicit knowledge is formed or what the rule is. Despite this lack of awareness, implicit knowledge can support judgments on whether new items (e.g., new stings of letters) conform to the ones previously presented. People can make such judgments with reasonable accuracy and confidence, and the phenomenological experience that accompanies such judgments is that of "intuition" rather than guessing. Such intuitive judgments are presumably based on familiarity (i.e., prior presentations). However, accurate judgments about new items make one aware of his or her knowledge, thus turning unconscious knowledge to conscious.

In the next chapter, Norman, Price, and Duff (this volume) introduce the concept of fringe consciousness that was originally described by William James (1890) and elaborated on by Mangan (1993). According to Norman et al. (this volume), fringe consciousness refers to the "transitive, fleeting, and inarticulate content of conscious experiences" (see Chap. 4). It is distinct from explicit, articulate contents of consciousness, and involves evaluation of implicit content knowledge and processing. Fringe consciousness reflects the monitoring of online cognitive processing and, therefore, it is a manifestation of online metacognition. It shares many similarities with the concept of experience-based metacognitive feelings introduced by Koriat (Koriat & Levy-Sadot, 1999). However, experience-based metacognitive feelings capture properties of cognitive processing rather than of content knowledge (feeling of knowing is an example of such an experience-based metacognitive feeling). Despite their similarities, fringe consciousness, according to Norman et al. (this volume), has features that go beyond those of experience-based metacognitive feelings: (a) fringe consciousness can reflect not only nonconscious processes but also nonconscious content; (b) it has a wider set of cognitive functions; and (c) the degree of introspective access is variable. These features suggest that fringe consciousness plays a mediating role between nonconscious and conscious cognitive processing.

The chapter authored by Bacon (this volume) is addressing the tip-of-the-tongue (TOT) phenomenon and the dissociation of the cognitive from the metacognitive (or phenomenological) aspect of TOT. The TOT state is a metamemory experience denoting the intense feeling of temporary inaccessibility of a target item (e.g., a word and/or the answer to a question) although it is available in memory. The phenomenological features of a TOT state comprise, besides the awareness that the target item is available in memory, the anticipation that retrieval of the correct response is imminent albeit currently inaccessible. Moreover, there is a persistent alternate (or part of the correct response) which blocks the retrieval of the target item and which the person recognizes as incorrect response. Using the amnesic drug lorazepam as a tool Bacon (this volume) showed that lorazepam induced temporary micro-amnesia, that is, inaccessibility of a target item available in memory and, at the same time, presence of a persistent alternate. In other words, lorazepam induced a cognitive TOT state. Yet, participants under lorazepam did not experience the phenomenological TOT state, namely the awareness that the persistent alternate is not the correct one and that it blocks the accessibility of the correct response. As a consequence, lorazepam participants made significantly more commission errors (i.e., responded with the persistent alternate as if it were the correct response) than placebo participants. However, when they were given feedback that their response was not correct, they experienced the phenomenological TOT, and this was associated with retrieval of the correct response as it happened with placebo participants.

The chapter by Efklides and Touroutoglou (this volume) further elaborates the distinction between cognition and metacognition. The authors refer to another metamemory phenomenon, namely the blank-in-the-mind (BIM) experience. This is an experience a person has in the context of prospective memory when there is temporary inaccessibility of the intention or the cue that should trigger the enactment of a purported action. Thus, BIM shares with TOT the temporary inaccessibility of a piece of information although it is available in memory. However, in the case of BIM the phenomenological experience denotes no awareness of a persistent alternate; instead there is awareness of having a blank in the mind; moreover, unlike TOT, the anticipation is that the missing intention or cue will not be retrieved despite one's efforts. Therefore, phenomenologically, the BIM is differentiated from TOT. The authors investigated the hypothesis that working memory load might underlie the failure to maintain in memory the prospective memory intention or cue. The results showed that working memory demands increased prospective memory failures but did not differentiate the retrospective ratings of BIM as compared to those of TOT. Yet, there was no correlation between BIM and TOT ratings. Thus, the evidence reveals two metamemory subjective experiences which, despite their shared cognitive basis, employ different processes leading to different metacognitive experiences, that is, to BIM and TOT.

Izaute and Bacon (this volume) also address metamemory and, specifically, judgments of learning (JOL) and allocation of study time in schizophrenics. However, the emphasis here is not on the processes underlying the formation of JOLs, but on the dissociation between the monitoring and control functions of metacognition. Schizophrenia is a mental disorder characterized by cognitive deficits besides the characteristic negative and positive symptoms that reflect emotional, social, and behavioral deficiencies. The question posed by the authors is whether there are metamemory deficits as well, and whether these deficits are found in the monitoring or the control of memory. Izaute and Bacon (this volume) present evidence from two of their own studies suggesting that monitoring is spared in schizophrenia whereas control is impaired. This finding is in line with evidence that schizophrenia impairs strategic regulation of behavior. More importantly, however, it shows that the usual assumption that there is a close relationship between monitoring and control in metacognition is not self-evident, because there can be a dissociation between these two.

Allwood (this volume) discusses a recurrent question in metamemory research, namely how veridical and accurate metamemory judgments can be. The author focuses on the realism of children's confidence judgments when they answer questions on their episodic memory – a topic of high relevance to eyewitness testimonies. Allwood (this volume) discusses two aspects of realism in metacognitive judgments – calibration and discrimination – and the factors that have been found to influence metacognitive accuracy. He then reviews evidence from studies with his collaborators in which some of the factors known to affect realism, such as type of measurement scale or type of question asked, were investigated. The evidence presented suggests that children can have a level of realism that is comparable to that of the adults.

The next chapter, authored by Touroutoglou and Efklides (this volume), turns our attention to a metacognitive experience that is characteristic of problem solving rather than of memory. The focus is on feeling of difficulty and the cognitive factors that affect it. Feeling of difficulty monitors the lack of fluency in cognitive processing (Efklides, 2002). Usually people attribute the difficulty they experience in problem solving either to task difficulty or to task complexity. However, the same task may be perceived as easy by some people and difficult by others. Efklides (2001, 2006) claims that feeling of difficulty is the interface between the person and the task. Moreover, feeling of difficulty is associated with negative affect and motivates effort exertion and strategy use. Touroutoglou and Efklides (this volume) tested two hypotheses regarding potential factors that may produce lack of fluency in cognitive processing, namely working memory load and cognitive interruption caused by events that are discrepant to an activated processing schema. Two experiments confirmed the working-memory-load hypothesis and a third showed that cognitive interruption does indeed increase feeling of difficulty. The authors also showed that cognitive interruptions due to discrepant events triggered not only feeling of difficulty but surprise as well. This is an important finding because it reveals the close relation between metacognition, in the form of feeling of difficulty, and emotions, such as surprise. Surprise serves the relocation of attention from the prevalent schema to the discrepant event. Feeling of difficulty along with surprise provide the input for better appraisal of the demands of the situation as well as for better control decisions.

The last chapter of the first part of the book (Chap. 10) is authored by Kinnunen and Vauras and deals with methodological issues in the measurement of reading metacomprehension. Methodological issues are of critical importance for the study of metacognitive phenomena because they can reveal different aspects of metacognitive phenomena as well as new phenomena and their underlying mechanisms. Kinnunen and Vauras (this volume) discuss the pros and cons of early metacomprehension measures and the use of methods that can provide data on online monitoring and regulation of reading comprehension. Hence, the emphasis of this chapter is on online methods and, particularly, on traced silent reading and eyetracking. With recent technological advancements, collection of online data has been greatly facilitated, allowing insights into children's growing awareness of their comprehension as they develop and become more proficient readers. Kinnunen and Vauras (this volume) provide evidence from their studies showing the potential of online data for understanding not only reading metacomprehension per se but also the effects that emotional and situational factors may have on the accuracy of comprehension monitoring.

1.2 Part II: Developmental and Educational Implications of Metacognition

The second part of the book comprises ten chapters. The first four are addressing the development of metacognition in young children, whereas the next four chapters deal with metacognition in school-aged children. The topics in these chapters are applied and related to basic skills, such as reading, spelling, and computer use. The last two chapters are focusing on adults, their self-awareness and how it can interact with self-regulation in their professional lives.

Whitebread, Almeqdad, Bryce, Demetriou, Grau, and Sangster (this volume) make the case that metacognition in young children should be studied in natural settings and with different methods (e.g., observation) than those used for older children or adults (e.g., self-reports). Moreover, metacognition in young children should not, according to these authors, be reduced to the acquisition of a theory of mind but should also include other facets of metacognition, such as metacognitive knowledge and metacognitive regulation of behavior. In young children metacognition is often based on implicit rather than explicit processes, and the role of social context (e.g., parents, peers) is of critical importance, because it may facilitate or hinder the development of metacognition. A series of studies that were carried out by Whitebread and his co-authors revealed interrelationships between early metacognition, executive functions (e.g., inhibitory control), theory of mind, conceptual development, as well as with learning and motor difficulties. In all cases, complex interactions of metacognition with cognition, motivation, and affect in the self-regulation of behavior were detected.

In the same line of reasoning, the next chapter, authored by Lyons and Ghetti (this volume), acknowledges monitoring and control processes even in very young children. The literature review on young children's monitoring of their mental activities suggests that, from the end of the 2nd year onwards they are able to use mental verbs referring to knowledge states (i.e., know, think) and to monitor their comprehension and mental imagery as well as their memory. Young children can also exercise control over their behavior by asking questions when they are uncertain about their knowledge or refraining from answering when they are not confident that their memory is accurate. Based on such evidence Lyons and Ghetti (this volume) propose a model of early metacognitive development and of the relations between monitoring and control in self-regulation. This model builds on the development of uncertainty monitoring and asserts that development of metacognition does not necessarily occur uniformly across domains, nor are monitoring and control functions closely linked.

The third chapter on the development of metacognition in children is authored by Misailidi (this volume). It discusses, first, why theory-of-mind (TOM) research has not been connected to research on the development of metacognition and presents arguments and research evidence that shows interrelations between TOM and metamemory. Misailidi (this volume) also connects TOM with comprehension of metacognitive language, which is an indispensable part of metacognition, and provides evidence that TOM precedes the comprehension of mental verbs.

The fourth chapter on children's development of metacognition is authored by Kleitman and Moscrop (this volume). The authors focus on school-age children's self-confidence as a person characteristic. It is known that self-confidence influences adults' confidence judgments but it is not known whether children's self-confidence acts in a similar way. The question Kleitman and Moscrop (this volume) set out to answer regards the sources of children's self-confidence and particularly the role of parent-child relationship. They present a study with 9–12-year-olds showing that self-confidence is a reliable and stable construct in this age population, and that this person characteristic, which is similar but not identical to self-efficacy, predicts school achievement beyond the effects of cognitive ability, age, and gender. Most importantly, however, parental care proved to be a predictor of children's self-confidence.

The educational implications of metacognition regard learning of basic skills such as reading and writing. Metacognition in reading is the topic of the next chapter, authored by Kolić-Vehovec, Bajšanski, and Rončević Zubković (this volume). The authors provide an overview of their recent studies with upper elementary and high school students. The emphasis is on age and gender differences in reading comprehension as well as in the various relevant components of metacognition. Metacognitive knowledge about reading was found to be related to increasing age but there were no systematic gender differences in it. On the contrary, comprehension monitoring was found to be a component of reading-related metacognition, in which gender differences, in favour of girls, were evident.

Interventions based on findings from research on metacognition in reading are of particular importance for education. The chapter by Csíkos and Steklács (this volume) is presenting two reading intervention programs that made use of prior research evidence on reading processes. The intervention programs were applied to Hungarian fourth-graders in the context of classroom reading instruction. Students were instructed how to use different reading strategies but they were also supported to develop metacognitive knowledge and awareness of strategy use. The results were encouraging and highlighted the importance of interventions that increase metacognitive awareness of reading processes.

Although reading and writing have attracted a lot of attention in metacognition research, spelling has not been extensively studied. The chapter by Vanderswalmen, Vrijders, and Desoete (this volume) presents a study in which the authors gathered data on the spelling performance of 1st-year college students. There were measures of three facets of metacognition as manifested in spelling, namely: (a) students' metacognitive knowledge of themselves as spellers, (b) students' use of metacognitive skills in spelling, and (c) students' metacognitive experiences after a spelling test. The authors categorized different spelling errors and tested the monitoring accuracy of spelling performance. The results showed that proficient spellers tended to underestimate their spelling performance more relative to poor spellers. Overall, students calibrated their judgments of how correct their spelling was vis-à-vis their spelling performance. Moreover, calibration of metacognitive judgments, metacognitive knowledge, and use of metacognitive skills predicted spelling performance.

The next chapter, authored by Swalander and Folkesson (this volume), looks at metacognition within the context of self-regulated learning. The authors observed how two teachers introduced computer use in their classrooms and how young students (aged from 6–7 to 9 years) learnt this new tool and integrated it into their everyday school practice. The data were content-analyzed in three major categories, that is, metacognition, motivation, and behaviour. As regards metacognition, it was found that children, in the main, were aware of processes in the new learning environment, self-monitored their working progress, planned their activities, and were aware of relations between different learning activities. However, there were some situations in which children's metacognition was not sufficient to help them deal with a problem; for example, children could not benefit from computer functions such as spelling or technical messages and wanted to be led by the teacher. At the same time, as children worked with the computer they experienced positive or negative affect, saw opportunities for making choices, were eager to learn, and viewed teacher demands either as positive or negative extrinsic motivation. At the behavioural level, children helped each other and actively participated in collaborative projects. This kind of evidence comes to add to claims that young children show not only metacognition but also self- and co-regulation of their behaviour in new learning environments.

The chapter by Bartimote-Aufflick, Brew, and Ainley (this volume) extends the picture of self-regulated learning to encompass the teacher. The question these authors pose is whether teachers engage in critical self-regulation and, if this is the case, whether this has effects on their instructional practices and their students' learning. Critical self-regulation is based on the assumption that adults (in this case university teachers) when found in dilemmatic situations engage in reflection, and reflective processes shape their inquiry and self-regulation. Thus, Bartimote-Aufflick et al. (this volume) extend Zimmerman's (2004) three-phase model to include a "prior" phase, before forethought, which is based on reflective processes and guides the subsequent self-regulation phases. This prior phase comprises teachers' "reflection on the basic premises of their instruction and consideration of higher-order instructional goals" (see Chap. 19). At the end of the self-regulation process (i.e., at the evaluation phase) critical reflection is advocated as a determinant of deeper self-reflection on one's professional work.

The last chapter of Part II, and of the book, deals with another issue that is pertinent to adults' metacognition with respect to their personal and professional lives. Colombo, Ianello, and Antonietti (this volume) explore the metacognitive knowledge people have of their decision-making processes, their related affective experiences as well as their intuitive vs. deliberative/analytic decision-making style. This is a novel topic in metacognition research. The findings of the study suggest that metacognitive knowledge of the self as decision-maker, of the occasions in which decision-making is required, and of the strategies used when making decisions is linked to the types of decisions the person is required to do in his/her professional (or personal) life and only partly to the intuitive or analytic decision-making style.

2 Convergences and Divergences

Having outlined the content of the chapters of the book, we now come to the discussion of major issues that cut across the various research paradigms and findings presented. In what follows, the distinction between the two parts of the book no longer holds; instead, the threads connecting basic and applied research, the facets of metacognition, and the development of metacognition will be highlighted. The points of convergence and, at the same time, of divergence between the chapters can be conceptualized as follows: (a) the importance of uncertainty monitoring; (b) conscious versus nonconscious processes in metacognition; (c) methodological issues in metacognition research; (d) individual differences in metacognition; and (e) metacognition and self-regulation.

2.1 The Importance of Uncertainty Monitoring

The basic finding of the studies reviewed by Beran et al. (this volume) is that non-human animals monitor uncertainty and this has implications for their behavior. Uncertainty monitoring is fundamental for successful adaptation because it informs whether one should respond or not to an external stimulus. Uncertainty may arise from stimulus features that do not allow discrimination between alternative stimuli; it may also arise in situations in which there is conflict of response and the proper response cannot be formed immediately. Uncertainty monitoring is, therefore, essential to inform the non-human animal for a go/no-go decision. According to Lyons and Ghetti (this volume) even young children monitor uncertainty as suggested by their information-seeking behavior and refraining from responding.

Despite the importance of uncertainty monitoring, most of the research with humans is based on the idea that confidence is the most important indicator of one's knowledge state. Confidence (or lack of confidence) in one's memory or comprehension emerges in early childhood (Allwood, this volume; Lyons & Ghetti, this volume; Whitebread et al., this volume) and is present throughout life (see, e.g., Vanderswalmen et al., this volume), although the present book does not go into older adults' metacognition. Confidence in one's memory is also present in clinical populations (Izaute & Bacon, this volume). However, confidence is quite different from uncertainty monitoring (see Allwood, this volume), although one could assume that lack of confidence is equivalent to uncertainty. Confidence is related to the monitoring of the correctness of one's response; for example, whether one can retrieve from memory the correct piece of information (e.g., feeling of knowing) or whether the response produced is correct or not (Efklides, 2002; Vanderswalmen et al., this volume). Thus, uncertainty monitoring and confidence seem to overlap only partly, that is, when the person is not confident that the answer is available in memory or is correct and, therefore, refrains from answering. This divergence between the two notions (uncertainty vs. confidence) implies that more research is needed to conceptually distinguish them and to consider how they could be integrated into a theoretical framework that could account for what are the critical features that constitute the object of the monitoring process.

2.2 Nonconscious Vs. Conscious Processes in Metacognition

Comparative research is important not only for tracing the origins of metacognition but also because it challenges the view that metacognition is solely a conscious process. In fact, a fundamental part of the puzzle of metacognition concerns the conscious versus nonconscious character of the monitoring and control processes, a dilemma that has important implications for the conceptualization of metacognition. Some of the manifestations of metacognition are declarative and expressible products of reflection (Bartimote-Aufflick et al., this volume; Colombo et al., this volume) and of analytic monitoring processes that can lead to deliberate control processes, whereas others are products of nonconscious (Beran et al., this volume; Scott & Dienes, this volume), non-analytic processes that monitor the state of one's knowledge or cognitive processing (see also, Norman et al., this volume). Humans make use of both forms of metacognition (i.e., conscious and nonconscious) but non-human animals and very young children share only the latter. The theoretical challenge in this case is not only to be able to delimit the nature of unconscious processes that can give rise to metacognitive judgments at a conscious level (Scott & Dienes, this volume) but also to be able to understand, first, the nature of monitoring and control at the nonconscious level and, second, the mechanism(s) that are responsible for the phenomenology of the various metacognitive feelings that are denotative of underlying cognitive states, contents, or features of cognitive processing. That is, the experiential qualities of fringe consciousness (Norman et al., this volume) or of metacognitive feelings (Efklides & Touroutoglou, this volume; Touroutoglou & Efklides, this volume) need to be explained along with their cognitive source.

Specifically, people "feel" uncertain, confident, or that something is familiar or novel; they "feel" they know something rather than they guess or remember; they "feel" that their judgment or decision is based on intuition rather than on analytical reasoning or on guessing; they experience a feeling of difficulty (FOD) or a tip-of-the-tongue (TOT) state, or blank in the mind (BIM). These metacognitive feelings constitute a "language" that people share and try to decode in order to be able to make informed decisions for control processes. Norman et al. (this volume), Touroutoglou and Efklides (this volume), Bacon (this volume), and Efklides and Touroutoglou (this volume) focus on such metacognitive feelings and their cognitive origins. It is evident that even in related metacognitive experiences (TOT state vs. BIM experience; Efklides & Touroutoglou, this volume) their phenomenology conveys different meanings that have different implications for control processes (e.g., persist or abstain from further memory search).

Obviously, in order to achieve a deeper understanding of the processes underlying the functioning of metacognition, the integration of different theoretical frameworks, such as those of fringe consciousness and metacognitive experiences, is needed (see Norman et al., this volume). In any case, it is important to distinguish metacognitive experiences and/or fringe consciousness – that originate from unconscious mechanisms – from other facets of metacognition, such as metacognitive knowledge or metacognitive skills (Efklides, 2006, 2008), that originate from conscious analytic processes.

Metacognitive knowledge and metacognitive skills originate from conscious analytic processes, although arguably they can be products of nonconscious inferential processes as well (Efklides, 2001, 2008). An important source of metacognitive knowledge is the observation of the self and others as agents capable of thoughtful and purposive action (Bartimote-Aufflick et al., this volume; Colombo et al., this volume); reflection is also a key process for the formation of explicit, declarative metacognitive knowledge. However, even young children possess metacognitive knowledge, even though they are not explicitly aware of it and cannot verbally express it. As Misailidi (this volume) points out, theory of mind (TOM) can be conceived of as a form of metacognitive knowledge, and there is evidence for relations between TOM and later metacognition (see also Lyons & Ghetti, this volume; Whitebread et al., this volume). Acquisition of this kind of metacognitive knowledge is facilitated by the frequent use of mental language in a child's family, which is a factor that contributes to the acquisition of TOM as well. This implies that metacognitive knowledge is facilitated by social interaction and communication with others through implicit learning processes rather than explicit or analytical/reflective processes. At school, on the other hand, a lot of metacognitive knowledge as well as metacognitive skills are acquired through direct instruction (Csíkos & Steklács, this volume). Therefore, metacognitive knowledge and metacognitive skills can be considered as the outcomes of either implicit or explicit processes but the nonconscious processes that give rise to them are not the same as the online monitoring processes that are at the background of metacognitive feelings and metacognitive judgments. In all cases, it is a major challenge for metacognition research to reveal the specific mechanisms that account for the formation of all the facets of metacognition at different ages or conditions.

2.3 Methodological Issues in Metacognition

A major issue in metacognition research concerns the development and use of methodologies that can reveal nonconscious (or unconscious) monitoring and control processes. Such methodologies are extensively discussed in this volume by Beran et al., Scott and Dienes, Norman et al., Kinnunen and Vauras, and Whitebread et al. The pitfalls of extant methodologies regarding judgments of confidence are pointed out by Allwood (this volume), whereas new methodologies are introduced by Bacon (this volume), such as use of amnesic drugs, and by Efklides and Touroutoglou (this volume) in the assessment of the BIM experience and FOD (Touroutoglou & Efklides, this volume). Colombo et al. (this volume) also introduced a new instrument for measuring metacognitive knowledge of decision making that comprises, besides self-report items, the use of analogies in order to describe one's self as a decision-maker. All the above developments underscore the fact that steps forward in metacognition research have been made to reveal new metacognitive phenomena and get to the heart of the differentiation of cognition from metacognition as well as of monitoring from control. Yet, it is still a long way to go, particularly if one introduces the role of social cognition in the formation of metacognition (Yzerbyt, Lories, & Dardenne, 1998).

2.4 Individual Differences in Metacognition

As already mentioned, this book provides a broad age perspective and shows that age-related differences might be due more to the methodologies used rather than to developmental effects. Prior knowledge is, however, an individual difference factor that has repeatedly been shown to have an impact on metacognition. For example, ignorance can have effects both at the cognitive and metacognitive level as Kolić-Vehovec et al. (this volume) and Vanderswalmen et al. (this volume) have shown. Prior knowledge is, of course, associated with age as well as with the mode of learning (implicit vs. explicit), but also with the context in which one learns – formal (e.g., school) vs. informal (e.g., family). For example, the learning environment in school and the teacher-pupil or peer interaction (Swalander & Folkesson, this volume) may facilitate the acquisition of metacognitive knowledge and skills that can support successful self-regulated learning even in early primary school students. However, even in sophisticated and well supported learning environments, such as the one described by Swalander and Folkesson (this volume), there can be obstacles arising either from the lack of prior knowledge or skills (e.g., in the use of computers) that can block metacognition or the willingness to invest on (meta)cognitive search and knowledge acquisition.

Moreover, there can be students who are not willing or able to invest on metacognition and self-regulated learning (Swalander & Folkesson, this volume). The question is whether students' (meta)cognitive development changes when the conditions of learning change. For instance, self-reflection and acquisition of metacognitive knowledge might be facilitated depending on teacher or peer support, or simply on increase of domain-related knowledge. Conversely, it may be hindered by a learning environment which poses demands that go beyond the person's cognitive resources, or when significant others (e.g., teachers, peers, parents) do not invest on metacognition and reflection because the learning environments requires rote learning and ready-made explanatory schemata. Thus, the factors that contribute to the acquisition of elaborated metacognitive knowledge and metacognitive skills is an issue that requires further investigation because some students may respond to interventions at the classroom level, such as those implemented by Csíkos and Steklács (this volume), while others may not. In the latter case, alternative instructional methodologies are required.

Individual differences in metacognition can also be detected in teachers or adults in the context of their work environment or professional life. Bartimote-Aufflick et al. (this volume) argue for the importance of critical self-reflection in university teachers. The question is whether all teachers engage in critical reflection about teaching and, if not, who are willing to do so. Motivation or the teachers' goals and strivings are important factors for channelling their critical reflection towards the content to be taught, to the instructional methods and means, or to the adaptation of their knowledge and skills to those of the learners, etc. The willingness to invest on critical reflection at all is also an issue of concern because, as with young students, there can be individual differences in the extent to which one invests on explicit engagement with metacognition and self-regulation.

Individual differences in prior knowledge can also be conceptualized as differences in domain knowledge (e.g., different professions) rather than as content differences within a knowledge domain (high vs. low level of knowledge). Such between domain differences explain, for example, why a medical doctor can differ in the metacognitive knowledge of decision-making processes from a student or a teacher or artist (Colombo et al., this volume). Since metacognition is a meta-process, it is obvious that differences in the data base on which metacognition builds leads to differences in the conceptualization not only of the self but also of how one behaves in specific situations, or how one makes decisions.

Another individual difference factor is gender. The chapter by Kolić-Vehovec et al. (this volume) directly addressed this factor in the metacognition of reading skills and reading comprehension. The question is whether it is gender per se that produces the monitoring differences found on metacomprehension, or it is the interaction of gender with prior knowledge, or motivation to engage in reading and metacomprehension processes. In gender differences the role of social context (peers, teachers, cultural environment) can be very powerful because it impacts the expectations and interests students develop about school subjects and their skills.

Finally, another source of individual differences in metacognition is person trait-like characteristics. Intuitive versus deliberative/analytic cognitive style is a person characteristic that is pertinent to decision making but also to metacognition because presumably a deliberative style facilitates awareness of the thinking processes employed in decision making or problem solving. However, the effect of this individual difference factor was not found to be so strong in the differentiation of metacognitive knowledge about the self as decision maker (Colombo et al., this volume).

Another person characteristic that has been shown in the past, but also in the Kleitman and Moscrop (this volume) study, to have implications for both cognition and metacognition is self-confidence. One's self-confidence influences the confidence in his or her knowledge and, consequently, to his or her self-regulation but is also influenced by task-specific confidence judgments in a reciprocal manner. What is interesting, however, is that in children this person characteristic is not only influenced by their knowledge state but also by the support and acceptance they receive from their social environment and particularly from their parents. Therefore, despite the importance of evidence showing the effects of person characteristics on metacognition and vice versa, the next step should be in the direction of theoretical explanations of the mechanism(s) through which social factors intervene to shape metacognition and self-regulation.

2.5 Metacognition and Self-regulation

Metacognition and self-regulation share two basic functions, namely monitoring and control. Metacognition represents the monitoring and control of cognition whereas self-regulation the monitoring and control of behavior guided by one's goal(s) (Carver & Scheier, 1998). In the context of self-regulated learning, self-regulation is conceived as a volitional process that involves the use of strategies for the monitoring and control of behavior, cognition, emotions, motivation, and of the environment as well, so that the individuals can achieve their learning goals (Dinsmore, Alexander, & Loughlin, 2008). This conceptual overlap, in the past, rendered metacognition as a purely top-down process, with metacognitive experiences and metacognitive knowledge being the two aspects of the monitoring function that inform the person on task/situation demands as well as on the state of one's cognition and cognitive processing. The control function of metacognition, such as increase of processing time or use of cognitive strategies (Nelson & Narens, 1994) was originally conceived of as being more or less an automatic process that responds to the input coming from the monitoring of one's cognition – for example, judgment of learning and allocation of study time (see Izaute & Bacon, this volume). Respectively, self-regulation is conceived of as a top-down process guided by the person's goals. However, as Koriat, Ma'ayan, and Nussinson (2006) have argued, metacognition is both a bottom-up and a top-down process (see also the agenda-based metacognition; Ariel, Dunlonsky, & Bailey, 2009), and there is an automatic shift from one mode of functioning to the other during self-regulation. Thus, monitoring at a nonconscious level may directly trigger a control process without the person being aware of it (e.g., one slows down reading when he/she is coming across an irregularity in a bottom-up fashion); or, the person's goal (e.g., to do well at the test) may trigger deliberate monitoring of irregularities so that study time is allocated accordingly in a top-down fashion.

The success of self-regulation, however, in both the bottom-up and top-down modes depends on the accuracy of the monitoring process, and on the availability of control processes (e.g., metacognitive/cognitive strategies). Csíkos and Steklács (this volume) tried to remedy the lack of strategy use in reading. However, the accuracy of monitoring one's own knowledge state is not to be taken for granted. The accuracy of monitoring is a major issue in metacognition research which is also addressed in the present book. Indeed, another point of convergence between the chapters of this book concerns the accuracy of metacognitive judgments. The veridicality of metacognition, as measured by calibration indices reflecting the relation of confidence – or estimate of the correctness of one's response – with performance (objective accuracy) was studied by Allwood (this volume) as well as by Vanderswalmen et al. (this volume). The evidence from the studies presented in these two chapters supports the assumption that people calibrate their metacognitive judgments, albeit not perfectly. Calibration is important because of its implications for both metacognitive control processes and self-regulation. For example, overconfidence (see Kruger & Dunning, 1999), which often occurs when people do not have knowledge or skills but are not aware of it, may decrease effort exertion when needed. On the contrary, people who are aware of knowledge constraints may underestimate their skills and exert effort to maintain a high level of performance. In the study by Vanderswalmen et al. (this volume) the students who had the best spelling performance tended to *underestimate* their performance, possibly because they were aware of the complexity of spelling rules, and overestimated the probability of error; consequently they were more observant of spelling errors and better regulators of their spelling behavior. The opposite happened with poor spellers despite the fact that they had metacognitive knowledge of themselves as being not good spellers.

The question, therefore, is whether children and young students calibrate their confidence judgments and whether they are aware of knowledge constraints. Allwood (this volume) showed that in the case of episodic memory children calibrate their confidence judgments; however, if we look at calibration in the school context this might not necessarily be the case because knowledge acquisition and learning is a complex and long-term process. Hence, confidence judgments have to rely on a number of cues that may change over time as new knowledge is gained, including metacognitive knowledge of task as well as of the self and others. Identifying the changes in the calibration of confidence judgments, while knowledge and metaknowledge develop, is a challenge for future research. Growth models can be particularly helpful in this direction.

However, does more accurate monitoring always entail better control? For example, schizophrenics could adjust their judgments of learning to task demands (Izaute & Bacon, this volume), that is, they demonstrated discrimination, which is another indicator of the accuracy of metacognition (Allwood, this volume). Yet, as Izaute and Bacon (this volume) found, schizophrenics could not regulate time allocation accordingly. This is possibly evidence of dissociation between monitoring and control processes that can be attributed to the pathology of the specific disease. The question is if such dissociation between monitoring and control can also occur in healthy individuals; for instance one may possess metacognitive knowledge of strategies but not use these strategies when working on a task. People may report use of strategies because they believe they are relevant but not because they actually used them. Of course, the opposite might be true as well; individuals may use control strategies in an automatized way without being aware of doing so. Concluding, the dissociation between monitoring and control is another challenge for future research in metacognition.

3 Theoretical Integration and Prospects

The discussion of the convergences and divergences of the various chapters of this book has made it clear that there are issues in metacognition research that cannot be fully accommodated by the prevalent model of metacognition as a meta-level representation of the object level (Nelson, 1996). Specifically, if we accept that there are nonconscious monitoring processes whose outcomes reach conscious awareness, then we should also accept that there is monitoring and control within the object level that do not reach conscious awareness (see also Efklides, 2008). Uncertainty monitoring or monitoring of interruption as opposed to fluency are examples of processes whose outcomes may reach awareness when there is need for conscious/controlled mode of processing (Touroutoglou & Efklides, this volume); familiarity monitoring is also important because it informs on novel situations in which familiarity does not suffice for triggering a response (Scott & Dienes, this volume). In other words, there is monitoring of "good functioning" (Frijda, 1986) and when there is deviation from it, then the information is reaching conscious awareness in the form of metacognitive feelings or affective response (e.g., surprise; Touroutoglou & Efklides, this volume). What is crucial in this case is that the outcomes of such nonconscious monitoring processes, that is, fringe consciousness or metacognitive feelings, are nonanalytic in nature although they can be the object of conscious analytic processes (Koriat, 2007; Koriat & Levy-Sadot, 1999). This implies that the meta-level, in the form of personal awareness (Efklides, 2008), hosts components that are nonanalytic in nature as well as analytic products of attention and metacognitive knowledge.

Moreover, at least some of the metacognitive feelings have an affective character (Efklides, 2001, 2006; Winkielman & Cacioppo, 2001) that informs, along with emotions (e.g., surprise; Touroutoglou & Efklides, this volume), on the positive or negative character of the task/situation. This affectivity triggers control processes such as effort regulation or persistence on (or quitting of) cognitive processing. From the point of view of metacognition, involving affect in self-regulation implies that there is no necessary link between metacognitive monitoring and control processes and that a possible mediator of the effects of monitoring on control is affect.

When we come to instruction of metacognitive processes, as in the study by Csíkos and Steklács (this volume), the question is how can one be aware of another's metacognition so that he or she can successfully direct the co- or other-regulation of the other person's cognition. Communication, perception of the others' cognitive and affective reactions, as well as theory of mind, are at the core of co- and other-regulation of cognition. This kind of metacognition, however, that has as its object the cognition of others could represent a meta-metacognition level, that is, a social level of metacognition (Efklides, 2008). The social level of metacognition builds on the level of personal awareness as well as on a representation of cognition which is primarily based on explicit, conscious processes that are mediated by shared metacognition (i.e., beliefs and knowledge people share with each other). Hence, teachers, parents, or peers who, for example, wish to regulate a child's cognition have to follow the route of the social level of metacognition to influence the level of the child's personal awareness and, through it, cognition. Or, they may influence the child's cognition by providing, for example, learning environments that will have an impact on the subjective experiences of the child and through them on his or her metacognition.

The above theoretical framework, besides being integrative, opens up prospects that go beyond what the discussion of convergences and divergences has already brought to the fore. Social metacognition and co-regulation of behavior are domains that deserve the attention of future research on metacognition, and so does the relationship of metacognition with affect in the self-regulation process. New phenomena in metacognition, such as feeling of difficulty, the blank-in-the-mind experience, self-confidence, or metacognitive knowledge of cognitive processes (such as decision making or spelling) are also topics that broaden the range of research on metacognition, along with new methodological tools which allow a fresh look at old questions. Finally, basic research on the processes that give rise to metacognition is, and should be, at the core of current and future research on metacognition and self-regulation. The present book hopefully serves the acknowledgement of some of the prevalent trends in current research and some of the prospects for the future.

References

Ariel, R., Dunlonsky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: When agendas override item-based monitoring. *Journal of Experimental Psychology: General, 138,* 432–447.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. Cambridge, UK: Cambridge University Press.

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review, 20,* 391–408.

Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation, and self-regulation. In A. Efklides, J. Kuhl, & R. M. Sorrentino (Eds.), *Trends and prospects in motivation research* (pp. 297–323). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2002). The systemic nature of metacognitive experiences: Feelings, judgments, and their interrelations. In M. Izaute, P. Chambres, & P.-J. Marescaux (Eds.), *Metacognition: Process, function, and use* (pp. 19–34). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*(4), 277–287.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–325). Hillsddale, NJ: Erlbaum.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34*, 906–911.

Frijda, N. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Holt.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289–325). Cambridge, UK: Cambridge University Press.

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 483–502). New York: Guilford.

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General, 135,* 36–69.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77,* 1121–1134.

Mangan, B. (1993). Some philosophical and empirical implications of the fringe. *Consciousness and Cognition, 2,* 142–154.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102–116.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: The MIT Press.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation leads to positive affect. *Journal of Personality and Social Psychology, 81,* 989–1000.

Yzerbyt, V. Y., Lories, G., & Dardenne, B. (Eds.). (1998). *Metacognition: Cognitive and social dimensions*. London: Sage.

Zimmerman, B. J. (2004). Sociocultural influence and students' development of academic self-regulation: A social-cognitive perspective. In D. M. McInerney & S. Van Etten (Eds.), *Research on sociocultural influences on motivation and learning: Vol. 4. Big theories revisited*. Greenwich, CT: Information Age.

Part I
Basic Research in Metacognition

Chapter 2
Metacognition in Nonhumans: Methodological and Theoretical Issues in Uncertainty Monitoring

Michael J. Beran, Justin J. Couchman, Mariana V. C. Coutinho, Joseph Boomer, and J. David Smith

1 Introduction

Humans usually know when they do not know. This occurs in a variety of contexts. We know immediately that we cannot give good directions to a location when asked, because we do not know how to get there. We know we cannot offer guidance on the right amount of butter to add to a recipe. We know when we cannot tell whether our automobile will fit into a small parking spot, and we know when we cannot tell which hole has been punched out of a ballot. Thus, human decision-making often is guided by our certainty (or uncertainty) about the accuracy of our own thought processes. In essence, we recognize that sometimes the right decision choice is not to choose, or to slow down our response until we are certain, or to ask for more information. Our judgments are made on the basis of how confident we are that our information processing routines have provided enough information or the correct information for an accurate response. This ability, called metacognition, is sometimes defined as "thinking about thinking", but it also refers to the monitoring of other, more basic "first-order" cognitive processes (Flavell, 1979). Thus, metacognition is a mental process that takes the results of a first-order process such as perception or memory activation and operates on the product of that first-order process for some second-order judgment (Proust, 2007).

Human metacognition is intricately linked to important aspects of mind, including cognitive control, self-awareness, and consciousness. For this reason, it is often held as one of humans' most sophisticated cognitive capacities, and it is widely accepted that humans are capable of metacognitive processing. Whenever humans reflect on what they know, re-evaluate their thought processes, or seek additional information, they are demonstrating their capacity for metacognition (Benjamin, Bjork, & Schwartz, 1998; Dunlosky & Nelson, 1992; Flavell, 1979; Koriat, 1993, 2007, 2008; Metcalfe & Shimamura, 1994; Nelson, 1992; Schwartz, 1994;

M.J. Beran (✉)
Language Research Center, Georgia State University, Atlanta, GA 30302-3965, USA
e-mail: mjberan@yahoo.com

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_2,

Serra & Dunlosky, 2005). As other contributors to this volume have demonstrated, human metacognition occurs in diverse situations and supports intelligent behavior.

Our contribution to this volume is to promote consideration of whether animals other than humans show evidence of metacognition. Historically, *Homo sapiens* alone was regarded as metacognitive, as other animals were considered to have little by way of mental lives, and they were considered much more bound in their behavior to the stimuli that they encountered and the outcomes that they experienced when they behaved (e.g., Morgan, 1906). Because metacognition is singled out as one of humans' most sophisticated cognitive capacities (Tulving, 1994; also see Metcalfe & Kober, 2005; Proust, 2007), the bar is set high for any potential demonstration of nonhuman animal metacognition. A heavy burden is placed on those studying nonhuman animals if claims of animal behavior are to be considered the result of metacognitive processes. However, these burdens also illustrate what the implications might be, should other animals be afforded recognition of their metacognitive abilities. It might require new perspectives on higher-order processes such as self-awareness (Gallup, 1982; Gallup, Povinelli, & Suarez, 1995; Humphrey, 1976; Parker, Mitchell, & Boccia, 1994), theory of mind (Byrne & Whiten, 1988; Smith, Shields, & Washburn, 2003), and consciousness (Koriat, 2007; Nelson, 1996). These implications excite those of us who are exploring the capability for metacognitive capacities that nonhuman animals can demonstrate.

Both proponents and critics of animal metacognition research agree that efforts to discern such abilities in animals are important for what they might illustrate about animal cognition and about the emergence of metacognition and other higher-order faculties in humans. However, a strong debate exists as to how convincing a case can be made for animal metacognition. Numerous research teams have provided data for this debate (Beran, Smith, Redford, & Washburn, 2006; Call & Carpenter, 2001; Foote & Crystal, 2007; Hampton, 2001; Inman & Shettleworth, 1999; Kornell, Son, & Terrace, 2007; Smith et al., 2003; Sutton & Shettleworth, 2008; Washburn, Smith, & Shields, 2006). Here, we review the basic methods that have been used to evaluate metacognition in nonhuman animals, and the results from those studies. Then, we focus on the method we have used extensively with monkeys – the uncertainty monitoring paradigm. We will highlight one of the major criticisms of the paradigm regarding the role of stimulus properties in the performance patterns that emerge. We will outline recent evidence that counters this criticism and other theoretical objections being raised against the possibility that monkeys' performances reflect metacognitive abilities.

2 Paradigms for Testing Animal Metacognition

2.1 Information Seeking

One way to assess what animals may or may not know about their own knowledge states involves letting animals choose whether to seek more information or not. For example, Call and Carpenter (2001) assessed chimpanzees' and orangutans' knowledge about their own visual perception. Specifically, they assessed whether the apes could seek more information when it was needed. The apes were presented with tubes into which they could reach and obtain any hidden food items. Sometimes, the location of the hidden item was seen by the animal, but sometimes it was hidden out of sight. In the latter case, the researchers suggested that the apes would be well served to look into the tubes before reaching, and in fact the apes did exactly this.

This paradigm is easy to adapt for use with other species, and recent efforts have shown some important differences across species in performance on this task. Rhesus monkeys, for example, behave much like apes in that they search more often for additional information about hidden items when they did not view the original hiding event compared to when they had (Hampton, Zivin, & Murray, 2004). However, capuchin monkeys do not show this pattern, suggesting more limited information seeking (and, perhaps, metacognitive ability) in this species (Paukner, Anderson, & Fujita, 2006).

2.2 *Retrospective Confidence Judgments*

Another approach has been to let nonhuman animals provide retrospective confidence judgments, or reports about how certain they are about an already given answer. This method, when given to humans, seems to offer a clear indication of metacognition in a variety of domains (e.g., Koriat, 2008; Koriat, Lichtenstein, & Fischoff, 1980; Nelson & Narens, 1990). Although only a few attempts have been made to use this method with nonhuman animals, the results indicate that animals too can rate the likelihood of having given a correct response. Shields, Smith, Guttmannova, and Washburn (2005) reported that rhesus monkeys used a confidence-rating scale to judge whether a response they had just made in a psychophysical discrimination was likely to be a correct response or not. Monkeys classified stimuli as “sparse” or “dense” without feedback and then made a confidence judgment by choosing different levels of timeout or reward to accompany the outcome of their primary discrimination response. Monkeys appeared to use these confidence responses appropriately (and similarly to how humans used them) as they risked longer timeouts to potentially gain larger reward amounts on trials in which they were more likely to be correct.

Son and Kornell (2005) also gave monkeys a confidence-judgment task. Their task included a psychophysical judgment in which nine lines were presented on a screen and the monkeys learned to select the longest line. This task also included a retrospective confidence judgment as the monkeys subsequently had to choose between a high-risk bet (two tokens gained or lost) or a low-risk bet (one token gained no matter whether the chosen line was longest or not). Monkeys chose the high-risk bet more often on trials that were objectively easier and on which they performed more accurately; they even transferred this performance pattern to new kinds of discrimination (Kornell et al., 2007).

2.3 Uncertainty Responses

Paradigms using variations of an uncertainty response are the most prevalent in animal metacognition research. In all cases, the uncertainty response acts as an alternative to some primary response that an animal has been trained to use. The primary tasks themselves can take many forms, and almost any task with a continuum of easy to difficult trials can incorporate an uncertainty response. To date, the uncertainty response has been included in tests of psychophysical discriminations in the visual, auditory, and temporal domains (Foote & Crystal, 2007; Smith, Beran, Redford, & Washburn, 2006; Smith et al., 1995; Smith, Shields, Allendoerfer, & Washburn, 1997; Smith, Shields, Schull, & Washburn, 1997), tests of list memory, item memory, and spatial memory (Hampton, 2001; Suda-King, 2008), judgments of quantity (Beran et al., 2006), judgments of sameness and difference (Shields, Smith, & Washburn, 1997), and tests for two-choice discrimination learning (Washburn et al., 2006).

In the first test using an uncertainty response, a dolphin performed an auditory discrimination (Smith et al., 1995). The dolphin learned to make a High response when a 2,100 Hz tone sounded or a Low response when the tone was lower than 2,100 Hz. He made Low and High responses appropriately for easy Low trials and true High trials, but he made more errors when the tone approached, but did not reach, 2,100 Hz.

The dolphin then was presented with a third response that allowed it to avoid making either of the primary responses. The dolphin used that response most often in response to those tones that were most difficult to discriminate from the 2,100 Hz tone. This result subsequently was replicated in other tasks and species. For example, in the psychophysical domain, rhesus monkeys judged the pixel density of rectangles presented on a computer monitor along a continuum from most sparse to most dense and used the uncertainty response most often for those trials on which they were most likely to make an error (Smith et al., 1997). The monkeys' performance closely mirrored that of humans given the same task (see Fig. 2.1). Foote and Crystal (2007) presented rats with a psychophysical temporal discrimination in which an uncertainty response also was available. Rats judged whether the duration of a noise was closer to the short category (2 s) or long category (8 s), with large food rewards for correct responses. Rats also could make an uncertainty response that allowed them to decline the trial and receive a smaller food reward instead. Difficulty was varied so that there were easy (2 or 8 s) and hard (e.g., 3.62 s) trials, and the rats chose the uncertainty response more often for the more difficult trials.

Hampton (2001) gave rhesus monkeys a memory test that included an uncertainty response. Monkeys had to remember a presented stimulus for some delay. In one condition, they next had to take the memory test. In another condition, they could choose to take the memory test or they could choose to avoid the test and move on to the next trial. Monkeys were more accurate when they chose to complete the memory test than when they were forced to complete the test. Hampton (2001) suggested that this result indicated that the monkeys knew when they remembered the correct response, and they declined trials when they were uncertain as to the correct

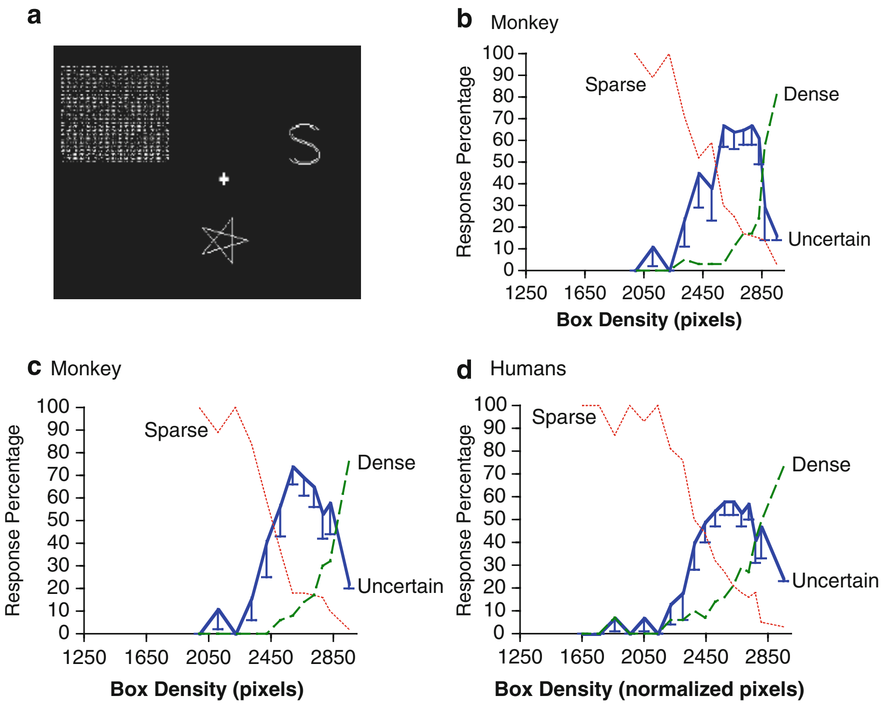


Fig. 2.1 (**a**) A trial in the Sparse-Dense discrimination of Smith et al. (1997). (**b**–**c**) The performance of two monkeys in this task. The "dense" response was correct for boxes with exactly 2,950 pixels – those trials are represented by the rightmost data point for each curve. All other boxes deserved the "sparse" response. The pixel-density of the box is shown on the horizontal axis, and the solid line represents the percentage of trials receiving the uncertainty response at each density level. The error bars show the lower 95% confidence limits. The percentages of trials ending with the "dense" response or "sparse" response are also shown. (**d**) The performance of seven humans performing the same discrimination given to the monkeys. The *panels* in **a**, **b**, and **c** are reprinted from Smith, J. D., Shields, W. E., Schull, J., & Washburn, D. A. (1997). The uncertain response in humans and animals. *Cognition, 62*, 75–97, with permission from Elsevier. The *panel* in **d** is reprinted from Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *The Behavioral and Brain Sciences, 26*, 317–373. Copyright 2003 by Cambridge University Press. Reprinted with permission

response. If true, this would be a demonstration of metamemory and provide another similarity between nonhuman and human behavior (Benjamin et al., 1998).

Suda-King (2008) also reported that orangutans monitored their memory for spatial locations of hidden food items. Items were hidden in clear view or were hidden out of view of the observing orangutan. When hidden in view, the orangutans typically chose a location. However, when the items were hidden out of view, the orangutans were more likely to avoid choosing a location and instead chose a smaller, but guaranteed, food reward. These data match other reports in which apes show equivocation in responding when information is sparse or discriminations are very difficult (Suda & Call, 2006).

3 Problems (and Challenges) for Testing Animal Metacognition Using the Uncertainty Paradigm

As evidence has accumulated to force the consideration of metacognition in animals, so too have criticisms of the paradigms that have been used. These challenges have been productive for the field, and they are outlined in great detail elsewhere (Carruthers, 2008; Crystal & Foote, 2009; Hampton, 2009; Jozefowiez, Staddon, & Cerutti, 2009; Proust, 2007; Smith, Beran, Couchman, & Coutinho, 2008; Smith, Beran, Couchman, Coutinho, & Boomer, 2009; Staddon, Jozefowiez, & Cerutti, 2007). One of the main challenges reflects the need for more refined methods of inquiry, and that is a positive direction for the field. Here, we address perhaps the most prevalent criticism for the uncertainty monitoring paradigm, and particularly its use with monkeys; specifically, uncertainty responses are not reflective of monitored confidence or uncertainty but instead are learned responses to specific stimuli or specific kinds of stimuli.

This criticism has two bases. The first is that objective stimulus cues might occasion uncertainty responses through traditional mechanisms of learning. The second is that transparent reinforcement necessarily makes some stimuli more or less preferred compared to others, and animals could use their preference or aversion to stimuli to guide their use of the uncertainty response rather than their uncertainty as to how to classify those stimuli.

How would these methodological characteristics account for the frequently reported ideal and efficient use of uncertainty responses by animals without the need to appeal to uncertainty experienced by those animals? Consider the “dense” and “sparse” responses by monkeys in the Smith et al. (1997) study. Those monkeys were trained to focus on the primary stimulus quality (density) over many thousands of trials, with only so many possible density levels that could be presented. This made it quite likely that specific, first-order stimulus properties could generate uncertainty responses. Formal models (Smith et al., 2008) even indicate that this is likely if monkeys are assumed to develop aversion and avoidance reactions to first-order stimuli. What occasions this aversion or avoidance? The use of a transparent schedule of punishment or reinforcement does. Again considering the Dense-Sparse paradigm, after each primary response (“dense” or “sparse”), there is immediate feedback in the form of timeout punishment or food reward. Very quickly, monkeys could learn to associate certain density levels with richer or leaner reward, and thus develop aversion or preference for responses made in the presence of those stimuli. Where aversion was felt, the uncertainty response could operate to alleviate that aversion (particularly if it offered some tangible positive outcome such as a smaller food reward), and it would necessarily be used most often for those trial levels that were most averse (and objectively most difficult for the monkeys).

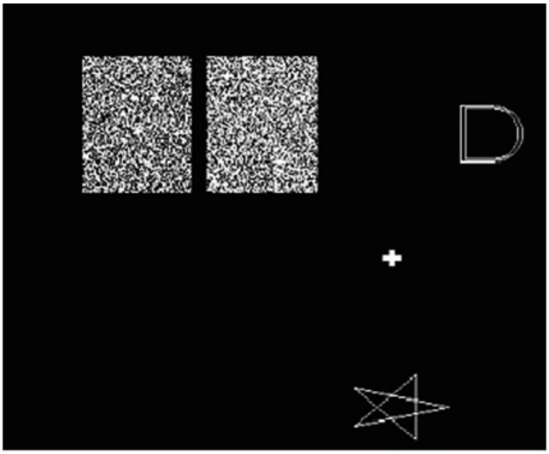
To overcome these concerns, one needs to either avoid the repeated presentation of stimuli or prevent the association of certain levels of aversion or preference with repeatedly presented stimuli. Recent investigations using the uncertainty monitoring paradigm with monkeys have done this, and we outline some of those investigations below.

3.1 Avoiding Stimulus Cues

Perhaps the best way to prevent objective stimulus cues from occurring is to ensure that tasks lift the variable to be judged or compared off the plane of the primary stimuli, so that judgments are about abstract relations between stimuli, or about remembered traces of stimuli. To achieve this, animals must be trained to make relative judgments or relational judgments of stimuli. For example, Shields et al. (1997) used the Same-Different (SD) task so that monkeys would be required to judge the relation between two stimuli rather than the stimuli themselves in isolation. Each trial contained two rectangles filled with lit pixels. Animals made "same" or "different" responses to pairs of rectangles that had the same or different density. To cause animals difficulty, the size of the difference on Different trials was adjusted to constantly challenge participants' discrimination abilities. Moreover, "same" and "different" trials at several absolute density levels were intermixed to ensure a true relational performance. Monkeys used the uncertainty response appropriately – to decline trials near their discrimination threshold – even though now the displays had to be judged relationally for their sameness or difference (see Fig. 2.2). In this case, uncertainty responses could not have been triggered by low-level stimulus cues because the relevant cue was abstract-relational in nature.

Smith et al. (1998) asked whether monkeys would be able to show memory monitoring during serial list retention tests. They presented monkeys and humans with a serial-probe recognition task in which a trial-unique list of items was presented in serial order followed by a probe item. One response was made if the probe was in the list, and another was made if it was not. Typically, items in the middle of lists are hardest to remember, and monkeys showed this pattern. They also showed the highest level of uncertainty responses for exactly those same serial positions. The studies by Smith et al. (1998) and Hampton (2001) both carefully controlled the presentation of all stimuli so that each stimulus served as either memory probe or foil with the same frequency. All stimuli were rewarded and nonrewarded following both primary responses in the same way. No stimulus cue ever indicated any response. The only thing that was relevant was whether a stimulus was a to-be-remembered item or not. Thus, these animals showed metamemory – they were monitoring the contents of memory to determine their response, but they were not relying on specific stimulus cues to determine that response.

Washburn et al. (2006) added an uncertainty response to Harlow's (1949) learning-set paradigm. Monkeys completed blocks of trials in which they made successive two-choice discriminations. The first trial of each new problem presented a difficult choice for the monkeys, because they could not know which stimulus was the S+. Although they could make a random response on Trial 1 and then have the information needed for the remainder of the problem, they also were offered an uncertainty response that provided a hint as to the correct response on Trial 1. The monkeys used this uncertainty response most often on the first trial of each new problem compared to other trial numbers within problems. The first trial within each problem always consisted of novel stimuli, and so there was no chance that monkeys were conditioned



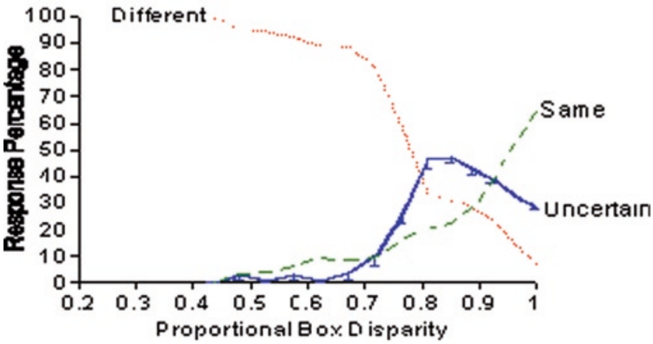


Fig. 2.2 (*Top*) The Same-Different discrimination task used by Shields et al. (1997). Monkeys had to determine whether two rectangles contained the same number of pixels or not. "Same" responses were made by moving the cursor into contact with the rectangles. "Different" responses were made by moving the cursor into contact with the "D." The Star at the bottom of the screen was the uncertainty response. (*Bottom*) Two monkeys used the uncertainty response most often on those trials for which they were equally likely to classify stimuli as being same or different. This figure is from Shields, W. E., Smith, J. D., & Washburn, D. A. (1997). Uncertain responses by humans and rhesus monkeys (Macaca mulatta) in a psychophysical same-different task. *Journal of Experimental Psychology: General, 126*, 147–164. Reprinted with permission from the American Psychological Association

to make uncertainty responses to specific stimuli. Monkeys also showed this pattern of choosing the uncertainty response more often during the first trial of each problem from the very first discrimination problems that they completed in the experiment. Thus, they did not learn to make uncertainty responses when new stimuli appeared, but instead they responded as if they recognized immediately that new problems were difficult on Trial 1, and therefore uncertainty responses were appropriate on those trials. A possible criticism of this result is that this recognition may have been the result of some sense of familiarity or unfamiliarity that the monkeys experienced (as with old or new stimuli), with monkeys making uncertainty responses when they experienced unfamiliarity with presented stimuli. If true, uncertainty responses would not have been based on specific features of stimuli that monkeys remembered but on the familiarity of stimuli. Whether this familiarity refers only to a first-order state (knowing that presented stimuli are old or new) or refers to a second-order state whereby monkeys are aware of the familiarity of stimuli is undetermined (see Chap. 3, this volume, for a more in-depth discussion of the role of familiarity in decision-making and metacognition).

Beran et al. (2006) showed that monkeys could make uncertainty responses adaptively when they faced difficult and uncertain numerical discriminations. Monkeys were shown dot quantities and had to classify those quantities as being "large" or "small" according to some arbitrary midpoint. For values closer to that midpoint, performance was lower than for values more discrepant from the midpoint, and the monkeys' use of the uncertainty response mirrored their performance in categorizing the quantities as large or small. The changing center values across days meant that given quantities varied as to their closeness to the center value, making their difficulty variable as well. Thus, on some days, a particular quantity would be difficult to classify as large or small because it was near the central value, and monkeys were likely to use the uncertainty response for that quantity. On other days, the same quantity was easier to classify as large or small because it was farther from the central value, and in that case, monkeys were unlikely to use the uncertainty response when presented with that quantity.

3.2 Making Reinforcement Opaque

Smith, Beran, Redford, and Washburn (2006) made the reward and punishment structure of a Dense-Sparse task opaque to rhesus monkeys. Monkeys first were trained to make the primary ("dense" or "sparse") response to pixilated stimuli. Then, the uncertainty response was offered, and it operated only to remove a trial from the screen, without providing any food reward, hint, or easier next trial. Monkeys also learned to tolerate deferred feedback. They would complete blocks of four trials, with each trial randomly chosen from the continuum from most sparse to most dense. At the end of the four trials, feedback was provided as a summary for performance in the entire block. All rewards for correct responses were delivered first, followed by all timeouts for incorrect responses. Trials in which the uncertainty response was chosen were not factored into the feedback, and in essence they operated to eliminate that trial from the scoring procedure. In this way, monkeys could not directly associate any response with specific feedback. Rather, they only received summary feedback regarding their overall performance (as reflected in the ratio of rewards to timeouts). This feedback routine is analogous to one in which humans take a multiple-choice test in which only those questions that are answered are scored, and students only receive an overall grade without feedback on the correctness of any given answer to a question.

Smith et al. (2006) asked whether monkeys might organize their response patterns appropriately given this opaque feedback structure. In this case, however, appropriate use of the uncertainty response meant something new. In previous experiments, with trial-by-trial feedback, appropriate use of the uncertainty response meant that it would mirror performance on the psychophysical judgment. In the Smith et al. (2006) study, however, this did not have to be true. Performance and uncertainty response frequency could be dissociated and still potentially reflect metacognitive processes, provided that the use of the uncertainty response instead mirrored the monkeys' own subjective structuring of the primary task. Remember that in this task, the monkeys could not know whether their responses were correct or not (on a trial-by-trial basis). All they could learn was the general structure of the task (i.e., that stimuli are sparse to dense) and then respond based on their impression of a specific stimulus as belonging to one of the stimulus categories. This necessarily meant that perceptual error could not be "corrected" by feedback that could be associated to specific stimuli. Thus, monkeys could come to structure the task in a subjective manner that did not necessarily reflect its objective structure. In a two-choice perceptual discrimination task, this would occur, for instance, if the monkeys came to respond as if the midpoint of the continuum was different than that afforded in the task parameters itself (i.e., monkeys mistakenly believed that some stimuli were dense when in fact they were sparse, or vice versa). One monkey, Murph, in the Smith et al. (2006) study provided just this evidence (Fig. 2.3, panels a, c).

For Murph there was no relationship between trial failure and trials declined (Fig. 2.3, panel b). Based on an associative hypothesis, the trial-decline rates should have been highest for the bins with the lowest proportion correct. The animal's uncertainty responses did not follow the task's reinforcement and associative patterning. For example, the monkey was equally likely to use the uncertainty responses on trials from stimulus levels in which he was 93% correct in making the primary response as he was on trials from stimulus levels in which he was only 25% correct (Fig. 2.3, panel a). His primary performance (choosing "dense" or "sparse") was radically different, but he declined trials at a similar rate. If Murph had tracked the reinforcement for responses made on specific stimulus levels he would have learned to avoid primary responses on those trials for which he was performing badly, and he rarely would have used the uncertainty response on those trials in which he was performing very well. But, he did not do this, confirming that he was not responding to associative cues. Rather, he was using the uncertainty response according to the task construal that he had established *despite* the feedback he was receiving. Where his subjective task construal indicated that the two discrimination responses were equally feasible is where his uncertainty response peaked (Fig. 2.3, panels c, d).

3.3 Answering the Criticism

Despite these outcomes from methods that make reinforcement opaque (as it pertains to individual responses) and prevent stimulus cues from occasioning uncertainty responses, we acknowledge that critics may not be fully reassured. Even in rearranging the order of feedback, monkeys still could associate their patterns of responding with general levels of reward and punishment. For example, after

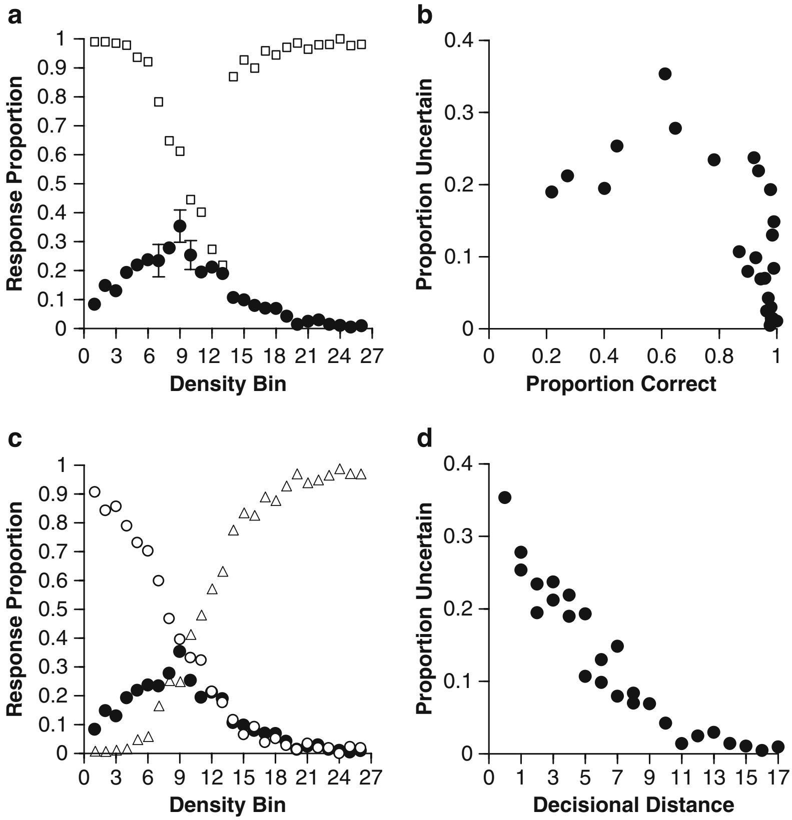


Fig. 2.3 (**a**) The performance of monkey Murph in the deferred-reinforcement uncertainty-monitoring task used by Smith et al. (2006). The horizontal axis indicates the density bin of the box. The "sparse" and "dense" responses, respectively, were correct for boxes at Density Bins 1–13 and 14–26. The open squares show the proportion of trials attempted that were answered correctly. The *dark circles* show the proportion of trials receiving the uncertainty response at each density bin. Representative 95% confidence intervals are shown for the peak of uncertainty responding near the task's midpoint and for the first bins to the right and left of that peak in which uncertainty responding was significantly reduced relative to the peak. (**b**) Murph's performance in the same task, with the proportion of trials declined in each density bin plotted against the proportion correct for that bin. There was no relation. (**c**) Murph's performance showing separately his use of the "sparse" response (*open circles*), "dense" response (*open triangles*), and uncertainty response (*dark circles*). (**d**) Murph's performance in this task with his proportion of trials declined in each density bin plotted against the decisional distance of the bin from his decisional breakpoint. This figure is from Smith, J. D., Beran, M. J., Redford, J. S., & Washburn, D. A. (2006). Dissociating uncertainty states and reinforcement signals in the comparative study of metacognition. *Journal of Experimental Psychology: General, 135*, 282–297. Reprinted with permission from the American Psychological Association

choosing the "dense" response most often in a block that included stimuli near the perceptual threshold, and then receiving only moderate reward levels, the animals might come to make different patterns of responses to the same kinds of stimuli in the future. Here, outcomes do influence future choices. In the case of trial unique stimuli or relational judgments among stimuli, one could argue that memory trace strength or concepts themselves operate as stimuli to be associated with outcome frequencies. But, how is any behavior (human or nonhuman) ever free from those constraints? We argue that it is not.

No organism behaves in a vacuum. The goal of behavior is to better match the organism to its environment, and that requires some form of monitoring whether only first-order in the form of Stimulus-Response-Outcome associations or second-order in the form of monitoring first-order processes so as to generate a second-order decision or judgment (Proust, 2007). Our point is that the level of the cognitive performance to which the uncertainty response attaches is critical in evaluating the cognitive level of the uncertainty response itself. Tasks that involve trial-unique stimuli, or responses to concepts such as number or sameness-difference, or the availability of a memory are different in psychological character from those that allow clear external stimulus control. In our view, responding to the indeterminacy of a conceptual relation or responding to the dimness of a memory are profoundly different from responding to a present, aversive stimulus. Tasks that require learning how groups of responses, presumably organized on the basis of the stimuli to which they are given, lead to overall levels of reward are different in psychological character to those that give immediate, clear feedback as to the correctness of a specific response to a specific stimulus (or stimulus class). It is important to see that animals reliably produce adaptive uncertainty-monitoring performances in many tasks that form a broader pattern and that transcend the criticisms one may have regarding any paradigm in particular. At some point, the most parsimonious claim becomes that animals have a general, uncertainty-monitoring mechanism that serves them well in many tasks and in many domains.

4 Summary

Given the speed with which the topic of animal metacognition has emerged and entrenched itself within the domain of animal cognition, we expect much progress in coming years. We believe that stimulus-based or reinforcement-based explanations of uncertainty responding no longer hold in all cases. Animals respond to more than just the stimulus in front of them. They respond to their own impressions of those stimuli, and to the certainty or uncertainty they feel about those impressions vis-à-vis the animals' subjective structuring of tasks. New data will be required to engage the remaining challenges in this field if we are to understand the nature and extent of animal metacognitive abilities. We anticipate the development of new techniques that will allow animals to have even better ways to report their uncertainty more broadly and more flexibly. We believe that future tests might demonstrate that animals, like humans, can use generalized, flexible, spontaneous uncertainty responses, and that such data will inform the debate regarding the internal states that provoke those responses and the role that metacognition plays in other mental states such as consciousness. This, to us, offers an exciting future and a better understanding of all minds, human and nonhuman.

Acknowledgements Many people have contributed to this research program, and we offer special thanks to David Washburn, Wendy Shields, Joshua Redford, Barbara Church, John Gulledge, Ted Evans, and Mary Beran. This research is supported by Grant BCS-0634662 from the National Science Foundation and Grants HD-38051 and HD 061455 from the National Institute of Child Health and Human Development. Correspondence can be sent to the first author at mjberan@yahoo.com.

References

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metacognitive index. *Journal of Experimental Psychology: General, 127*, 55–68.

Beran, M. J., Smith, J. D., Redford, J. S., & Washburn, D. A. (2006). Rhesus macaques (Macaca mulatta) monitor uncertainty during numerosity judgments. *Journal of Experimental Psychology: Animal Behavior Processes, 32*, 111–119.

Byrne, R. W., & Whiten, A. (1988). *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans.* New York: Oxford University Press.

Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition, 4*, 207–220.

Carruthers, P. (2008). Meta-cognition in animals: A skeptical look. *Mind and Language, 23*, 58–89.

Crystal, J. & Foote, A., (2009). Metacognition in animals. *Comparative Cognition and Behavioral Reviews, 4*, 1–16.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed JOL effect. *Memory & Cognition, 20*, 374–380.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911.

Foote, A., & Crystal, J. (2007). Metacognition in the rat. *Current Biology, 17*, 551–555.

Gallup, G. G. (1982). Self-awareness and the emergence of mind in primates. *American Journal of Primatology, 2*, 237–248.

Gallup, G. G., Povinelli, D. J., & Suarez, S. D. (1995). Further reflections on self-recognition in primates. *Animal Behaviour, 50*, 1525–1532.

Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences, 98*(9), 5359–5362.

Hampton, R. R. (2009). Multiple demonstrations of metacognition in nonhumans: Converging evidence or multiple mechanisms? *Comparative Cognition and Behavioral Reviews, 4*, 17–28.

Hampton, R. R., Zivin, A., & Murray, E. A. (2004). Rhesus monkeys (Macaca mulatta) discriminate between knowing and not knowing and collect information as needed before acting. *Animal Cognition, 7*, 239–246.

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review, 56*, 51–65.

Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Cambridge, UK: Cambridge University Press.

Inman, A., & Shettleworth, S. J. (1999). Detecting metamemory in nonverbal subjects: A test with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 25*, 389–395.

Jozefowiez, J., Staddon, J. E. R., & Cerutti, D. (2009). Metacognition in animals: How do we know that they know? *Comparative Cognition and Behavioral Reviews, 4*, 29–39.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*, 609–639.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289–325). Cambridge, UK: Cambridge University Press.

Koriat, A. (2008). Subjective confidence in one's answers: The consensuality principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 945–959.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 107–118.
Kornell, N., Son, L., & Terrace, H. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science, 18*, 64–71.
Metcalfe, J., & Kober, H. (2005). Self-reflective consciousness and the projectable self. In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 57–83). New York: Oxford University Press.
Metcalfe, J., & Shimamura, A. (1994). *Metacognition: Knowing about knowing.* Cambridge, MA: Bradford Books.
Morgan, C. L. (1906). *An introduction to comparative psychology.* London: Walter Scott.
Nelson, T. O. (Ed.). (1992). *Metacognition: Core readings.* Toronto, Canada: Allyn and Bacon.
Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102–116.
Nelson. T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation, 26*, 125–141.
Parker, S. T., Mitchell, R. W., & Boccia, M. L. (1994). *Self-awareness in animals and humans: Developmental perspectives.* New York: Cambridge University Press.
Paukner, A., Anderson, J. R., & Fujita, K. (2006). Redundant food searches by capuchin monkeys (Cebus apella): A failure of metacognition? *Animal Cognition, 9*, 110–117.
Proust, J. (2007). Metacognition and metarepresentation: Is a self-directed theory of mind a precondition for metacognition? *Synthese, 2*, 271–295.
Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin and Review, 1*, 357–375.
Serra, M. J., & Dunlosky, J. (2005). Does retrieval fluency contribute to the underconfidence-with-practice effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1258–1266.
Shields, W. E., Smith, J. D., & Washburn, D. A. (1997). Uncertain responses by humans and rhesus monkeys (Macaca mulatta) in a psychophysical same-different task. *Journal of Experimental Psychology: General, 126*, 147–164.
Shields, W. E., Smith, J. D., Guttmannova, K., & Washburn, D. A. (2005). Confidence judgments by humans and rhesus monkeys. *Journal of General Psychology, 132*, 165–186.
Smith, J. D., Beran, M. J., Couchman, J. C., Coutinho, M. V. C., & Boomer, J. (2009). Animal metacognition: Problems and prospects. *Comparative Cognition and Behavioral Reviews, 4*, 40–53.
Smith, J. D., Beran, M. J., Couchman, J. C., & Coutinho, M. V. C. (2008). The comparative study of metacognition: Sharper paradigms, safer inferences. *Psychonomic Bulletin & Review, 15*, 679–691.
Smith, J. D., Beran, M. J., Redford, J. S., & Washburn, D. A. (2006). Dissociating uncertainty states and reinforcement signals in the comparative study of metacognition. *Journal of Experimental Psychology: General, 135*, 282–297.
Smith, J. D., Schull, J., Strote, J., McGee, K., Egnor, R., & Erb, L. (1995). The uncertain response in the bottlenosed dolphin (Tursiops truncatus). *Journal of Experimental Psychology: General, 124*, 391–408.
Smith, J. D., Shields, W. E., Allendoerfer, K. R., & Washburn, D. A. (1998). Memory monitoring by animals and humans. *Journal of Experimental Psychology: General, 127*, 227–250.
Smith, J. D., Shields, W. E., Schull, J., & Washburn, D. A. (1997). The uncertain response in humans and animals. *Cognition, 62*, 75–97.
Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *The Behavioral and Brain Sciences, 26*, 317–373.
Son, L. K., & Kornell, N. (2005). Metaconfidence judgments in rhesus macaques: Explicit versus implicit mechanisms. In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 296–320). New York: Oxford University Press.
Staddon, J. E. R., Jozefowiez, J., & Cerutti, D. (April 13, 2007). Metacognition: A problem not a process. *PsyCrit*, 1–5.
Suda, C., & Call, J. (2006). What does an intermediate success rate mean? An analysis of a Piagetian liquid conservation task in the great apes. *Cognition, 99*, 53–71.

Suda-King, C. (2008). Do orangutans (Pongo pygmaeus) know when they do not remember? *Animal Cognition, 7*, 239–246.

Sutton, J. E., & Shettleworth, S. J. (2008). Memory without awareness: Pigeons do not show metamemory in delayed matching-to-sample. *Journal of Experimental Psychology: Animal Behavior Processes, 34*, 266–282.

Tulving, E. (1994). Preface. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 6–8). Cambridge, MA: MIT Press.

Washburn, D. A., Smith, J. D., & Shields, W. E. (2006). Rhesus monkeys (Macaca mulatta) immediately generalize the uncertain response. *Journal of Experimental Psychology: Animal Behavior Processes, 32*, 85–89.

Chapter 3
The Metacognitive Role of Familiarity in Artificial Grammar Learning: Transitions from Unconscious to Conscious Knowledge

Ryan Scott and Zoltán Dienes

1 Introduction

We distinguish two fundamentally different ways of learning; one which can occur in the absence of metacognition but may lead to its emergence, the other which is reliant on metacognition from the outset. The first method is shared with animals and new born babies; by this method, one learns using representations whose function is to indicate how the world IS. The second method involves using representations that can indicate possibilities or counterfactuals. According to Perner (1991), the latter representational capacity arises at around 18 months in humans, when we acquire the ability to use "multiple models" of the same object. This capacity makes possible some of the requirements of episodic memory (the same object simultaneously considered as it was and as it is) and of hypothetical reasoning (the same object considered in the different ways it might be).

The first method can be illustrated by the learning in a standard connectionist network (see, e.g., Shanks, 2005a). The input activation pattern functions to indicate what is there; the weights in the network function to indicate the enduring statistical structure in the environment. When the presented stimulus changes, then the input pattern changes; when the relevant statistical structure changes, the asymptotic weights change. Both activation and weight representations function to track how reality is. Perner (1991) called such representations "single updating models", that is, each update over-writes the last. The second method can be illustrated by hypothesis testing. In this case, representations are explicitly marked as possible rather than factual. We can entertain any hypothesis within the limits of our imagination (for example, more than first-order statistical properties) and consider it confirmed in as little as one trial, or never, depending on the limits of our reasoning (and data).

R. Scott (✉)
Department of Psychology, University of Sussex,
Pevensey Building, Falmer, Brighton BN1 9QH, UK
e-mail: r.b.scott@sussex.ac.uk

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_3,

The two methods are related to the distinction between conscious and unconscious knowledge, that is, between knowledge in the presence and absence of metacognition. According to Dienes and Perner (1999), knowledge being conscious requires not only making explicit the possibility/fact distinction we have been discussing, but also making explicit one's attitude of knowing. That is, for a mental state of knowing to be conscious, one should represent that one is in that state (Rosenthal, 1986). However, the correspondence between these dualities (conscious vs. unconscious knowledge and single updating vs. possibility-explicit learning) is not one-to-one. We will investigate their relation in this chapter.

In assessing whether knowledge is conscious or unconscious it is important to distinguish first-order and second-order mental states. A first-order mental state is a state about the world. The state is a conscious state when we are aware of being in it (cf. Carruthers, 2000; Rosenthal, 2005). Thus, the state is conscious when we have another mental state, a second-order state, which asserts we are in the first-order state. Showing that a person can accurately respond to objective properties in the world shows the presence of a first-order state; it shows knowledge but not awareness of knowing. Showing that a person can accurately tell what mental state they are in shows the presence of a second-order state; it shows knowing that one knows and hence conscious knowledge or meta-knowledge. These considerations form the basis of our methodology for determining the conscious status of knowledge states (for more discussion see Dienes, 2008).

The artificial grammar learning paradigm (Reber, 1967) provides a convenient task in which learning using a single updating model or the use of hypotheticals can be explored. Participants are asked to look at or memorise strings of letters. The order of the letters within each string obeys a complex set of rules, but the strings look more-or-less random. After a few minutes exposure participants are told about the presence of rules, but not the rules themselves, and asked to classify new strings as obeying the rules or not. Participants can make these well-formedness judgments reasonably well (typically with about 65% accuracy) even while complaining that they have ruined the experiment because they did not know anything relevant. Such learning has been modelled with various connectionist networks (for a review, see Cleeremans & Dienes, 2008). Reber (1976) has also informed participants of the existence of the rules in the training phase. When deliberately trying to search for rules, participants consider possible rule structures and can acquire conscious knowledge of the structure.

It can be shown that participants merely looking at or memorising strings in the training phase acquire knowledge they do not know they have. Participants can be asked on each classification trial to state whether they guessed or knew to some degree. When participants say they are guessing, classification accuracy is typically above baseline, showing unconscious knowledge by the *guessing criterion* (Dienes, Altmann, Kwan, & Goode, 1995). Dienes et al. (1995) argued the guessing criterion was useful in identifying two different methods of learning. Plausibly the weights of a connectionist network can update without us knowing we are learning; and predictions can be made on the basis of those weights without us being aware of having knowledge. Conversely, most of the time when we make possibilities explicit we are prone to represent the relevant mental state explicitly as well; so gaining knowledge by reasoning with hypotheticals typically leads us to be aware we have knowledge (Dienes & Perner, 1999). These considerations suggest that the guessing criterion may be useful in identifying which method of learning – single updating model vs. possibility-explicit representations – is being employed.

However, there is no logical reason why we may not recognise the answers produced by a connectionist network as being knowledge. In the connectionist networks of artificial grammar learning to date, the network produces a continuous output that reflects the degree to which the structure of the test item matches the structure of the training items. This output represents an objective property of the world and hence constitutes a first-order state; we will call this state (first-order) familiarity. Familiarity enables people to make worldly discriminations, namely, in the typical artificial grammar learning paradigm, whether a test item is grammatical or not. Familiarity might also give rise to various second-order states. For example, first-order familiarity may also be used by the processes that make higher-order thoughts (the HOT box) for generating confidence judgments about those discriminations. When a judgment of grammaticality is given to a string with particularly high familiarity, the judgment can be given a high confidence (and likewise for a judgment of non-grammaticality for a string with very low familiarity). In this case the participant has conscious knowledge of the grammaticality of the string; they know the grammaticality of the string and know that they know. Similarly, one may become aware of knowing the string is familiar to a certain degree, so the familiarity itself can be a conscious state. Consistently, both Cleeremans and Jiménez (2002) and Shanks (2005a) have proposed "single process" models (i.e., connectionist models) that lead to conscious knowledge.

The content of the conscious knowledge gained about grammaticality judgments is simply the judgment that the string is (or is not) grammatical. Dienes and Scott (2005) called this *judgment knowledge*. But the knowledge embedded in the weights can remain unconscious even when judgment knowledge is conscious. The knowledge embedded in the weights is knowledge of the structure of the strings as a whole. Similarly, in learning by hypothesis testing, the hypotheses contain knowledge of the structure of the strings. Dienes and Scott (2005) introduced a simple way of determining whether people were aware of their *structural knowledge*. On each trial participants indicated whether they made their decision without either conscious judgment or structural knowledge (guessing), based on some conscious judgment knowledge but without conscious structural knowledge (intuition), or based on both conscious judgment and structural knowledge (rules or memory). In brief, we believe one method of learning, the connectionist single updating model, produces unconscious structural knowledge, and judgment knowledge that may or may not be conscious. Conversely, the other method of learning, involving the consideration of possibilities, typically produces both structural and judgment knowledge that are both conscious.

We propose two methods of learning but not that they operate in isolation (cf. Sun, 2002). For the remainder of the chapter we will spell out their interactions. First we give a flow diagram of the processes involved in learning and judging test strings when participants learn about their common structure either incidentally (by simply looking at or memorising the strings) or deliberately (by actively seeking to discern their common structure). We then consider the processes by which unconscious knowledge can become conscious and hence the process by which metaknowledge emerges. Finally, we discuss in detail recent evidence for the proposed model.

2 Dual Process Model of Artificial Grammar Learning

Figure 3.1 provides a simplified illustration of the proposed dual process model of artificial grammar learning. It is not intended to capture all relationships but rather to outline the key processes. The model identifies separate processes proposed to be active during either training or test phases and according to whether learning is either incidental or deliberate. It demonstrates how deliberate (possibility-explicit model) and incidental (single updating model) learning processes might interact and how knowledge may make the transition from unconscious to conscious. The different routes to acquiring knowledge and the processes involved are first described before the evidence for each is presented.

2.1 Incidental Learning

2.1.1 Training Phase

Incidental learning proceeds by a passive familiarisation with the training strings, that is, without consciously intending to determine their common structure. Though participants are sometimes encouraged to memorise the strings during training, there can be considerable learning even when participants are directed to simply look at each string for a few seconds. This passive exposure is anticipated to result in familiarity with various common features of the training strings (Higham, Vokey, & Pritchard, 2000; Kinder, Shanks, Cock, & Tunney, 2003; Servan Schreiber & Anderson, 1990). These features will include localised characteristics such as bigrams and trigrams, as well as global characteristics such as the repetition structure. The focus of attention determines the input to the learning network and hence what features are learnt and abstracted. The familiarity of any such feature will roughly reflect the frequency with which it occurs.

2.1.2 Test Phase

Participants are set the task of distinguishing grammatical from non-grammatical test strings. They are informed that the training strings obeyed a complex set of rules and that, while all the test strings are new, exactly half will obey the same rules as the training strings. At this stage participants are familiar with the common features

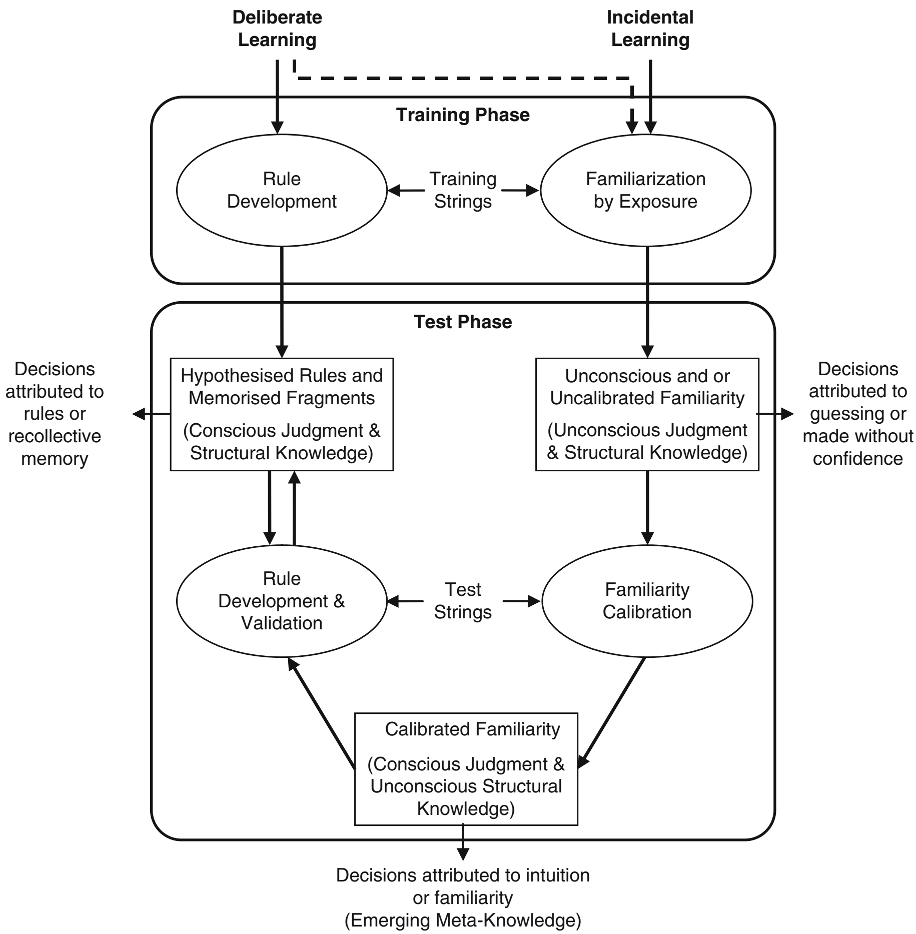


Fig. 3.1 A dual process model of artificial grammar learning

of the training strings but have no means to know how similar those training strings will be to either valid or invalid test strings. All the strings are compiled from the same letter set and many features may be common to both. In the tests we have conducted, after each classification of a string as grammatical or non-grammatical, participants gave their confidence in the classification. This can be used to assess the conscious status of judgment knowledge, as we have indicated. Participants also stated whether the basis of the classification was use of random responding, intuition, familiarity, rules, or recollection. This attribution assesses the conscious status of the structural knowledge

While some judgments may be made based on specific features that participants either recall seeing or not seeing during training and hence be attributed to rules or recollection, the primary basis for decisions after incidental learning is familiarity. Familiarity reflects the unconscious structural knowledge embedded in the weights of the network: A stimulus or part of a stimulus being familiar does not entail consciously knowing why it is familiar (Norman, Price, & Duff, 2006). Consider the common experience of feeling someone has changed in appearance in some way; they are strangely unfamiliar, without knowing why. The familiarity itself also need not be conscious. However, participants are often (though not always) aware of familiarity and that it is a basis for their decisions.

When familiarity is a basis of decisions, a logical strategy would be to endorse strings with familiarity greater than the mean as grammatical and those less as non-grammatical; the mean being the best estimate for the intersection between the familiarity distributions of grammatical and non-grammatical strings. This is illustrated in Fig. 3.2. Further, as long as the network was adapted to the structures it was learning, if confidence in grammatical decisions increased with increasing familiarity and confidence in non-grammatical decisions increased with decreasing familiarity, confidence would appropriately track accuracy. The process by which "guess" vs. "sure" confidence judgments become more tightly linked to familiarity we call calibration. Calibration requires obtaining ever more reliable estimates of mean familiarity, and also an assessment of the reliability of that estimate. If the mean familiarity estimated is uncertain, confidence in the grammaticality decision should also reflect this uncertainty. Thus, initially confidence in grammaticality decisions may be low when the mean familiarity is not taken to be estimated reliably. Lau (2008), also adopting a higher-order thought theory, describes a similar calibration process (of estimating signal and noise distributions in visual perception) as being the basis of conscious perception.

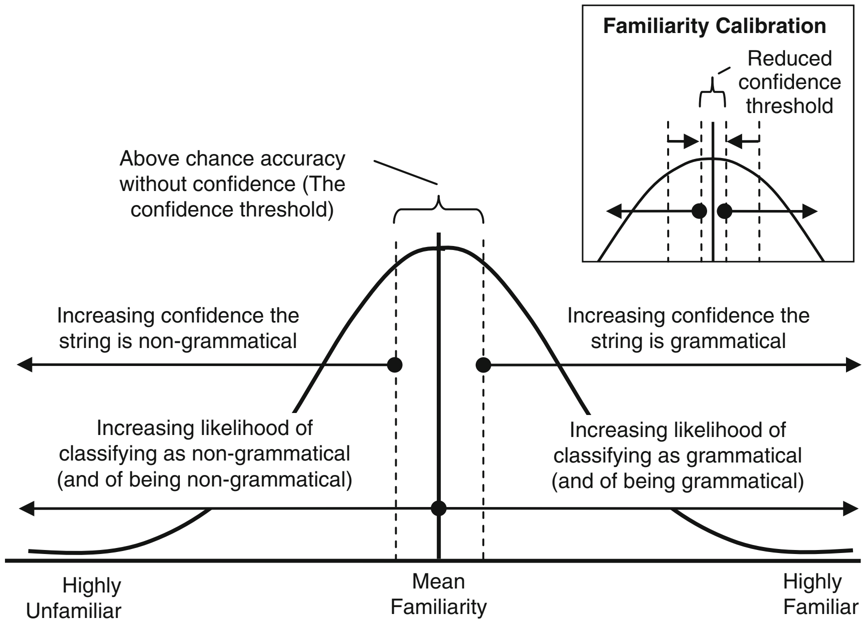


Fig. 3.2 Familiarity based grammaticality judgments and their relation to confidence

As the test phase proceeds, exposure to the test strings will permit the continuously updating estimate of the mean familiarity to be taken, as ever more reliable. This is represented in Fig. 3.1 by the familiarity calibration process. The resulting *calibrated familiarity* gives rise to conscious judgment knowledge (with unconscious structural knowledge). The assessed reliability of mean familiarity permits participants to have confidence when classifying some of the strings. Initially confidence will arise only when classifying strings furthest from the mean. As the assessed reliability improves, however, the minimum familiarity difference associated with confidence (the confidence threshold) will gradually shrink (see Fig. 3.2 inset). In instances where the mean estimate is poorly bounded and a string's familiarity is not highly distinctive, participants may experience confidence in their judgment without being aware of employing familiarity to make it. Judgments of this sort would logically be attributed to intuition. As the mean estimate improves, however, it will become increasingly apparent to participants that their judgments and confidence reflect differences in familiarity.

Once participants are able to distinguish the likely grammaticality of the test strings based on their familiarity, they can employ that knowledge to derive conscious structural knowledge. This is illustrated in Fig. 3.1 by the test phase rule development and validation process. As test strings are encountered participants may identify features which are common to those they classify as grammatical or non-grammatical. That is, they can use the test strings to discern the possible objective bases for their familiarity differences and in doing so derive conscious structural knowledge. Familiarity is initially a unidimensional continuous output returned from the current focus of attention. Thus, whole strings can be divided into classes when the whole string is attended (grammatical and non-grammatical for high and low familiarity strings) and this classification can be used to test rules about the difference between grammatical and non-grammatical strings. For example, if I notice that none of the strings I have classified as grammatical contain "XXX", I can subsequently use this explicit fragment knowledge to make judgments that are independent of familiarity. Knowledge of this sort is most likely to be attributed to the use of rules (e.g., "I have a rule that strings containing this trigram are non-grammatical"). However, the same knowledge may also feasibly be attributed to recollection (e.g., "I don't recall seeing this trigram in any of the training strings"). Indeed if a simple rule is repeatedly found to be applicable then that may itself increase a participant's conviction that the fragment did or did not appear during training. This, in turn, would increase the likelihood of subsequent judgments being attributed to the use of recollection (evidence for a trend from rule to recollection attributions will be presented subsequently).

Just as one can determine the familiarity of whole strings by attending to them, attention to different parts of strings can determine the relative familiarity of those parts. Further, after sufficient exposure to training materials, attention may be drawn to parts of strings with high differential familiarity. The notion that participants may ultimately be capable of distinguishing more or less familiar string fragments is consistent with their noted ability to underline grammatical elements (Dulany, Carlson, & Dewey, 1984). Crucially, the differential familiarity of individual features, while not essential to hypothesising about rules, would logically provide a considerable advantage (Evans & Over, 1999; Perruchet & Vinter, 2002). For example, if the fragment ABB in the string XABBCCD feels more familiar than other elements, then it would be logical to use it to derive a rule or to attempt to memorise it. Attention being drawn to features will naturally suggest rules to the participant. Attention being drawn to a part of a string does not in itself constitute conscious knowledge of that structure (it does not per se entail knowing that one knows anything); but when the rule development mechanism makes use of the attentional focus, conscious knowledge readily follows. When the structure is not about chunks but more global properties of a string it should be harder to explicate the structural knowledge – see Kuhn and Dienes (2005) who demonstrated implicit learning of musical inversions.

2.2 Deliberate Learning

2.2.1 Training Phase

In deliberate learning participants are aware that training strings embody a complex set of rules and are actively seeking to discern them. This is illustrated in Fig. 3.1 by the training-phase rule development process which utilises the ability of participants to consider the possible as well as the actual. Participants hypothesise potential rules and memorise what they believe to be pertinent fragments. The development of rules and the memorising of fragments cannot be wholly separated as rules will most often require that fragments be committed to memory. For example, common rules might include which letters can start or end a string, or which combinations of letters can or can't be repeated. The process is directly dependent on attentional resources, with participants focusing on what they take to be the most relevant features of the strings. However, the process is not entirely separable from familiarity. First, those aspects of a string receiving attention in order to derive rules will doubtlessly become the most familiar. In addition, to the extent that attention is imperfect during learning, some degree of incidental learning through passive familiarisation will also occur. This is indicated in Fig. 3.1 by the dashed line connecting deliberate learning with the familiarisation process and will be shown to be predictably related to personality differences.

2.2.2 Test Phase

Participants will start the test phase with hypothesised rules and memorised string fragments from the training-phase rule development process, and with some uncalibrated familiarity arising through incidental learning. The majority of conscious structural knowledge expressed after deliberate learning is the result of rule development during training. However, existing knowledge will be refined and new knowledge will be developed during the test phase. This is envisioned as occurring via two mechanisms. Firstly, to the extent that incidental learning takes place during training the subsequent familiarity calibration and rule development associated with that process will again be active. Secondly, as participants seek to classify the test strings they will validate or invalidate the rules acquired during training.

The test-phase rule-development and validation processes are interrelated and hence depicted as a single process in Fig. 3.1. Rule validation takes place by at least two methods. In its simplest form rules derived during training but subsequently found to apply to either too many or too few test strings will be dropped. For example, participants may memorise the limited set of starting letters used in training strings only to subsequently discover that all the test strings conform to this constraint. The second means of validation results from the interaction of deliberate and incidental learning processes. The calibrated familiarity differences driving test-phase rule-development under incidental learning may also be used to validate rules derived during deliberate learning. If a codified rule indicates that a string is valid and yet the familiarity of strings conforming to that rule indicates otherwise then participants may re-examine the rule in question.

3 Mapping the Transition from Unconscious to Conscious as a Metacognitive Process

The proposed model postulates a process whereby, under incidental learning, there is a transition from unconscious to conscious knowledge. Early grammaticality judgments reflect unconscious judgement knowledge drawing on unconscious structural knowledge; accurate judgments are made without confidence. Over time conscious judgement knowledge emerges through a process of familiarity calibration – judgments subsequently being attributed to intuition and familiarity. Finally, conscious structural knowledge emerges, as participants discern the objective bases for the familiarity differences, resulting in judgments attributed to rules and recollection.

We combined data from two separate experiments ($N=60$) – a pilot study and Experiment 1 of Scott and Dienes (2009) – to examine the transition in participants' reported basis for their grammaticality judgements. Participants memorised 48 training strings (16 strings repeated three times) for five seconds each and subsequently classified each of 32 test strings twice in consecutive blocks. None of the test strings appeared in training but exactly half conformed to the training grammar. Participants were asked to classify the strings as grammatical or non-grammatical and to indicate the basis for their judgments according to the following decision strategies: random selection, familiarity, intuition, rules, or recollection.

We examined how the decision strategies used to classify each of the test strings changed between the two time points. The phi correlation coefficient was computed for the relationship between each decision strategy at time one and each alternative strategy at time two. Figure 3.3 shows all the significant positive associations resulting from those analyses. The decision strategies are arranged in order of increasing meta-knowledge from left to right. Attributing a response to random selection is taken to indicate the absence of both conscious judgment knowledge and conscious structural knowledge. Attribution of responses to either intuition or familiarity

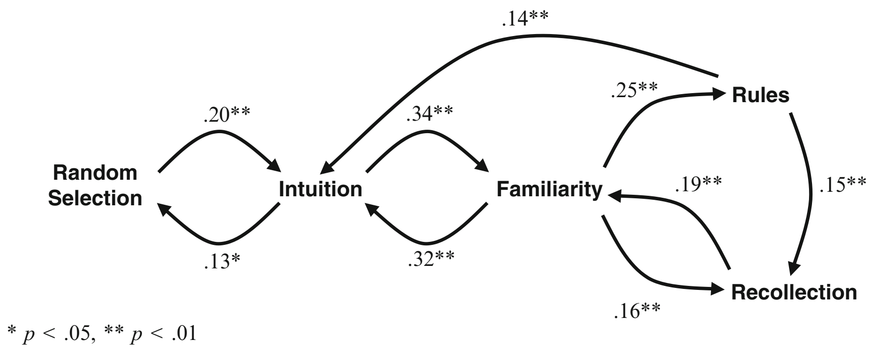


Fig. 3.3 Change in reported basis for grammaticality judgments made for the same test strings classified at two consecutive time points. All positive associations are shown with phi correlation coefficients and associated probabilities. $N = 60$. Degrees of freedom for individual associations range between 20 and 56

indicates conscious judgment knowledge in the absence of conscious structural knowledge, that is, the absence of knowledge of the specific features which make the strings grammatical or non-grammatical. These attributions differ in that intuition responses indicate that participants are unaware of any basis for their confidence, whereas familiarity attributions indicate that they are aware of using the overall familiarity of the strings to distinguish them. Finally, responses attributed to either rules or recollection indicate both the presence of conscious judgment knowledge and conscious structural knowledge. The mean accuracy for judgments attributed to random selection was 60%, significantly greater than chance indicating unconscious judgment knowledge as measured by the guessing criterion.

The change in reported basis for participants' grammaticality judgments is consistent with the proposed model. The changes in decision strategies show a clear pattern of increasing metacognition. There is a transition in reported decision strategy from random selection to intuition, from intuition to familiarity, and from familiarity to both rules and recollection. Strings classified on the basis of rules are also subsequently likely to be classified based on recollection. The association between rules and intuition suggests that when rules are found to be unreliable participants may subsequently rely on intuition. Similarly, recollection flows back to familiarity, familiarity to intuition, and intuition to random selection.

4 Familiarity as the Basis of Incidental Learning

The proposed model holds that subjective familiarity is the primary source of knowledge under incidental learning conditions. Research has long suggested that familiarity plays an important role in Artificial Grammar Learning (AGL). However, its role has generally only been evaluated indirectly. The contribution of familiarity has variously been inferred from the following: (a) from the relationship between grammaticality judgment and fragment frequency (Knowlton & Squire, 1996; Meulemans & Van der Linden, 1997; Servan Schreiber & Anderson, 1990); (b) from the conformance of signal detection analyses to familiarity based models (Kinder & Assmann, 2000; Lotz & Kinder, 2006); and (c) from the effects on performance resulting from fluency manipulations (Kinder et al., 2003).

We sought to adopt a more direct means of evaluating both the basis of subjective familiarity in AGL and its role in making grammaticality judgments (Scott & Dienes, 2008). In a series of three AGL experiments we had participants provide subjective ratings of familiarity for each test string in addition to providing grammaticality judgments, confidence ratings, and reports of the subjective basis for their decisions. The results were replicated across all three experiments and provide strong support for the proposed model.

Consistent with the suggested primary use of familiarity for grammaticality decisions both grammaticality judgments and confidence ratings were reliably predicted by familiarity ratings. More familiar strings were substantially more likely to be endorsed as grammatical ($r=0.73$), and their associated confidence was predicted by the extent to which their rated familiarity differed from the mean ($r=0.48$). Signal detection models assume overlapping probability distributions of some continuous variable; our approach permits those distributions to be examined directly. Figure 3.4 shows the probability distribution of z-transformed familiarity for grammatical and non-grammatical strings. These distributions are consistent with the observed decision processes using an estimate of the mean z-familiarity as their criterion.

The subjective familiarity ratings were themselves reliably predicted by objective measures of similarity between training and test strings. In combination the similarity measures accounted for 20% of the variation in familiarity. Similarity measures included statistics evaluating the frequency with which fragments of test strings

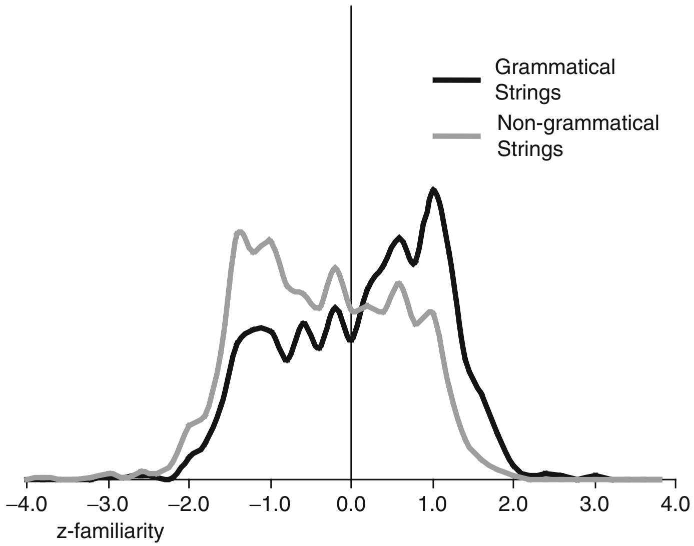


Fig. 3.4 The probability distribution of z-transformed familiarity scores for grammatical and non-grammatical strings

appeared during training, as well as global similarities such as their repetition structure. Importantly, measures of both fragment frequency and repetition structure were significant predictors of familiarity. Lotz and Kinder (2006) demonstrated that Receiver Operating Characteristics (ROCs) remain consistent with a familiarity-based process for AGL under transfer conditions – where the letters used to construct the grammar strings are changed between training and test. The contribution of repetition structure to subjective familiarity is consistent with that finding as repetition structure is preserved when surface features are changed. Thus subjective familiarity can account for accurate grammaticality judgments made under transfer conditions.

It is standard to associate familiarity with fluency in the memory literature. Jacoby and Dallas (1981) proposed that when processing an item with relative ease, or *fluently*, people may attribute this to the item having been seen before and experience it as being more familiar. Similarly, in AGL, test strings containing features in common with the training strings may be processed more fluently resulting in greater familiarity. Perceptual processing fluency – as a potential basis for grammaticality judgments in AGL – was experimentally explored by Buchner (1994) and Kinder et al. (2003). However, neither study incorporated a measure of subjective familiarity which would be necessary to assess their relative contributions.

In a sequence of four experiments we evaluated the relationship between perceptual processing fluency and subjective familiarity ratings in AGL, and the capacity of each to predict grammaticality judgments (Scott & Dienes, 2009). Perceptual fluency was found to have a small influence on ratings of subjective familiarity ($r=0.07$). However, when the complexity of grammatical and non-grammatical strings was counterbalanced, perceptual fluency was found to be unrelated to grammaticality. As such, perceptual fluency could not account for the accuracy of grammaticality judgments. In contrast, the relationship between familiarity ratings and grammaticality was substantial ($r=0.40$), and the relationship between familiarity and grammaticality judgements replicated that found in the previous sequence of experiments ($r=0.73$). The observed relationships between grammaticality and processing fluency (as measured by reaction times in a perceptual clarification task), and between grammaticality and familiarity are illustrated in Fig. 3.5.

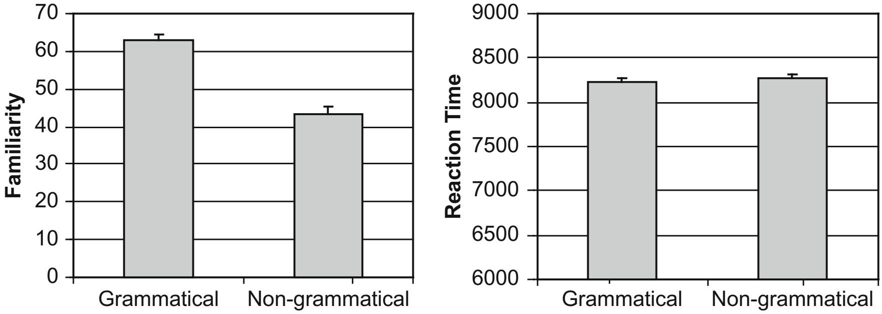


Fig. 3.5 Mean familiarity ratings and reaction times (from a perceptual clarification task) for grammatical and non-grammatical test strings

In sum, while we believe familiarity plays a major role in the expression of implicit knowledge, we do not believe that it is derived from perceptual processing fluency.

5 Different Sources of Conscious Structural Knowledge

The dual process model proposes that both the nature of conscious structural knowledge and the time at which it is derived, differs according to the learning condition. Under deliberate learning conditions conscious structural knowledge is predicted to arise primarily from hypothesising about rules during training. In contrast, conscious structural knowledge under incidental learning conditions is thought to develop during the test phase and to be fairly directly derived from familiarity differences. Once input has been attended, learning and application of knowledge by a standard connectionist network does not on the face of it require executive resources. Conversely, the often metacognitive process of considering multiple models plausibly requires working memory. Conscious structural knowledge acquired during deliberate learning should be both more reliant on executive resources available during training and result in knowledge less directly related to familiarity.

Dienes and Scott (2005) examined the effect of divided attention during the deliberate and incidental learning of artificial grammars. During training, participants were either instructed to memorise the strings or made aware of the presence of rules and asked to try to identify them. In addition, half of the participants were required to divide their attention between the learning task and the generation of random numbers in time with a metronome. At test, participants were required to make grammaticality judgements and to report the subjective basis of those judgements by attributing them to one of four decision strategies – guess, intuition, rules, or memory. Analysis revealed a significant three-way interaction between learning condition, attentional demands, and type of decision strategy on the accuracy of participants' grammaticality judgments (see Fig. 3.6). Divided attention significantly

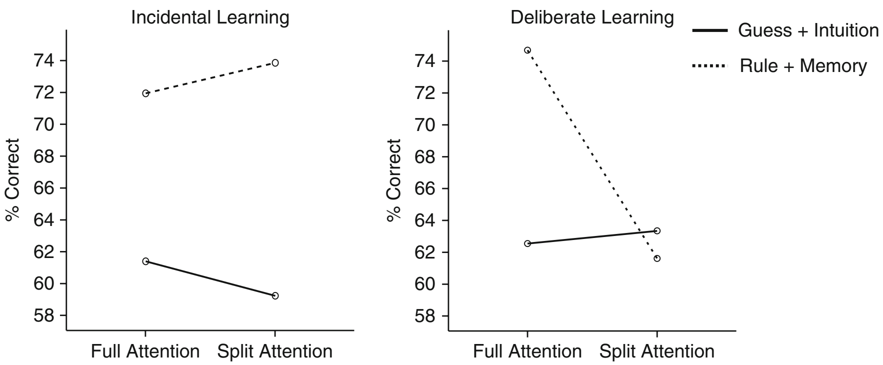


Fig. 3.6 The effect of attentional demands, learning condition, and the reported basis for grammaticality judgements on the percentage of grammaticality judgments correct

reduced the accuracy of only those judgments attributed to sources indicating conscious structural knowledge (rules and memory), and did so only under deliberate learning conditions. The results are consistent with the development of conscious structural knowledge being reliant on attentional resources during training for deliberate but not for incidental learning.

The same pattern of results was observed where the time available to process training strings was restricted. A separate analysis of Experiments 1 and 3 from Scott and Dienes (2008) revealed that when a blank delay between training strings was removed the development of conscious structural knowledge was impeded under deliberate but not incidental learning conditions. At test, both the number and accuracy of attributions attributed to the rule and memory categories were significantly reduced in the absence of a delay between training strings. The accuracy of other attributions was not significantly affected (see Fig. 3.7).

In common with the effect of divided attention, the absence of a delay appears also to reduce participants' ability to derive accurate rules and consolidate recollective memories, thus impeding the development of conscious structural knowledge under deliberate learning conditions. This effect has a direct parallel in the serial reaction-time (SRT) task. Destrebecqz and Cleeremans (2001) employed an SRT task while manipulating the time interval between providing a response to one stimulus and the appearance of the next stimulus – the response stimulus interval (RSI). The assumption behind the manipulation is that a shorter RSI gives less time for stimulus processing, which selectively impairs the development of explicit representations. Participants were asked to generate a sequence that was not the training sequence. Destrebecqz and Cleeremans (2001) found that, with an RSI of zero (RSI-0), participants were not able to comply with these instructions, whereas with an RSI of 250 ms (RSI-250) they were able to refrain from generating the training sequence at above baseline levels. They interpreted these findings as indicating that learning is implicit at RSI-0, but explicit at RSI-250. Although Wilkinson and Shanks (2004) failed to replicate this finding, Fu, Fu, and Dienes

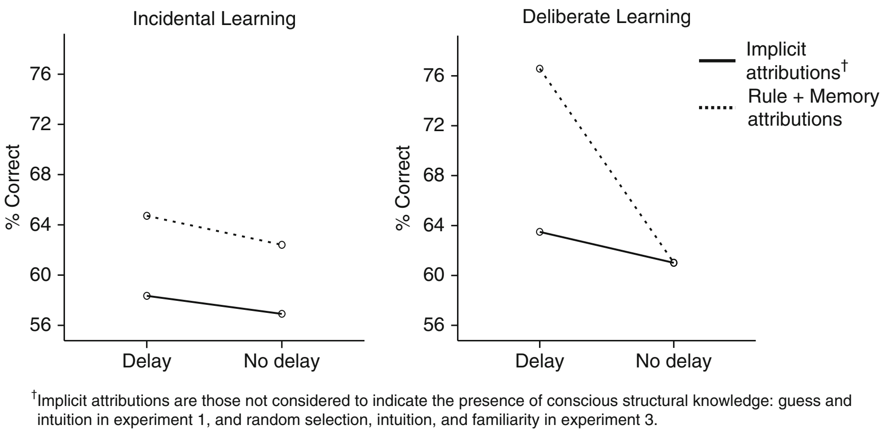
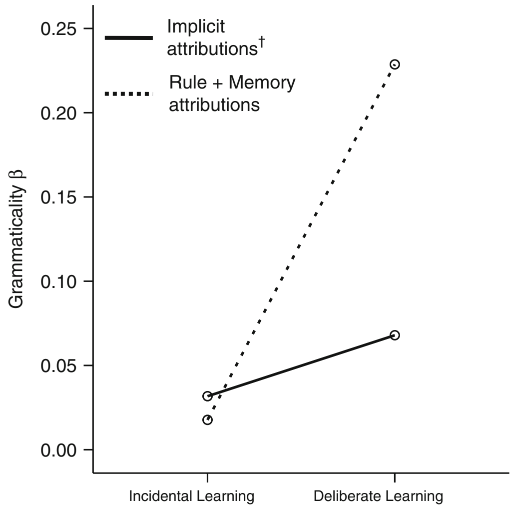


Fig. 3.7 The effect of a delay between training strings, learning condition, and the reported basis for grammaticality judgements on the percentage of grammaticality judgments correct

(2008) both replicated the original finding and provided an account for Wilkinson and Shanks' (2004) failure to do so.

Evidence for a qualitative difference in the nature of conscious structural knowledge derived under deliberate and incidental conditions is provided by Scott and Dienes (2008). Each of three experiments used familiarity ratings, grammaticality judgments, and the reported basis for those judgments as indicated by their attribution to guess, intuition, rules, or recollection (or random selection, intuition, familiarity, rules and recollection in Experiment 3). Multiple regression analysis was used to examine the relative contribution made to grammaticality judgements of familiarity (as captured by familiarity ratings) and grammaticality (whether or not the strings were grammatical). These analyses were conducted separately for responses attributed to each of the decision strategies. Familiarity was found to make a significant contribution irrespective of learning condition or the reported basis of grammaticality judgments. In contrast, the contribution of grammaticality, controlling for familiarity, revealed a significant interaction between learning condition and type of decision strategy (see Fig. 3.8). A reliable additional contribution of grammaticality over and above that of familiarity was only observed in the deliberate learning condition, and then only for judgments attributed to rule and memory attributions. Consistent with the proposed model, conscious structural knowledge derived under incidental conditions appears to be fairly directly derived from familiarity differences. In contrast, conscious structural knowledge derived under deliberate learning conditions reflects knowledge of the grammar that is additional to that derived from familiarity.

The existence of two alternative routes to the development of conscious structural knowledge may explain how, where rules are complex, explicit attempts to learn



†Implicit attributions are those not considered to indicate the presence of conscious structural knowledge: guess and intuition in experiments 1 and 2, and random selection, intuition, and familiarity in experiment 3.

Fig. 3.8 The effect of learning condition and reported basis of grammaticality judgments on the contribution that grammaticality (controlling for familiarity) makes to grammaticality judgments can actually impede performance (Reber & Lewis, 1977). Conscious attempts to hypothesise about rules prior to extended exposure would not be guided by differential familiarity and may consequently result in more inaccurate inferences. In contrast, rules devised during the test phase after incidental learning will benefit from the knowledge inherent in familiarity differences.

6 Familiarity Calibration and the Emergence of Metacognition

It has been established that the familiarity of strings relative to the mean familiarity reliably predicts both participants' judgments and their confidence in those judgments. Their confidence, and hence their conscious judgment knowledge, is thought to develop during the test phase by the process of familiarity calibration (cf. Redington, Friend, & Chater, 1996). As calibration proceeds, the amount by which familiarity has to differ from its mean before participants feel they are not guessing should shrink. That is, their *confidence thresholds* should decrease over trials.

Scott and Dienes (2008; Experiment 2) examined the proposed calibration process and whether it could be accelerated by positive feedback. A sample of 160 participants – trained under either incidental or deliberate learning conditions – provided grammaticality judgments, confidence ratings, and reported the basis for their judgments in the usual way attributing them to guess, intuition, rules, or memory. In addition, half of the participants rated the familiarity of each test string. The experimental manipulation took place throughout the test phase with half of the participants given positive feedback intended to encourage them to be more confident. Feedback took the form of on-screen warnings that they had thus far been under-confident in their grammaticality judgments; this was given irrespective of actual performance.

An estimate of the confidence threshold, as illustrated in Fig. 3.2, was obtained from the standard deviation of familiarity ratings for judgments where participants indicated having no confidence (guesses).[1] Consistent with the model, the width of participants' confidence thresholds was found to be related to the presence of conscious judgment knowledge. Specifically, the width estimated from the entire test-phase was significantly positively related to both the number and accuracy of judgments attributed to guessing ($r = 0.20$ and $r = 0.21$ respectively), and negatively related to the number of judgments attributed to intuition ($r = -0.34$). To examine changes over time the confidence threshold was estimated at three time-points; for the first, second, and final 20 grammaticality judgments. Here again, the results were consistent with predictions. The mean confidence threshold reduced over time and was also substantially lower for participants given positive feedback (see Fig. 3.9). Furthermore, consistent with the adoption of a narrower confidence

[1]Familiarity ratings were standardised (z-transformed).

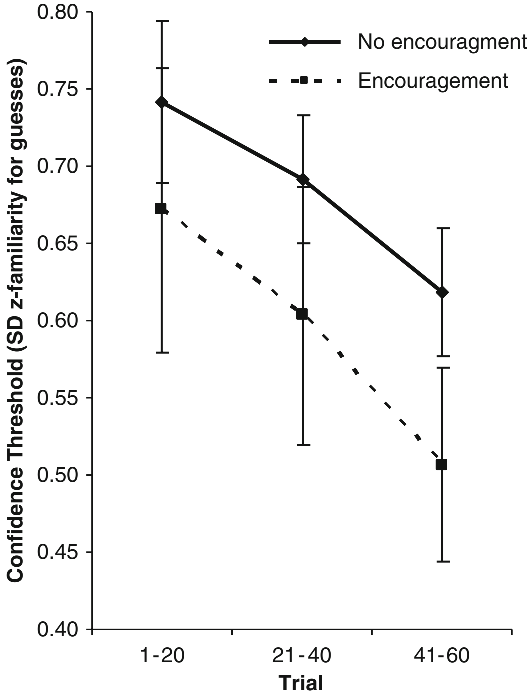


Fig. 3.9 The reduction in confidence threshold over time for participants encouraged to be more confident (warned they are being underconfident) or not given encouragement

threshold, both the number and accuracy of judgments attributed to guessing were significantly lower for participants encouraged to be more confident.

These results strongly support the notion that conscious judgment knowledge emerges as familiarity differences are calibrated during the test phase – see Lau (2008) who discusses a similar process for perception. The full model further proposes that knowledge gained in this way is then exploited to derive conscious structural knowledge in the form of rules. This process would predict that as the calibration process proceeds more grammaticality judgements will be attributed to rules. The same experiment reported above revealed exactly this pattern. After incidental exposure to the training strings, the number of grammaticality judgments attributed to rules increased significantly over precisely the same time-frame that the confidence threshold was observed to decrease.

We contend that the familiarity calibration process in AGL is just one example of a more general process whereby metacognition emerges as our assessed predictive ability breaches some threshold. It is proposed that we continually, and unconsciously, attempt to predict our environment, with feedback permitting those predictions to be evaluated and refined (cf. Gray, 1995). At some threshold of assessed predictive ability the higher-order representation of that ability is established and meta-cognition results. Prior to that threshold, predictions will influence behaviour without conscious awareness; this is an adaptive process as many predictions can only be evaluated if acted upon (albeit acted upon unconsciously). By this account unconscious knowledge will be apparent when our predictions are above chance but below our predictive threshold, or when they are above that threshold but there has been insufficient feedback for predictive ability to be assessed. In the context of AGL the automatic predictions relate to the frequency with which test strings with differing degrees of familiarity will be encountered, and feedback is gained as more test strings are seen. However, this same framework – the higher-order representation of predictive ability – can be applied to any conscious experience, with the variety of phenomenology mirroring the multitude of predictive contexts (cf. Chrisley & Parthemore, 2007). For example, conscious vision can be seen as the assessed ability to predict how sensory input to the visual system changes over time or with our movements (O'Regan & Noe, 2001).

7 The Conscious and Unconscious Use of Familiarity

The findings presented thus far provide compelling evidence that familiarity plays a central role in guiding participants' grammaticality judgments. However, participants' awareness of familiarity and of employing it has not been addressed. Familiarity is an ambiguous word. On the one hand, it can refer to a first-order state, namely knowing that an object or its features are old to some extent (the output of the learning network). In this sense, a participant indicating the familiarity of a stimulus may be guessing its objective properties and in that sense, the guesses may reflect unconscious knowledge of stimulus properties in the same way that the guesses of a blindsight patient do. On the other hand, familiarity also refers to a feeling we can have about a mental state. In this sense it implies a second-order state: an awareness of familiarity as a first-order state. We asked participants to rate the *feeling* of familiarity elicited by each test string. Prima facie then the familiarity ratings we obtained measure conscious familiarity. However, we have not tested this claim thoroughly. Just as a blindsight patient asked to indicate the direction in which they felt an object was moving may sometimes (in cases where they had no feeling) *guess* the direction instead of reporting it, it is also possible our participants did the same. This will be a matter for future testing.

Given familiarity itself is conscious, an additional question is whether participants are aware that familiarity played a role in their grammaticality decisions. As Chomsky (1957) indicated, grammaticality and first-order familiarity are quite separate properties (in the famous example, "colourless dreams sleep furiously" may be very unfamiliar to a person yet judged as completely grammatical). Participants need not consciously assume that grammaticality decisions were or should be based on familiarity. Scott and Dienes (2008 Experiment 3) explored participants' awareness of using familiarity to make grammaticality judgments. After training under incidental or deliberate conditions, participants made grammaticality judgments and reported confidence ratings in the usual way. When reporting the subjective basis for their grammaticality judgments, they were asked to choose from random selection, intuition, familiarity, rules, and recollection. Participants were required to report their confidence and the basis for each grammaticality judgment separately. Thus participants were able to report using systematic strategies, such as familiarity, to make grammaticality judgments while simultaneously indicating that they had no confidence in those decisions. The proportion of grammaticality judgments attributed to familiarity increased significantly over time, consistent with awareness of its use increasing as familiarity calibration proceeded. Familiarity was also the most commonly reported decision strategy, on average reported as the basis for 33% of grammaticality judgments, including 20% of judgments made without confidence. That is, participants were often aware of exploiting familiarity to make grammaticality judgments and did so even when they lacked confidence in those decisions. However, the same experiment also provided evidence that familiarity may at times influence participants' judgements without their awareness. Controlling for the contribution of grammaticality (whether or not a string was grammatical) familiarity ratings were found to reliably predict participants' grammaticality judgments even for those judgments reportedly based on selecting responses at random ($\beta=0.29$). This surprising result has since been replicated in each of four additional experiments with a mean familiarity β of 0.43 (Scott & Dienes, 2009). In sum, participants often are aware of using familiarity to make grammaticality decisions, but sometimes are not.

The use of familiarity in AGL is also consistent with findings in the serial reaction-time task. Norman, Price, Duff, and Mentzoni (2007) employed an SRT task where participants completed learning, recognition, and generation tasks and were subsequently required to indicate the feelings they had experienced. The most frequently selected response was a "feeling of familiarity", with a 56% probability of this feeling being reported after any given task. The use of subjective familiarity in both AGL and SRT studies suggests that feelings of familiarity may be central to implicit learning in general rather than specific to a particular experimental framework.

8 Metacognition and Individual Differences in Learning Style

The concepts of implicit and explicit learning in cognitive psychology are paralleled by numerous concepts in the personality and social psychology literatures. Among others, these include the experiential and rational learning styles proposed by Epstein (1983). A number of self-report measures have developed around these concepts and have demonstrated behavioural implications in contexts ranging from impression formation (Cacioppo & Petty, 1982) to non-optimal responses in games of chance (Pacini & Epstein, 1999). If the parallel between these concepts and those of implicit and explicit learning is the result of more than superficial similarity then these, or similar measures, might be expected to predict behavioural differences in AGL. The interaction between implicit and explicit processes in the model we propose prevents simple performance predictions. However, it was anticipated that the different routes by which participants acquire metaknowledge may reflect their preferred learning styles.

Though not reported in that paper, Experiment 1 of Scott and Dienes (2008) used the Rational-Experiential inventory devised by Epstein, Pacini, Denes Raj, and Heier (1996) to measure participants' experiential and rational learning styles and assess how their scores related to responding in an AGL task. The inventory contains two sub-measures: (a) a modified form of the Need for Cognition (NFC) scale evaluating analytical-rational thinking, and (b) the Faith in Intuition (FI) scale evaluating intuitive-experiential thinking. Persons high vs. low in NFC report greater inclination towards, and ability in, rational-analytical tasks. In contrast those high vs. low in FI report greater inclination towards, and ability in intuitive-experiential responding. The two measures are uncorrelated, so it is feasible for individuals to be high or low on either or both measures. In theory the two learning styles should facilitate different processes in the proposed AGL model. The strong analytical ability associated with higher NFC should aid the development of conscious structural knowledge by assisting the rule development processes. The strong intuitive processing associated with higher FI should aid the development of conscious judgment knowledge based on feelings (intuition attributions). Consistent with these simple predictions, significant positive correlations were found between both NFC and the number of judgments attributed to rules ($r=0.25$), and between FI and the number of judgments attributed to intuition ($r=0.23$).

Norman et al. (2006) explored similar concepts using the personality measure Openness to Feelings which was chosen to reflect the ability to introspect on what they term *fringe feelings* of consciousness. These fringe feelings include feelings of rightness or wrongness as well as feelings of familiarity. They found that the familiarity ratings of participants higher in openness to feelings more accurately differentiated old from new sequence fragments in an SRT task – though contrast Norman et al. (2007). Importantly, this difference was observed with RSI-0 but not with RSI-250. That is, the benefit was apparent only under the condition intended to impede the development of explicit representations (RSI-0), which is broadly equivalent to an incidental learning condition in AGL.

We were intrigued to examine whether participants' inclination towards an experiential learning style, as measured by FI, was similarly related to a greater sensitivity to feelings of familiarity. Consistent with Norman et al. (2006), a relationship was only observed under incidental learning conditions where FI was significantly related to participants' average familiarity ratings ($r=0.49$). The higher a participants' FI the more familiar they found the test strings. Furthermore, again under incidental learning conditions, FI was significantly related to how well participants' familiarity ratings reflected the objective similarity of training and test strings ($r=0.46$).[2] Our model would predict that an increased sensitivity to familiarity, like

[2]For each participant, familiarity was regressed on seven measures of the structural similarity between training and test strings. The adjusted R^2 from those regressions was used as the measure of how well familiarity was predicted by structural similarity. FI was significantly related to the adjusted R^2 values.

that observed with higher FI, should permit the adoption of a narrower confidence threshold. This in turn would increase conscious judgment knowledge and thus account for the greater proportion of judgments attributed to intuition. Indeed higher FI was associated with narrower confidence thresholds but this relationship was only marginal ($r = -0.25$, $p = 0.08$).

The results presented are from a single study and need to be replicated before strong conclusions can be drawn. However, their consistency with related research suggests a potentially fruitful reconciliation between dual process accounts of learning and similarly dichotomised theories of personality.[3]

9 Discussion

In this chapter we have presented an account of learning artificial grammars based on two processes or methods: learning by a single updating model that has the function to reflect how reality is, and learning by the use of considering possibilities. The first method results in unconscious structural knowledge and is what we take the term implicit learning to refer to: The process of acquiring implicit structural knowledge. The process does produce unconscious knowledge, but not exclusively so, that is, it can produce conscious judgment knowledge and ultimately conscious structural knowledge too, because the outputs of the process are available to the mechanisms that produce conscious knowledge. We suggest that the output of the process is familiarity, which is often experienced consciously and can more or less directly guide the development of conscious structural knowledge.

Our account is consistent with that of Reber (1989), who regarded the results of implicit learning to be largely but not exclusively conscious; and yet our account is also largely consistent with Dulany (1997) and Shanks (2005b), barring a terminological difference. Both Dulany and Shanks regarded implicit learning as an associative process which gives rise to conscious feelings. According to Dulany's (1997) account, the associative processes themselves are not conscious (being processes rather than states), and so structural knowledge implicit in their operation can evoke certain feelings, such as familiarity. Similarly, Shanks (2005b) conceives implicit learning as being based on connectionist processes. For Dulany (1997), and also for Perruchet and Vinter (2002), mental states are conscious even if we are not conscious of them; for these authors, first-order familiarity and the accurate guesses of blind sight patients constitute conscious knowledge (even in the absence of any relevant higher-order thoughts). If one wishes to view the possession of relevant higher-order thoughts as indicating introspective or reflective consciousness rather than conscious awareness per se, then various claims of ours concerning making

[3]FI was also found to correlate with reported experience of déjà vu, $\rho(80) = 0.25$, $p = 0.024$, consistent with the theory that déjà vu experiences may result from misattributed familiarity (Jacoby & Whitehouse, 1989).

knowledge conscious can be read as claims concerning making knowledge reflectively conscious rather than conscious per se. For the purposes of this chapter that would be fine; our point is not to quibble over mere words. (Nonetheless, we believe higher-order thought theories provide a more natural use of words.) The aim is not to *classify* states as conscious or not for the sake of it; but to *identify* meaningful psychological processes. We hope the review of research here shows that considering the conscious status of judgment and structural knowledge may allow us to identify two different learning processes and their interaction. Indeed, it is possible that contradictory results in the implicit learning literature, for example concerning the effects that secondary tasks demanding executive resources have on implicit learning (cf. Jiménez, 2003), may be rendered consistent when measures of the conscious status of structural knowledge are taken to dissociate underlying processes more cleanly – see, also, Dienes (2008) for further examples and discussion.

Our use of the term familiarity is not the same as some other authors, such as Jacoby (1991). Jacoby (1991) defined familiarity as the memory process that occurs regardless of intentions. We define it as the continuous indication of oldness that emerges from a learning network and can be reflected in ratings of familiarity given by a person. Whether or not such familiarity is sensitive to intentions is then a contingent rather than definitional issue. In fact, there is evidence that familiarity is sensitive to the intentions of the person. Wan, Dienes, and Fu (2008) trained participants on two grammars and then asked them to endorse strings from only one of them. Participants also rated how familiar each string felt and reported whether or not they used familiarity to make their grammatically judgment. Participants proved able to endorse the strings of one grammar and ignore the strings from the other. Importantly, when participants said they were using familiarity, the rated familiarity for test strings consistent with their chosen grammar was greater than that for strings from the other grammar. Familiarity, subjectively defined, was sensitive to intention. Hence, counter-intuitively, familiarity can be used as a basis for metacognitive control in deciding whether one body of knowledge applies or another. Similarly, it may be that people can focus on a certain level of structure, for example, global or local (see Tanaka, Kiyokawa, Yamada, Dienes, & Shigemasu, 2008) and have familiarity reflect the level attended to. This is a matter for future testing.

Many authors in the implicit learning literature, who favour single system models (e.g., Cleeremans & Jiménez 2002; Shanks, 2005b), focus on the fact that the learning network that produces unconscious structural knowledge can produce conscious judgment knowledge (and ultimately conscious structural knowledge). However, we do not believe they would object that the considering of possibilities is rather different from the way a first-order connectionist network attempts to represent just how reality is (compare Shanks & St John's, 1994, distinction between exemplar-based and rule-based learning). Nor in turn do we object that ultimately the brain is a neural network that as one system accomplishes both types of learning. The challenge is not to count systems but to specify in detail the components and their interactions. In this vein, we have argued for the fruitfulness of distinguishing learning by a connectionist network that just represents reality from learning by the consideration of hypotheticals.

References

Buchner, A. (1994). Indirect effects of synthetic grammar learning in an identification task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(3), 550–566.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116–131.

Carruthers, P. (2000). *Phenomenal consciousness: A naturalistic theory.* Cambridge, UK: Cambridge University Press.

Chomsky, N. (1957). *Syntactic structures.* The Hague, The Netherlands: Mouton.

Chrisley, R., & Parthemore, J. (2007) Synthetic phenomenology: Exploiting embodiment to specify the non-conceptual content of visual experience. *Journal of Consciousness Studies 14*, 44–58.

Cleeremans, A., & Dienes, Z. (2008). Computational models of implicit learning. In R. Sun (Ed.), *Handbook of computational cognitive modeling* (pp. 396–421). Cambridge, UK: Cambridge: University Press.

Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness* (pp. 1–40). Hove, UK: Psychology Press.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin and Review, 8*(2), 343–350.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Models of Brain and Mind: Physical, Computational and Psychological Approaches, 168*, 49–64.

Dienes, Z., Altmann, G. T. M., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(5), 1322–1338.

Dienes, Z., & Perner, J. (1999) A theory of implicit and explicit knowledge. *Behavioural and Brain Sciences, 22*, 735–755.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research, 69*(5–6), 338–351.

Dulany, D. E. (1997**).** Consciousness in the explicit (deliberative) and implicit (evocative). In J. Cohen & J. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 179–212). Mahwah, NJ: Erlbaum.

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General, 113*(4), 541–555.

Epstein, S. (1983). The unconscious, the preconscious and the self-concept. In J. Suls & A. Greenwald (Eds.), *Psychological perspectives on the self* (Vol. 2, pp. 219–247). Hillsdale, NJ: Erlbaum.

Epstein, S., Pacini, R., Denes Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology, 71*(2), 390–405.

Evans, J. S. B. T., & Over, D. E. (1999). *Rationality and reasoning*. Hove, UK: Psychology Press.

Fu, Q., Fu, X., & Dienes, Z. (2008). Implicit sequence learning and conscious awareness. *Consciousness and Cognition, 17*(1), 185–202.

Gray, J. A. (1995). The contents of consciousness: A neuropsychological conjecture. *Behavioral and Brain Sciences, 18*(4), 659–722.

Higham, P. A., Vokey, J. R., & Pritchard, J. (2000). Beyond dissociation logic: Evidence for controlled and automatic influences in artificial grammar learning. *Journal of Experimental Psychology: General, 129*(4), 457–470.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*, 306–340.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General, 118*(2), 126–135.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*(5), 513–541.

Jiménez, L. (Ed.). (2003). *Attention and implicit learning.* Amsterdam: Benjamins.

Kinder, A., & Assmann, A. (2000). Learning artificial grammars: No evidence for the acquisition of rules. *Memory and Cognition, 28*(8), 1321–1332.

Kinder, A., Shanks, D. R., Cock, J., & Tunney, R. J. (2003). Recollection, fluency, and the explicit/implicit distinction in artificial grammar learning. *Journal of Experimental Psychology: General, 132*(4), 551–565.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 169–181.

Kuhn, G., & Dienes, Z. (2005). Implicit learning of nonlocal musical rules: Implicitly learning more than chunks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1417–1432.

Lau, H. C. (2008). A higher order Bayesian decision theory of consciousness. *Models of Brain and Mind: Physical, Computational and Psychological Approaches, 168*, 35–48.

Lotz, A., & Kinder, A. (2006). Transfer in artificial grammar learning: The role of repetition information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 707–715.

Meulemans, T., & Van der Linden, M. (1997). Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(4), 1007–1028.

Norman, E., Price, M. C., & Duff, S. C. (2006). Fringe consciousness in sequence learning: The influence of individual differences. *Consciousness and Cognition, 15*, 723–760.

Norman, E., Price, M. C., Duff, S. C., & Mentzoni, R. A. (2007). Gradations of awareness in a modified sequence learning task. *Consciousness and Cognition, 16*, 809–837.

O'Regan, J. K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939–973.

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology, 76*(6), 972–987.

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT Press.

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences, 25*(3), 297–388.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*(6), 855–863.

Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 88–94.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219–235.

Reber, A. S., & Lewis, S. (1977). Implicit learning: An analysis of the form and structure of a body of tacit knowledge. *Cognition, 114*, 14–24.

Redington, M., Friend, M., & Chater, N. (1996, July). *Confidence judgments, performance, and practice, in artificial grammar learning.* Paper presented at the Eighteenth Annual Conference of the Cognitive Science Society, Mawah, New Jersey.

Rosenthal, D. (1986). Two concepts of consciousness. *Philosophical Studies, 49*(3), 329–359.

Rosenthal, D. M. (2005). *Consciousness and mind.* Oxford, England: Clarendon.

Scott, R., & Dienes, Z. (2008). The conscious, the unconscious, and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1264–1288.

Scott, R., & Dienes, Z. (2009). Fluency does not express implicit knowledge of artificial grammars. *Manuscript submitted for publication.*

Servan Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(4), 592–608.

Shanks, D. R. (2005a). Connectionist models of basic human learning processes. In G. Houghton (Ed.), *Connectionist models in cognitive psychology* (pp. 45–82). Hove, UK: Psychology Press.

Shanks, D. R. (2005b). Implicit learning. In K. Lamberts & R. Goldstone (Eds.), *Handbook of cognition* (pp. 202–220). London: Sage.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral & Brain Sciences, 17*, 367–447.

Sun, R. (2002). *Duality of the mind.* Mahwah, NJ: Erlbaum.

Tanaka, D., Kiyokawa, S., Yamada, A., Dienes, Z., & Shigemasu, K. (2008). Role of selective attention in artificial grammar learning. *Psychonomic Bulletin and Review, 15*, 1154–1159.

Wan, L. L., Dienes, Z., & Fu, X. L. (2008). Intentional control based on familiarity in artificial grammar learning. *Consciousness and Cognition, 17*(4), 1209–1218.

Wilkinson, L., & Shanks, D. R. (2004). Intentional control and implicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 354–369.

Chapter 4
Fringe Consciousness: A Useful Framework for Clarifying the Nature of Experience-Based Metacognitive Feelings

Elisabeth Norman, Mark C. Price, and Simon C. Duff

1 Introduction

This chapter discusses how research on the type of metacognitive monitoring referred to as *experience-based metacognitive feelings* (Koriat, 2000, 2007) can benefit from integration of ideas and methods from recent research on the Jamesian concept of *fringe consciousness* (James, 1890).

What is special about experience-based metacognitive feelings is that they occur in relationship to *implicit/unconscious* cognitive activity. Thus, the concept is narrower than Flavell's (1979) concept of metacognitive experiences which also refers to feelings reflecting conscious aspects of ongoing cognition. A large body of contemporary metacognition research has explored various subclasses of experience-based feelings, including tip-of-the-tongue states (for an overview, see Brown, 1991 and Schwartz, 2002), the feeling-of-knowing judgement (for an overview, see Metcalfe, 2000), feelings-of-difficulty in learning and problem solving (Efklides, 2001; Chap. 9, this volume) and judgements of the future memory for items during study – so-called judgements-of-learning (see, e.g., Nelson & Dunlosky, 1991). The concept of experience-based metacognitive feelings was introduced by Koriat (2000, 2007) to distinguish this class of metacognition from information-based metacognition. The latter refers to judgements that are based on a person's explicit beliefs or theories about the capacities and limitations of their cognitive processes. One example is the subjective estimate of how much time or effort one needs to invest in the task at hand (Efklides, 2006). This class of metacognition reflects conscious analytic processes associated with goal-directedness and self-control (Koriat, 2000). In contrast, so-called experience-based metacognitive

E. Norman (✉)
Faculty of Psychology, University of Bergen and Haukeland University Hospital, Christies gate 13, 5020 Bergen, Norway
e-mail: Elisabeth.Norman@psysp.uib.no

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_4,

feelings are described as feelings that reflect unconscious-automatic processes (Koriat, 2000, 2007). One of their characteristics is that they involve dissociations between subjective and objective indices of knowing (Koriat, 2000). Even though there is conscious awareness of a feeling, the feeling derives from cognitive processes of which the person is not currently consciously aware. The subjective feeling can be seen as the "end product of processes that lie below awareness" (Koriat, 2007, p. 298).

The phenomenal qualities of experience-based metacognitive feelings are different from other conscious experiences. From a first-person perspective it *feels* like they involve little deliberate cognitive mediation/evaluation. The felt quality is, instead, one of a "hunch" or a "direct, unmediated experience" (Koriat, 2000); it is, sometimes, characterised as a "fleeting" event (Koriat, 2000). It has been suggested that the validity of the feeling is taken for granted by the person – it has a "self-evident quality" (Koriat & Levy-Sadot, 1999, p. 488).

The neighbouring concept of fringe consciousness originated in the writings of James (1890). He referred to the transitive, fleeting and inarticulate content of conscious experiences as "the fringe", which stands in contrast to the substantive, lingering and articulate aspects of consciousness that he referred to as "the nucleus". One of James' examples is the anticipatory feeling that sometimes arises as a person tries, with initial lack of success, to recall a forgotten name. Another classic example is the feeling of anticipation in a train of thought. In spite of James' emphasis on the importance of "re-instatement of the vague to its proper place in our mental life" (James, 1890, p. 254), the concept was largely neglected for nearly a century, until it was re-introduced and elaborated by Mangan (1993a, b, 2001, 2003), who referred to it as "fringe consciousness".

Koriat (2000, 2007) broadly defines experience-based metacognitive feelings as reflecting unconscious cognition. However, his more detailed descriptions of this class of feelings seem restricted to conscious summary signals driven by heuristics which reflect processes of memory retrieval but are not themselves conscious. Evaluation of knowledge *content* is presented as a conscious activity which characterises so-called information-based judgements. By contrast, fringe consciousness is assumed to also involve evaluation of the content of unconscious knowledge, henceforth referred to as *implicit content*, and has the functions of summarising aspects of this knowledge in a way that may reveal properties of knowledge contents and facilitate their retrieval.

Before comparing fringe consciousness and experience-based metacognitive feelings in more detail, we first address an ambiguity in Koriat's theoretical descriptions and empirical studies of one type of experience-based feeling, namely the feeling-of-knowing. We then argue that integrating the concept of experience-based metacognitive feelings with ideas and operational criteria from research on fringe consciousness can help resolve this ambiguity. In addition, we argue that such integration can also broaden the scope of theoretical accounts of feeling-of-knowing.

1.1 An Ambiguity in Koriat's Treatment of Experience-Based Feelings

Feeling-of-knowing refers to the feeling of having potential access to a certain piece of knowledge that is not currently retrieved. The classic procedure for measuring feeling-of-knowing is the recall-judge-recognise (RJR) procedure (Hart, 1965). Participants are first presented with a set of general knowledge questions or questions about newly learned knowledge. For the subset of items for which the participant is unable to provide the correct answer, s/he is required to rate the likelihood of recognising the correct item if later presented with a set of alternatives on a forced-choice test. This rating is taken as the measure of feeling-of-knowing. Participants can also be asked to report partial knowledge, which refers to fragmentary target knowledge such as the target's initial letter, or knowledge that is semantically related to the searched-for target.

A central assumption made by Koriat is that experience-based feelings do not reflect *implicit content*, but merely the implicit evaluation of various aspects of the *retrieval process*. This would mean that in a situation where explicit retrieval fails (as in the tip-of-the tongue or feeling-of-knowing), but where there is either unconscious activation of the memory target or some degree of activation of relevant partial knowledge, the type of knowledge should not influence the quality or strength of the feeling. For example, the feeling should not be influenced by whether activated knowledge is correct or incorrect. Instead, the feeling would simply reflect various aspects of the retrieval process, including processing fluency, retrieval fluency, cue familiarity, amount of accessed information and subjective confidence (Koriat, 2000). These are "contentless mnemonic cues that pertain to the quality of processing, in particular, the fluency with which information is encoded and retrieved" (Koriat, 2007, p. 298).

A related assumption is that when explicitly retrieved partial knowledge, henceforth *explicit content*, is taken into account, what was previously an experience-based feeling now turns into more of an information-based judgement. The monitoring process then "changes its quality from an automatic, nonanalytic process to a deliberate, inferential process of probability estimation" (Koriat, 1993, p. 632). A feeling-of-knowing that is influenced by deliberate evaluation of explicit content rather than implicit evaluation of the retrieval process is therefore no longer a prototypical experience-based feeling. Instead, it should be rejected as a feeling-of-knowing (Koriat, 1993) or be regarded as a subtype, for example a "late feeling-of-knowing" (Koriat & Levy-Sadot, 1999).

This suggests that evaluation of explicit content could be seen a criterion for rejecting an experience as a true feeling-of-knowing. Therefore, the "contentlessness" of feeling-of-knowing should ideally be tested in situations where partial knowledge is not explicitly retrieved, but where implicit content might play a role. It is then surprising that evidence of the "contentlessness" of feeling-of-knowing largely derives from studies where partial knowledge is either not measured, or is known to be explicitly retrieved, making it difficult to exclude the possibility that the feeling-of-knowing derives from a more deliberate evaluation of explicit content. Take for example studies looking at the so-called *covert accessibility* of partial knowledge, where the to-be-retrieved targets are selected from semantic categories that differ in size (e.g., Koriat & Levy-Sadot, 2001). It has been found that stimuli belonging to larger semantic categories are associated with higher feeling-of-knowing. The explanation is that that the mere amount of knowledge activated during memory search will influence feeling-of-knowing and that this amount will, by definition, increase with increasing category size. However, since partial knowledge is not specifically measured, it is impossible to disentangle the possible influence of implicit versus explicit content.

Studies of so-called *overt accessibility* can be equally problematic. These specifically ask participants to report any partial knowledge they might have for targets that are not successfully retrieved (Koriat, 1993, 1995). Feeling-of-knowing is rated at the end of each trial. This recall phase of the experiment is later followed by a recognition task. Koriat found that feeling-of-knowing strength was related to the amount of reported partial knowledge, but importantly it did not seem to matter whether reported partial knowledge was correct or incorrect. Again, it is questionable whether these findings can be generalized to situations where partial knowledge is activated but not explicitly retrieved. The study has also been criticised for artificially increasing the likelihood that participants make deliberate inferences about explicitly retrieved partial knowledge because feeling-of-knowing was rated immediately after partial recall (Norman, Price, Blakstad, Johnsen, & Martinsen, 2009).

The question of whether implicit content can influence the quality of feeling-of-knowing is complicated even further by a lack of clarity over how to empirically distinguish between situations where explicit content influences the monitoring process and situations where it does not. First, it seems unclear how explicit content is operationally defined. The distinction between implicit and explicit content appears to be somewhere between situations where detailed partial knowledge is verbally reported and situations where partial knowledge is not "articulate enough to support an analytic inference" (Koriat, 2000, p. 159). In other words, explicitness is defined in terms of the verbal report criterion of consciousness. However, it has been suggested that feelings which are vague and difficult to describe verbally might nevertheless be regarded as conscious if they can be expressed on more objective, forced-choice discrimination measures (Price, 2002). Therefore, the distinction implied in Koriat's (2000, 2007) work does not necessarily reflect the degree of conscious accessibility to partial knowledge in an accurate manner. Furthermore, even if it is possible to distinguish explicit from implicit content, it also seems unclear *when* and under which *conditions* explicit content influences monitoring. Koriat (1993) has made one suggestion, namely that content considerations are more likely to come into play when the task is concerned with real-world knowledge as opposed to novel information. However, this hypothesis has not been tested.

In sum, the ambiguity we want to address concerns the lack of clarity over whether feeling-of-knowing can reflect implicit content. In the empirical examples taken to support the "contentlessness" of feeling-of-knowing it seems difficult to exclude the possible confounding influence of explicit content. It is also unclear what the criterion is for defining partial knowledge as implicit or explicit.

Two empirical examples serve to support these concerns by questioning Koriat's position that experience-based feelings do not reflect knowledge content. First, Koriat's central finding that feeling-of-knowing is related to the amount, but not the accuracy of reported partial knowledge (1993, 1995) has proven difficult to replicate. The aforementioned study by Norman et al. (2009) has pointed out that measuring feeling-of-knowing after, rather than before partial knowledge report, might encourage intentional evaluation of explicit content. In addition, using a partial knowledge measure which restricts participants to narrowly report structural fragments of the target, might, in combination with the use of non-word targets as stimuli, artificially increase the influence of incorrect explicit content on feeling-of-knowing (for a more detailed discussion, see Norman et al., 2009). Using a procedure that avoided these possible confounds, it was found that correct partial knowledge influenced feeling-of-knowing more strongly than incorrect partial knowledge. If the correctness of activated partial knowledge is considered an aspect of the content of the knowledge, then these data show feeling-of-knowing is influenced by knowledge content after all. In particular, this was found in an experimental situation which maximised the chances that feeling-of-knowing was experience-based as opposed to information-based.

A second finding that we take to indicate that feeling-of-knowing might reflect implicit content actually comes from Koriat's own research (Koriat, Levy-Sadot, Edry, & de Marcas, 2003). Koriat et al. (2003) studied recall for a set of pseudo-Somali words that each had been paired with a different Hebrew word during an initial training phase. For words that were not successfully recalled during a subsequent test phase, participants nevertheless showed accurate judgements of the *valence* of the words. They rated each word on the dimensions "good-bad", "strong-weak" and "active-passive". These valence judgements were defined as a form of partial recall. However, in our view, they could also be seen as a form of *metacognitive feeling* that accurately reflected implicit content. The interpretation that valence judgements should be seen as experience-based feelings rather than partial knowledge gets some support from the fact that the valence judgements showed little or no decay after 1 week in contrast to recall performance that showed a substantial decay. In our opinion, this experiment not only illustrates the potential influence of implicit content on subjective feelings. It also exemplifies the difficulty of empirically distinguishing between feeling-of-knowing and partial knowledge.

We now turn to whether the concept of fringe consciousness can help resolve this ambiguity in Koriat's treatment of experience-based feelings.

2 The Concept of Fringe Consciousness

According to James (1890), the stream of consciousness contains more than just the clear experiences of our focal awareness. It also contains a *fringe* of more transient, vague and elusive experiences. His central assumption is that this fringe of consciousness binds focally attended sensory information, referred to as the *nucleus*, to relevant contextual background information. Even though fringe consciousness is by definition consciously experienced, it reflects knowledge that is not currently consciously accessible, but nevertheless relevant for our ongoing cognitive processing.

Fringe consciousness is proposed to have a set of special cognitive functions (Mangan, 2001). First, it represents large amounts of information in a condensed format, to avoid exceeding the limited capacity of consciousness – referred to as a summary function. Second, it has a function in monitoring and control of ongoing cognitive activity, that is, a metacognitive function. Finally, it has been suggested that because fringe feelings signal the presence of relevant context information, the person who experiences the feeling may be able to redirect attention to the previously nonconscious source of those feelings – i.e., fringe consciousness also has a retrieval function.

Its constant presence in our stream of consciousness makes fringe consciousness a very wide concept. In previous work we have tried to operationally define a subclass of fringe consciousness that we variously refer to as “cognitive feelings” (Price & Norman, 2009) or “intuitive feelings” (Price & Norman, 2008).

2.1 Similarities Between Fringe Consciousness and Experience-Based Metacognitive Feelings

First consider some suggested properties of fringe consciousness which are shared with experience-based metacognitive feelings, namely that they have a *metacognitive function* in providing a conscious metacognitive signal that reflects unconscious cognitive activity, and are characterised by *vagueness* and *elusiveness*. These and other similarities have been addressed in some detail by Price and Norman (2008).

2.1.1 The Metacognitive Function

Fringe consciousness is assumed to reflect monitoring of ongoing cognitive activity. For example, a *feeling of rightness* might indicate that a flow of thought is coherent (Mangan, 2001). Fringe consciousness is also proposed to have a metacognitive control function in regulating ongoing cognition. Mangan (2003) compares feelings of rightness and wrongness to a “feedback device”, whose role is not only to provide feedback about the current state of cognitive activity, but also to regulate this activity. There is a high degree of overlap between this proposed functional role of fringe consciousness and the concept of *online metacognition,* which refers to the monitoring and control of one’s ongoing cognitive activity (Nelson, 2001), and of which experience-based metacognitive feelings is one example (Koriat, 2000). At a broader level, it has been suggested that metacognitive experiences interact with affect in the self-regulation of behaviour (Efklides, 2001, 2008).

Because fringe consciousness refers to conscious, metacognitive signals that reflect unconscious information-processing antecedents, it can be seen as a *mediator* between conscious and unconscious aspects of cognitive activity. This is very much like experience-based metacognitive feelings, which are consciously experienced even though they relate to *implicit* monitoring of ongoing cognitive activity.

2.1.2 Vagueness

Both fringe consciousness and experience-based metacognitive feelings are characterised by a certain *vagueness*, although there is some disagreement over exactly what is meant here by *vague*. It has been suggested that fringe consciousness is vague in the sense that it is experienced less clearly than other conscious experience (Mangan, 2001), or that it is not claimed to be clearly conscious (Baars, 1993). However, others have argued that fringe consciousness can be experienced both vividly and clearly (Galin, 1993; Price, 2002). The vagueness of fringe consciousness can instead be understood in terms of the experience being difficult to verbally describe to oneself and to others (Price, 2002). In addition, we agree with Galin (1993) that whether an experience is considered vague depends on our purposes for the information it presents. For example, the purpose of a feeling-of-knowing is to retrieve a certain piece of knowledge. A feeling-of-knowing would therefore be regarded as vague because it does not itself contain the to-be-retrieved knowledge. Price and Norman (2008) have suggested a related explanation, namely that the vagueness of fringe consciousness lies in the unexpectedness of the gap between the conscious feeling and its non-conscious information-processing antecedents. Take the situation where seeing a person's face elicits a feeling-of-familiarity but where one fails to retrieve the person's name. If feeling-of-familiarity is normally accompanied by conscious access to the sought-after information, such retrieval failure might be perceived as unexpected (Price, 2002).[1]

Other experience-based metacognitive feelings also seem to be vague in this sense. Consider the tip-of-the-tongue phenomenon, where a person feels that a certain piece of knowledge is potentially available in memory although not currently retrieved. The tip-of-the-tongue experience can be very intense and strong. The subjective feeling can therefore be claimed to be conscious, corresponding to Baars' (1993) operational definition of *non-fringe* consciousness. However, if this experience occurs in a situation where searching for a piece of information normally succeeds, the tip-of-the-tongue might be regarded as vague because it does not provide us with the expected information.

[1]It should be noted that the unexpectedness of the gap between metacognitive feelings and their information-processing antecedents can be experimentally induced. Specifically, in an experimental study exploring metacognition in arithmetic operations (Chap. 9, this volume) illustrate how feelings-of-difficulty increase as the result of cognitive interruption, where a discrepant event blocks the progress toward solution.

2.1.3 Elusiveness and Instability

In the writings of James (1890) the fringe of consciousness is always described in terms of its fleeting or transient quality. This general phenomenological elusiveness is also emphasised by others (Baars, 1988; Mangan, 1993b). Elusiveness refers to the instability of fringe experience which arises because it is short lasting, and because it cannot be introspected upon over time without undergoing some change (Bailey, 1999). According to Mangan (2001) sustained attention will change the phenomenological quality of fringe consciousness.

In research on experience-based metacognitive feelings elusiveness is more often described as a property of the searched-for target information than of the feeling itself (e.g., Koriat, 2000). Nevertheless, research on the feeling-of-knowing judgement indicates that the experience itself also has elusive quality in the sense that it has low stability: The distinction between rapid and late feeling-of-knowing suggests that the nature of the feeling (and its determinants) might change over time, and that different stages of the feeling-of-knowing might even be best understood as consisting of a series of different feelings. Whether the change in phenomenology is the result of holding the feeling in focal attention is an empirical question.

In spite of these similarities, there are also some differences in the way elusiveness is treated within the two research traditions. Whereas the fringe consciousness framework acknowledges that some feelings are too short-lived and fleeting to be introspected (Norman, 2002), the possibility of introspectively reporting experience-based feelings like the feeling-of-knowing seems to be taken for granted in the metacognition literature. This will be addressed in the next section, where we address differences between fringe consciousness and experience-based feelings.

2.2 Differences Between Fringe Consciousness and Experience-Based Metacognitive Feelings

Three important differences between fringe consciousness and experience-based metacognitive feelings are that fringe consciousness is assumed to reflect not just nonconscious processes but also *nonconscious content*, that it has a wider set of *cognitive functions*, and that the degree of *introspective access* is regarded as a variable rather than static property.

2.2.1 Fringe Consciousness Can Reflect Implicit Content

The nonconscious information-processing antecedents of fringe consciousness are normally referred to as *nonconscious context information*. Baars (1988, p. 139) describes this context broadly as "the *inner* world that shapes our experience". The core idea is that fringe feelings can be influenced by a broad array of implicit cognitive activity, including *implicit content*. This is different from Koriat's (2000, 2007) perspective, where it is assumed that experience-based feelings do not reflect implicit content. However, it is compatible with Efklides' (2006) model, according to which metacognitive experiences can sometimes reflect nonconscious evaluation of intrinsic features of the learning material.

A central assumption in the fringe consciousness framework is that properties of the information-processing antecedents of a fringe feeling – including their content – actually influence the *quality* of the resulting fringe feeling. More specifically, it has been suggested that the quality of the feeling can provide information concerning the type of implicit content involved: McGovern (1993) suggests a link between emotional feelings and goal contexts, feeling-of-knowing and conceptual contexts, and feeling-of-familiarity and autobiographically specific knowledge. From this perspective, the nature of relevant implicit content can be inferred from the quality of the fringe feeling, even though it is not accessible in detail at the time the feeling occurs. Thus the influence of implicit content on fringe feelings is functionally important because it contributes to their summary function.

An example indicating that unconsciously represented knowledge might be reflected in fringe consciousness comes from research on implicit learning (Norman, Price, Duff, & Mentzoni, 2007). In a version of the well-known serial reaction time (SRT) task (Nissen & Bullemer, 1987), participants learned to make fast motor responses to indicate a series of positions of a target figure that moved between four positions on a computer screen according to a complex rule. Participants' reaction time patterns indicated that they had learned the complex rule. In traditional SRT tasks, the movement of the target is the only aspect of the stimulus displays that varies between individual trials. We added random variation in the colour and shape of targets and target position indicators. This allowed participants to develop incorrect hypotheses about the rule, for example, that the movements of the target were related to previous colours, shapes, or a combination. It then became possible to separate between those participants who had conscious awareness of the general nature of the rule and those who did not. It was found that even participants who expressed no awareness of the general nature of the rule underlying sequence movements, nevertheless showed behavioural responses that indicated knowledge of the hidden rule structure. Flexible control over sequence knowledge was measured on a separate task where, on each trial, participants had to predict the next target position in a short sequence of target movements. Participants indicated the location which was rotated from the location of the actual target by one or two positions around the square layout, as instructed by a number cue presented centrally *after* the end of the presented sequence. This cue, which varied randomly from trial to trial, could be positively (+1, +2) or negatively (−1) signed, indicating clockwise or anticlockwise rotation, and varied randomly from trial to trial. This task was assumed to require considerable cognitive flexibility. Interestingly, even those participants who expressed no awareness of the general rule could respond in a manner that required conscious, flexible control over what they had learned. Since flexible control is often used as a criterion for consciousness (Baars, 1988), it was concluded that behaviour was not guided by nonconscious automatisms. Rather, it was guided by genuinely conscious feelings that reflected rule knowledge whose details were not themselves consciously accessible, that is, implicit content. According to a related theoretical framework (Dienes & Scott, 2005; Scott & Dienes, 2008) this could also be seen as an example of conscious judgement knowledge in the absence of conscious structural knowledge. An example of how the consciousness of judgement versus structural knowledge can be assessed within a different implicit learning paradigm, namely artificial grammar learning, is found elsewhere in this book (Chap. 3, this volume).

2.2.2 Fringe Consciousness May Facilitate Access to Previously Implicit Content

Mangan (1993b, 2001) uses the term *retrieval* to refer to the process whereby information-processing antecedents that are unconscious when the fringe feeling arises might later become accessible to conscious awareness. According to Mangan (1993b), the summary and retrieval functions are closely intertwined: Because fringe feelings signal the presence of relevant context information (cf. the summary function described above), it might enable the person to redirect his/her attention to the previously nonconscious source of those feelings. Whereas Mangan suggests that retrieval of relevant context information occurs in a largely automatic manner whenever attention is directed to the fringe feeling, others (Baars, 1993; May, 2004) describe the retrieval function as a more voluntary and controlled process. For example, Baars (1993) claims that the tip-of-the-tongue state helps maintain the word-retrieval process that first created it. His description of this retrieval process seems to require both sustained effort and attention – the tip-of-the-tongue will endure as long as the person does not give up and is not distracted. A more active process is also implied in May's (2004) interpretation of fringe consciousness as reflecting relevant but unattended representations. Here, fringe consciousness signals the potential availability of certain information of relevance to the current situation, and to which the person could direct attention if s/he wished.

2.2.3 Introspective Access Is Seen as a Variable Property

As illustrated by a variety of examples in the literature on fringe consciousness (see, e.g., Mangan, 2001), and as pointed out by Price (2002), fringe conscious experience can be hard to describe verbally, both to others and to oneself. Norman (2002) has suggested that the degree of conscious accessibility to fringe consciousness might not be constant, but instead is inversely related to the degree of potential conscious accessibility of information-processing antecedents, which can be understood as a continuum. At one end of the continuum is highly accessible context information, which can easily be brought into awareness, and which is nonconscious only in the sense that it is not currently represented in consciousness in any detail. Fringe feelings relating to this form of context information are of a *fleeting* nature, and cannot be introspectively reported. At the other end is highly inaccessible context information, which in its current form cannot be brought into conscious awareness. However, it can give rise to feelings which have a more long-lasting, *frozen*, nature and which can therefore more easily be introspected.

This view of introspective accessibility as a variable property seems quite different from the view expressed in Koriat's (2000, 2007) writing, where introspective access and verbal reportability of experience-based feelings seem to be taken for granted. However, Norman (2002) suggests that even within traditional experimental paradigms for studying metacognitive feelings, the introspective accessibility might vary. For example, a rapid feeling-of-knowing, collected immediately before the retrieval attempt (Reder & Ritter, 1992), could be seen as an example of a fleeting feeling relating to relatively accessible context information. In contrast, the feeling-of-knowing measured following retrieval failure in the recall-judge-recognise paradigm could be seen as a more frozen feeling relating to context information with an intermediate level of accessibility.

2.3 How the Fringe Consciousness Framework Can Help to Resolve Inconsistencies in Research on Experience-Based Feelings

We have addressed an ambiguity in Koriat's (2000, 2007) treatment of experience-based metacognitive feelings, concerning the proposed role of knowledge content. Even though Koriat's general view seems to be that implicit knowledge content does not influence the nature of the resulting feeling-of-knowing, the empirical examples taken to support this view focus largely on situations where the influence of explicit content cannot be discounted. Moreover, the criteria for explicit versus implicit content are not clearly defined. Finally, certain empirical examples go against Koriat's proposal because they indicate that the nature of experience-based feelings sometimes reflects implicit content. The concept of fringe consciousness outlined above can help to resolve these apparent inconsistencies in the following ways.

2.3.1 Conscious Feelings Can Reflect Implicit Content

First, the fringe consciousness framework clearly states that conscious feelings can be shaped by nonconscious context information, including implicit content. This broader view of conscious feelings, as also reflecting summary representations of implicit content, supplements and extends Koriat's view that experience-based metacognitive feelings can act as a "conscious summary representation of a variety of unconscious processes" (Koriat, 2000, p. 163).

2.3.2 To Test Whether Conscious Feelings Reflect Implicit Content, We Must Measure the Degree of Unawareness of Relevant Knowledge Content

The fringe consciousness framework seems more inclusive than Koriat's (2000, 2007) model in the sense that it also includes implicit knowledge content as a category of information-processing antecedents, in addition to properties of ongoing cognitive processing. However, the specific criteria for what counts as "nonconscious context information" could also be seen as being more conservative within the fringe consciousness framework: Even though a broader set of candidates would be considered as potential information-processing antecedents of a feeling, the information-processing antecedent must fulfil stricter criteria for the experience to be counted as fringe consciousness. Specifically, according to operational definitions of fringe consciousness (Norman et al., 2007; Price, 2002; Price & Norman, 2008, 2009) there must be lack of conscious awareness of the information-processing antecedents of the feeling at the time the feeling occurs.

Therefore, it becomes necessary to develop methods for testing whether relevant knowledge content is truly implicit/unconscious. The serial reaction time experiment on implicit learning that was mentioned above (Norman et al., 2007), measured not only the conscious accessibility of the feeling, but importantly also whether participants were unaware of the nature of the rule knowledge on which the feeling was presumably based. As far as we are aware, this type of measurement is not common in research on experience-based feelings like the feeling-of-knowing. Since research on experience-based feelings seems not to have acknowledged the possible role of implicit content, it is perhaps unsurprising that this type of measurement is rare. What is more surprising is that conscious awareness of the monitoring of the retrieval process is also not measured. Instead, the "implicitness" of this monitoring – which is a crucial aspect of Koriat's model – is simply taken for granted. As a side note, it is also interesting that the amount of retrieved partial knowledge, assumed to be implicitly monitored, is often inferred from the size of the semantic category to which the target belongs rather than being measured directly (Koriat & Levy-Sadot, 2001).

If one is interested in measuring whether feeling-of-knowing is influenced by implicit *content*, as would be hypothesized from a fringe consciousness perspective, then methods for assessing the degree of conscious awareness of such content need to be developed. Perhaps the fringe consciousness framework might also inspire researchers to develop more precise measurements of the extent to which participants are consciously aware of various aspects of their ongoing monitoring.

2.3.3 If Experience-Based Feelings Reflect Implicit Contents They Might Be Seen as Having an Additional Functional Role

If experience-based feelings can reflect implicit content, one implication is that they might have an additional functional role – namely a retrieval function as suggested by Mangan (2001). When large amounts of implicit content are condensed into a consciously experienced feeling, this enables the person to redirect attention to relevant parts of the knowledge domain, referred to by Mangan (2001) as retrieval. This is compatible with previous claims that experience-based feelings not only have a monitoring and control function, but also a *motivational* function in memory search (Koriat, 2000; Smith, 1994), and with recent hypotheses that metacognitive feelings are involved in effort regulation due to their affective nature (Efklides, 2008). Including a retrieval function in a theoretical model of experience-based feelings would also involve more detailed specification of the *nature* of the interplay between conscious and unconscious processes that characterises this subclass of experience.

2.3.4 Specification of the Relationship Between the Accessibility of Knowledge Content and the Nature of the Feeling Has Implications for the Choice of Experimental Approach

The role of accessibility of information plays a central role in Koriat's work. For example, according to the *accessibility account* of feeling-of-knowing, the mere accessibility of relevant knowledge is believed to affect subjective feelings in a memory situation (Koriat, 1993, 1995). Koriat clearly seems to acknowledge that the conscious accessibility of knowledge can vary. However, whereas Koriat seems to discriminate between knowledge that is consciously available (i.e., explicit content) versus knowledge that is not (i.e., implicit content), Norman (2002) discriminates between different levels of *potential* conscious accessibility to unconscious knowledge. Koriat's model does not specify how potential accessibility might influence the nature of the feeling-of-knowing, over and above the proposed direct relationship between conscious accessibility of partial knowledge and feeling strength. For example, Koriat does not address whether the introspective accessibility of different feelings varies. As we pointed out earlier, the lack of clarity over whether inaccessibility to knowledge contents should be seen as a defining property of experience-based feelings is one of the more problematic aspects of Koriat's model.

Norman's (2002) distinctions between subclasses of fringe consciousness and subclasses of nonconscious contexts, based on the relative accessibility of each, might provide some clarification. As mentioned above, Norman (2002) claims there is variation in the degree of conscious accessibility of fringe consciousness. Frozen fringe feelings are relatively accessible to introspective awareness because they can be held in focal attention. They relate to highly inaccessible context information which cannot be brought into conscious awareness without extensive time or effort. Such feelings might be studied in situations of implicit learning (see also Chap. 3, this volume), blindsight and subliminal perception. In contrast, fleeting fringe feelings have a lower intensity, duration and/or degree of specificity. They cannot be accessed directly, but can only be studied by subjective comparison. This is because focally attending the feeling will lead the previously unconscious context information to become consciously available. One example is a feeling-of-wrongness in a patient with anxiety when presented with anxiety-provoking stimulus. Another is the rapid feeling-of-knowing. The traditional feeling-of-knowing following retrieval failure is according to Norman (2002) associated with intermediate level of context accessibility.

Norman's (2002) model can help clarify some of the problematic aspects of Koriat's model in several ways. First, it extends Koriat's (2000, 2007) more general hypothesis that conscious feelings are influenced by the accessibility of relevant knowledge. Norman proposes that a specific property of subjective feelings, namely their degree of introspective accessibility, is influenced by the *potential* accessibility of knowledge. Furthermore, it is specifically suggested that a feeling can qualify as *fringe consciousness* even when the degree of context accessibility is intermediate or high, as long as the feeling is captured before the knowledge itself becomes consciously accessible. This requires that conscious access to relevant context information is assessed at regular intervals throughout the experiment.

This view of context accessibility has implications for what type of feelings the researcher might want to target, and maybe even more important – how feelings should be measured. Feelings relating to highly inaccessible context information are best suited for empirical exploration because they can be more easily introspected and verbalised than other feelings (Norman, 2002). Research addressing *fleeting* feelings, related to relatively more accessible context information, need to use more refined measurements of subjective feelings. The suggestion by Price (2002) that people may sometimes be unable to verbally report their own experience, even though the feelings are consciously experienced, might apply in particular to feelings in the *fleeting* category. For example, in the context of experiments on subliminal perception, transient aspects of visual stimulus displays might prompt a fleeting fringe feeling, but since participants lack confidence in what they see, or lack a framework to categorise and describe their experience, they may report they see nothing. Price (2002) therefore suggests that verbal report measures are supplemented by objective forced-choice discrimination measures which are also less likely to be contaminated by response bias.

2.3.5 Integration of Ideas from the Fringe Consciousness Framework Has Broader Implications for Theoretical Discussions About the Origin of Feeling-of-Knowing

Koriat and Levy-Sadot (2000) make a strong claim about the trace-access hypothesis of feeling-of-knowing. They claim that this hypothesis, which proposes that feeling-of-knowing reflects a direct monitoring of the target information, i.e., implicit content, is not testable: "In fact, we cannot envision an experimental procedure that would allow us to reject the direct-access hypothesis outright" (Koriat & Levy-Sadot, 2000, p. 195). The argument seems to be that as long as it can be shown that feeling-of-knowing reflects the implicit evaluation of retrieval processes, whether or not it is also influenced by activated but implicit content is not central, as it is impossible to separate the influence of content versus process. We disagree with this claim that the influence of monitoring versus knowledge cannot be distinguished. Let us again turn to implicit learning research.

In the example we have presented above (Norman et al., 2007), we showed how implicit learning is a situation where it is possible to assess both the implicitness of content (i.e., rule knowledge) and the consciousness of the subjective feeling (i.e., whether it can be used flexibly). However, how can it be ruled out that the subjective feeling reflects implicit content rather than implicit monitoring of information-processing, for example increased processing fluency for stimuli that follow the learned rule?

A different implicit learning experiment illustrates how this can in fact be done. Dienes, Altmann, Kwan, and Goode (1995) have developed an experimental procedure where participants implicitly learn two complex rules rather than only one. More specifically, participants are presented with two sets of non-word letter strings. Within each set, letter strings are constructed on the basis of an artificial grammatical structure that specifies the ordering of individual letters. In a subsequent test phase, participants are presented with a set of unseen test items that they are asked to classify according to only one of the two grammars. The ability to flexibly control which grammar to apply is indicative of conscious, strategic control. Importantly, since participants have had equal exposure to both sets of rules, flexible control over the two grammars cannot be attributed to fluency effects. The flexible application of one of two rules, in the absence of explicit access to the rules, therefore implies that the hunches which drive performance in the test phase of the experiment are sensitive to implicit content. These hunches can then be seen as manifestation of fringe consciousness.

3 Conclusion

In the first part of this chapter, we pointed to an ambiguity in Koriat's (2000, 2007) treatment of feeling-of-knowing as an experience-based feeling, related to a lack of clarity concerning the role and measurement of implicit content. We then presented a theoretical concept closely related to experience-based feelings, namely fringe consciousness. Even though experience-based feelings and fringe consciousness have a number of characteristic features in common, related to their phenomenological quality and metacognitive role, there are also some central differences. Importantly, fringe consciousness is seen as providing a summary signal of various aspects of implicit cognitive activity that also includes *implicit content*. We suggested that a closer integration with ideas and methodology from research on fringe consciousness might help resolve the ambiguity in Koriat's treatment of experience-based feelings, and thus benefit research on experience-based feelings at a methodological and a theoretical level.

At a methodological level, we have shown that the fringe consciousness framework might be useful in constructing experimental situations that facilitate experience-based feelings. We have also provided examples of how the "implicitness" of implicit content can be empirically measured, and how the influence of implicit monitoring and implicit content can be distinguished. In our view, the most important challenge for future research on experience-based feelings is not how to measure the subjective quality of metacognitive feelings, but how to empirically verify the unconscious nature of their information-processing antecedents. At a theoretical level, we have hypothesized that experience-based metacognitive feelings might have an additional functional role related to the retrieval of previously implicit content. Finally, the central role of implicit content implied in the fringe consciousness framework suggests the need to re-evaluate theoretical models of experience-based feelings that acknowledge the influence of implicit content. One such model is the trace access hypothesis of feeling-of-knowing (Hart, 1965).

Acknowledgements This research was partly supported by a postdoctoral grant (911274) to the first author from the Western Norway Regional Health Authority (Helse Vest).

References

Baars, B. J. (1988). *A cognitive theory of consciousness.* New York: Cambridge University Press.
Baars, B. J. (1993). Putting the focus on the fringe: Three empirical cases. *Consciousness and Cognition, 2*, 126–136.
Bailey, A. R. (1999). Beyond the fringe: William James on the transitional parts of the stream of consciousness. *Journal of Consciousness Studies, 6*, 141–153.
Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109*, 204–223.
Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research, 69*, 338–351.
Dienes, Z., Altmann, G. T. M., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1322–1338.
Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation, and self-regulation. In A. Efklides, J. Kuhl, & R. M. Sorrentino (Eds.), *Trends and prospects in motivation research* (pp. 297–323). Dordrecht, The Netherlands: Kluwer.
Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.
Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*, 277–287.
Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911.
Galin, D. (1993). Beyond the fringe. *Consciousness and Cognition, 2*, 113–118.
Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology, 56*, 208–216.
James, W. (1890). *The principles of psychology.* New York: Holt.
Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*, 609–639.
Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124*, 311–333.

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition, 9*, 149–171.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *Cambridge handbook of consciousness* (pp. 289–325). New York: Cambridge University Press.

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgements: Information-based and experience-based monitoring of one's own knowledge. In Y. Trope & S. Chaiken (Eds.), *Dual-process theories in social psychology* (pp. 483–502). New York: Guildford.

Koriat, A., & Levy-Sadot, R. (2000). Conscious and unconscious metacognition: A rejoinder. *Consciousness and Cognition, 9*, 193–202.

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 34–53.

Koriat, A., Levy-Sadot, R., Edry, E., & de Marcas, S. (2003). What do we know about what we cannot remember? Assessing the semantic attributes of words that cannot be retrieved. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1095–1105.

Mangan, B. (1993a). Some philosophical and empirical implications of the fringe. *Consciousness and Cognition, 2*, 142–154.

Mangan, B. (1993b). Taking phenomenology seriously: The "fringe" and its implications for cognitive research. *Consciousness and Cognition, 2*, 89–108.

Mangan, B. (2001). Sensation's ghost: The non-sensory "fringe" of consciousness. *Psyche, 7*. Retrieved April 28, 2009 from the World Wide Web: http://psyche.cs.monash.edu.au/v7/psyche-7-18-mangan.html.

Mangan, B. (2003). The conscious "fringe": Bringing William James up to date. In B. J. Baars, W. P. Banks, & J. B. Newman (Eds.), *Essential sources in the scientific study of consciousness* (pp. 741–759). Cambridge, MA: The MIT Press.

May, J. (2004). An information processing view on fringe consciousness. *Psyche, 10*, 1–9. Retrieved April 28, 2009 from the World Wide Web: http://psyche.cs.monash.edu.au/symposia/mangan/pdf/may.pdf

McGovern, K. (1993). Feelings in the fringe. *Consciousness and Cognition, 2*, 119–125.

Metcalfe, J. (2000). Metamemory: Theory and data. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 197–211). London: Oxford University Press.

Nelson, T. O. (2001). Psychology of metamemory. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (Vol. 14, pp. 9733–9738). Amsterdam: Elsevier.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect". *Psychological Science, 2*, 267–270.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*, 1–32.

Norman, E. (2002). Subcategories of "fringe consciousness" and their related nonconscious contexts. *Psyche, 8,* 1–15. Retrieved April 28, 2009 from the World Wide Web: http://psyche.cs.monash.edu.au/v8/psyche-8-15-norman.html

Norman, E., Price, M. C., Duff, S. C., & Mentzoni, R. A. (2007). Gradations of awareness in a modified sequence learning task. *Consciousness and Cognition, 16*, 809–837.

Norman, E., Price, M. C., Blakstad, O., Johnsen, Ø., & Martinsen, S. K. (2009). *Measuring the relationship between feeling-of-knowing metamemory judgements and partial knowledge.* Manuscript submitted for publication.

Price, M. C. (2002). Measuring the fringes of experience. *Psyche, 8*, 1–24. Retrieved April 28, 2009 from the World Wide Web: http://psyche.cs.monash.edu.au/v8/psyche-8-16-price.html.

Price, M. C., & Norman, E. (2008). Intuitive decisions on the fringes of consciousness: Are they conscious and does it matter? *Judgment and Decision Making, 3*, 28–41.

Price, M. C., & Norman, E. (2009). Cognitive feelings. In P. Wilken, T. Bayne, & A. Cleeremans (Eds.), *Oxford companion to consciousness* (pp. 141–144). Oxford, UK: Oxford University Press.

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 435–451.

Schwartz, B. L. (2002). *Tip-of-the-tongue states: Phenomenology, mechanism, and lexical retrieval.* Mahwah, NJ: Erlbaum.

Scott, R. B., & Dienes, Z. (2008). The conscious, the unconscious, and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1264–1288.

Smith, S. M. (1994). Frustrated feelings of imminent recall: On the tip of the tongue. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 27–45). Cambridge, MA: Bradford.

Chapter 5
Further Insight into Cognitive and Metacognitive Processes of the Tip-of-the-Tongue State with an Amnesic Drug as Cognitive Tool

Elisabeth Bacon

When I remember forgetfulness there are present both memory and forgetfulness, memory, whereby I remember, forgetfulness, which I remember. Then is forgetfulness retained by Memory.

[St Augustine, Confessions]

1 Introduction

Everybody has probably experienced at least once in their life the frustration of being unable to retrieve a particular word at the desired time. To have the word on the *tip of the tongue* (TOT) is a very common oblivion, but it also reveals the awareness we have of the content of our own memory. States such as TOT may be viewed as transient and reversible micro-amnesia commonly affecting healthy people. Might it be possible to use amnesic drugs (e.g., lorazepam) to decipher this phenomenon? As in a TOT state, individuals under lorazepam would momentarily have no access to a known piece of information but would retrieve a word closely related to the target answer and provide it as the response to the question. Moreover, unlike normal people, who are well aware that this information that comes to mind is *not* the correct answer and have a strong sense of having the correct word on the tip of their tongue, we hypothesised that under the effect of lorazepam there would be a dissociation between the cognitive and metacognitive components of the TOT experience, and that lorazepam-treated participants would inappropriately attribute a high confidence rating to the intrusive incorrect word. In this chapter, we present why we were prompted to suggest the TOT model as an explanation for the peculiar pattern of temporary semantic memory/metamemory impairment induced by the amnesic drug lorazepam and we describe in detail how we experimentally verified this hypothesis (Bacon, Schwartz, Paire-Ficout, & Izaute, 2007).

E. Bacon (✉)
Psychiatry Department, National Institute for Health and Medical Research (U 666 INSERM) and University Hospital, BP 426, 67091 Strasbourg Cedex, France
e-mail: bacon@alsace.u-strasbg.fr

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_5,

1.1 Drugs as Tools for Exploring Memory Functioning

Our current knowledge about cognitive processes and functions stems from research performed with different populations (Danion, 1994). The more conventional studies involve healthy participants and highlight fundamental notions common to all. Other studies recruit individuals suffering from traumatic or organic memory pathologies, such as Alzheimer's disease. However, the study of clinical populations is likely to be problematic because the nature and extent of the brain lesion may vary from one individual to the next, and patients may be suffering from additional pathologies or taking drugs that may complicate interpretation of the observations.

A growing number of studies have been conducted originating from a theoretical viewpoint, but involving the administration to healthy participants of amnesic drugs viewed as tools for revealing functional principles of normal cognitive processing. Drugs from the benzodiazepine family were first described as having an amnesic effect in 1965 and are now widely used as tools for the purpose of memory studies (Duka, Curran, Rusted, & Weingartner, 1996). Their amnesic effect, particularly on episodic memory, is well known. The amnesic episode induced by benzodiazepines is transitory, lasting only for a few hours. In the case of episodic memory, healthy participants administered a benzodiazepine experience anterograde amnesia, and it is the acquisition of new information that is impaired by the drug (for review see Beracochea, 2006; Curran, 1999). Lorazepam is particularly interesting as a benzodiazepine, because it has no active metabolites. During the amnesic episode, participants are not aware of their memory deficit, and lorazepam has also been shown to induce some metamemory impairments (Bacon et al., 1998; Izaute & Bacon, 2005).

1.2 Effects of the Amnesic Drug Lorazepam on Semantic Memory

Very few drugs have been shown to alter semantic memory. It has long been taken for granted that benzodiazepines do not alter semantic memory (Curran, 1991, 1999). These conclusions relied mostly on unimpaired performance in verbal fluency tasks, where participants were required to provide as many items as possible from a given semantic category within a set time (Curran, 1991; File, Sharma, & Shaffer, 1992; Fluck, File, Springett, Kopelman, Rees, & Orgill, 1998; Vermeeren et al., 1995). However, findings with sentence verification tasks were found to be contradictory. Allen, Curran, and Lader (1993) and Green, McElholm, and King (1996) found that lorazepam did not affect the accuracy of semantic retrieval, whereas Vermeeren et al. (1995) reported that lorazepam-treated participants made more mistakes in these tasks than placebo participants. In addition, File et al. (1992)

showed that the benzodiazepine midazolam impaired word completion performance, and they observed that participants under benzodiazepine generated more low frequency exemplars than common words when retrieving categorical information from memory, which could be due to the fact that the most common, high frequency answers were temporarily not accessible and that the participant had to recruit more uncommon words from his or her semantic store to complete the task.

Some researchers observed (Bacon et al., 1998) and replicated (Izaute, Paire-Ficout, & Bacon, 2004; Massin-Krauss, Bacon, & Danion, 2002) an impairment in semantic memory, when healthy participants were under the effect of the benzodiazepine lorazepam. In these studies, participants were presented with general knowledge questions (e.g., What is the capital of Greece?) and had to recall the answer (e.g., Athens). Those under the effect of lorazepam produced as many recall answers as the participants under placebo, but gave more incorrect answers (commission errors) (see Table 5.1).

The impairment that lorazepam induces in semantic memory is reversible. Benzodiazepines are the most commonly consumed drugs in the western world because of their effects on anxiety, insomnia and muscle relaxation, and if the semantic impairment was permanent, clinical and daylife observations would have been noticed. However, to confirm the reversibility of the amnesic effect on semantic memory, 2 years after the experiment of Bacon et al. (1998), three participants who had taken lorazepam and one who had taken the placebo were retested in their usual or sober state. The transitory nature of the amnesic episode was confirmed as the performance of the ex-lorazepam participants improved (see Sect. 5.1), whereas that of the ex-placebo participant was similar to his performance 2 years earlier. Furthermore, lorazepam participants were more likely to experience a common

Table 5.1 Means (and *SD*) of performance on free recall and recognition tasks and of confidence level accuracy in the placebo and lorazepam groups (adapted from Bacon et al., 1998) and gamma correlations

	Group	
	Placebo	Lorazepam
	M (*SD*)	*M* (*SD*)
Free recall task		
Proportion of answers	61 (14)	57 (16)
Proportion of correct answers	82 (8)	60 (12)*
Recognition task		
Proportion of correct answers	58 (9)	49 (12)
Confidence level accuracy	82 (11)	80 (9)
For correct answers	87 (8)	88 (7)
For incorrect answers	57 (14)	68 (12)**
Gamma correlation		
Between confidence level accuracy and recall performance	0.65	0.61

*Significant difference at $p<0.05$

**Marginally significant difference at $p=0.07$

semantic illusion, the "Moses Illusion"[1] (Erickson & Mattson, 1981; Reder & Kusbit 1991). They also provided more incorrect recalls for the filler questions in the Moses paradigm, and this observation is an additional argument in favor of the existence of an impairment of semantic memory induced by lorazepam (Izaute et al., 2004).

It must also be borne in mind that the pattern of cognitive impairment induced by benzodiazepines may vary from one molecule to the next (Giersch, Boucart, Elliott, & Vidailhet, 2010). Mintzer, Kleykamp, and Griffiths (2010) observed that another benzodiazepine, triazolam, had no effect on performance in a general information task. The pattern of semantic memory impairment induced by lorazepam also differs from that induced by another potentially amnesic drug, ethanol. In a general information task, healthy participants under the effect of ethanol produced fewer recall answers compared to placebo participants (Nelson, McSpadden, Fromme, & Marlatt, 1986), whereas under lorazepam participants provide the same number of recall answers, but with a higher error rate.

1.3 The Peculiar Pattern of Memory/Metamemory Impairment Induced by Lorazepam for Semantic Memory

In their study, Bacon et al. (1998; see Table 5.1) used the classic recall-judgment-recognition paradigm (Hart, 1965). Participants were 12 placebo and 12 lorazepam (0.038 mg/kg) indivilduals. They were presented with 120 general knowledge questions and asked to recall the answers. For each answer they provided, they had to rate their retrospective confidence level that the answer given was correct. The lorazepam-treated participants seemed to selectively overestimate their retrospective confidence level for incorrect recalls, which was marginally higher than that of the placebo particpants, $t(22) = -1.91$, $p = 0.07$. However, the treated participants were still able to discriminate between correct and incorrect answers, as their confidence was higher for correct than for incorrect answers, and their gamma correlations between confidence levels and free recall performances were no different from those of the participants who had received a placebo. Thus, the drug seemed to induce a selective impairment of their monitoring ability. The same pattern of a higher rate of incorrect recall coupled with an overestimated confidence level for incorrect recalls ($p < 0.001$) and preserved monitoring accuracy was also observed in the context of a forced-recall task with respect to general knowledge questions (Massin-Krauss et al., 2002).

Evidence of impaired recall performance in a general knowledge task suggests that the control process might be impaired too. The drug might have altered the way participants make decisions and may have induced a desinhibitory state leading

[1]The Moses illusion is as follows: When asked "How many animals of each kind did Moses take on the ark?" people fail to notice the distortion introduced by the impostor "Moses" and respond "two". This semantic illusion, which is known as the Moses illusion, has proved to be quite robust and can be generalized across other materials and conditions.

them to output answers they might otherwise keep to themselves. In that case, one would expect lorazepam participants to provide more recall answers than placebos in a free recall task. However, the number of answers produced by the lorazepam participants was no different to the number provided by the placebos (Bacon et al., 1998), casting some doubt on this view. Furthermore, the effects of lorazepam on the processes involved in the strategic regulation of memory accuracy (Koriat & Goldsmith, 1996) have been investigated (Massin-Krauss et al., 2002). Control sensitivity, that is, the extent to which volunteering an answer is affected by confidence judgments, was only slightly impaired by the drug. Consequently, defective control sensitivity cannot explain all the extra commission errors produced under the effect of lorazepam in a semantic task.

Within the context of a reversible semantic memory impairment and relative preservation of decision-making, it is possible to re-phrase the question of why lorazepam participants provide an incorrect recall when they actually know the right answer, that is, under what circumstances is the memory of healthy participants temporarily impaired to such an extent that they are prompted to give an incorrect answer when they actually know the correct answer? There are everyday situations where individuals do behave in this manner, that is, when they are in a TOT state. Specifically, when a person is experiencing a TOT state, she or he cannot retrieve a known piece of information, and sometimes the TOT state is accompanied by an intrusive incorrect blocking word.

2 The TOT State

In everyday life we may all experience ordinary memory defects (Schacter, 1999), that may be either permanent or transient. The blank-in-the-mind (BIM) experience is a very common memory failure (see Efklides & Touroutoglou, Chapter 6). One of the most spectacular transitory memory impairments is probably the TOT state. The TOT state may occur for both semantic and episodic memory (Schwartz 1998; Schwartz, Travis, Castro, & Smith, 2000). When a person is experiencing a TOT state, she or he cannot retrieve a known piece of information. At the same time, the person has the strong and frustrating feeling that the missing target word is on the verge of being retrieved (Schwartz, 2002a, b). Schwartz (1999) wondered whether this experience is really universal and whether the "tongue" metaphor is used in other languages too to describe this peculiar state. He observed that, out of 51 languages, as many as 45 employed an expression using the "tongue" metaphor to describe this feeling of not being able to retrieve a known word. Brennen, Vikan, and Dybdahl (2007) observed that speakers of an unwritten Guatemalan language were able to recognize a description of the phenomenology associated with tip-of-the-tongue states and that TOT states could also be induced in this particular group of participants.

A TOT state is a relatively stressful and emotional situation often coupled with a feeling of frustration. The TOT state is a transitory state of inaccessibility of a known piece of information and accurate predictor of later recall and recognition (Schwartz, 2002b), that is, when rememberers experience TOT states, they are likely to retrieve the correct answer eventually, since 89–95% of the missing words are subsequently retrieved (Burke, MacKay, Worthley, & Wade, 1991; Schwartz, 2002a; Schwartz et al., 2000; for reviews see also Brown, 1991; Schwartz, 2002b).

Diary studies and laboratory tasks also show that 50–70% of TOT states are accompanied by intrusive blocking words, also known as "interlopers", or persistent alternates (Burke et al., 1991; Reason & Lucas, 1984). For example, in diary studies, Reason and Lucas (1984) found that over 50–70% of the resolved TOT states were preceeded by intrusive blocking words. Burke et al. (1991) observed that in a sample of young adults 67% of the TOT states were accompanied by what they called "persistent alternates" – the term we will use throughout this chapter. Furthermore, Burke et al. (1991) found that nearly 90% of the persistent alternates were from the same syntactic category as the missing world. These alternates were recognized as incorrect by the participants, who, however, were unable to retrieve the correct target in the meantime. Laboratory studies show higher rates of both resolution and persistent alternates among TOT states than among non-TOT states (Smith, 1994). Recognition of the correct target following a TOT state is much more likely than recognition of the correct target when rememberers are not experiencing a TOT (Schwartz, 1998, 2001; Schwartz et al., 2000; Schwartz & Smith, 1997; Smith, 1994).

The phenomena underlying TOT experiences are at the intersection between memory, language, and metamemory models and have been the subject of numerous studies by researchers from various disciplines (Schwartz, 1999, 2001). For psycholinguists and memory theorists, the TOT state and word retrieval are triggered by the same retrieval process. TOT states are interesting because they serve as "windows" to the retrieval process (Biedermann, Ruh, Nickels, & Coltheart, 2008; Brown, 1991). The metacognitive view is that TOT state and retrieval process are dissociable (Schwartz, 2001). The TOT state is classifiable as a *metacognitive judgment*, whereas retrieval is a *cognitive process*.

3 TOT as a Cognitive and Metacognitive Experience

We shall focus here on the metamemory perspective regarding the TOT state. Schwartz et al. (2000) used Nelson and Narens' (1994) model to explain the TOT state. Object-level cognition (encoding, imaging, retrieving…) is separate from meta-level cognition (feeling of knowing or judgment of learning). Monitoring is the flow of information from the object-level to the meta-level, and control is the flow of information from the meta-level to the object-level. The TOT state plays a monitoring role by informing rememberers when an item may be retrievable. It may serve to alert the rememberers that more time may be needed to retrieve an item and to warn them not to terminate the search. Thus, it provides rememberers with useful information that can then be used to control mnemonic behaviour. The TOT state differs from a strong feeling-of-knowing judgment, because different brain areas are activated during TOT states and feeling-of-knowing judgments (Maril, Simon, Weaver, & Schacter

2005). Moreover, by manipulating working memory load during retrieval of general knowledge questions, Schwartz (2008) obtained data supporting the view that a TOT state and a feeling-of-knowing judgment are separable metacognitive entities.

We shall distinguish between two aspects of the TOT experience. Firstly, the cognitive state of TOT is defined as the failure of the retrieval process to produce a known word (Burke et al., 1991; Miozzo & Caramazza, 1997; Vigliocco, Antonini, & Garrett, 1997). This cognitive process is about word retrieval and the failure of that process. On the other hand, the phenomenological experience of a TOT state will be defined as the strong and frustrating feeling that a particular target word is on the verge of being retrieved (Brown & McNeill, 1966; Schwartz et al., 2000). This experience is a metacognitive one, as it involves a feeling of future memorability.

From a study of the literature, it would appear that the research conducted to date may support such a distinction between cognitive and phenomenological TOT states (Schwartz, 2002b). Research suggests that not all temporary retrieval failures are accompanied by a TOT state and that not all phenomenological TOT states are accompanied by the eventual retrieval of a target (Schwartz, 1998; Schwartz et al., 2000). Furthermore, research has demonstrated dissociations between retrieval and the number of TOT states (Schwartz & Smith, 1997; Widner, Smith, & Graziano, 1996). Throughout this chapter, we use the term "cognitive TOT" to refer to the temporary amnesia associated with a known word retrieval failure and the term "phenomenological TOT" to refer to the subjective experience of feeling that a word is retrievable.

A TOT state reveals a conflict between the metacognitive judgement, that is, the certainty that the information is known, and the cognitive level, that is, the temporary inability to retrieve a known target from long-term memory. The TOT state is regarded as the slowing down of a memory process and may be viewed as momentary and reversible "micro-amnesia" occurring naturally and occasionally in healthy people.

To summarize, TOT states appear to be very common in everyday life, are quite similar across language groups, and easy to induce in laboratory. Participants in a TOT state are momentarily unable to retrieve a known piece of information and may sometimes provide an incorrect answer, referred to as a "persistent alternate". And this is exactly what we hypothesized that has happened in participants under the effect of lorazepam with some items in a general knowledge task.

4 The Amnesic Effect of Lorazepam on Semantic Memory and TOT State

4.1 Similarities and Divergences Between Lorazepam-Induced Amnesic Episode and Naturally Experienced TOT

Under the effect of lorazepam, as well as when naturally experiencing TOT, healthy participants are temporarily unable to retrieve some known information and sometimes retrieve information closely related to the target answer. There are striking differences, however, between what occurs in individuals experiencing TOT in everyday life and what occurs in participants under the effect of the benzodiazepine. First, when in a TOT state, and if a persistent alternate comes to mind, healthy individuals in an undrugged TOT state recognize this information as not being the correct answer. They have the feeling that the persistent alternate impedes their access to the correct answer, and they also feel very strongly that they know the target answer, and that retrieval of the target is imminent. Under lorazepam, however, participants do not recognize the incorrect item that comes to mind as being incorrect and so seem not to have the phenomenological experience of recognizing the incorrect item as a persistent alternate. They do not experience the phenomenological TOT. Second, the effect of lorazepam on participants increases the likelihood that they will give an incorrect answer despite knowing the correct response, thereby making a commission error.

We wondered then whether individuals under the effect of lorazepam could sometimes experience a kind of "dissociated" TOT. Specifically, they would experience the cognitive TOT (i.e., the correct target would be momentarily inaccessible and a persistent alternate would come to mind) but they would not reject the persistent alternate as such and would provide it as the target answer without recognizing the blocker nature of this response; in the meantime they would not spontaneously experience the phenomenological TOT, that is, the feeling that the correct answer is on the verge of being retrieved. We suggest that monitoring would be impaired, in that participants would not experience the TOT phenomenology, but that monitoring effectiveness (the ability to distinguish between correct and incorrect answers) would be preserved, insofar as the participants are still able to recognize the correct answer among distractors. Lorazepam would impair control at the point in time when they have to provide an answer (selection of the correct target in the recall step). This dissociation between monitoring and control has already been observed in the strategic regulation of memory accuracy under the effect of lorazepam (Massin-Krauss et al., 2002). Moreover, various patterns of memory and metamemory dysfunctioning have been reported in patients with brain lesions (Bäckman & Lipinska, 1993; Janowsky, Shimamura, & Squire, 1989a, b; Nelson et al., 1986; Pappas et al., 1992; Shimamura 1994; Shimamura & Squire, 1986, 1988). They suggest that memory and metamemory are not inextricably linked. The possible dissociation of the cognitive and the phenomenological TOT has already been evoked in the literature, that is, experience of the phenomenological TOT without its subsequent resolution has been referred to as "subjective TOT" by Jones and Longford (1987) or "negative TOT" by Vigliocco et al. (1997). On the other hand, commission errors followed only later by the phenomenological TOT were referred to as "commission TOT" by Schwartz et al. (2000). This is what we suspect occurs under the effect of lorazepam. We hypothesized that the incorrect recall answers provided by participants having experienced a commission TOT are the "blockers" or "persistent alternates" often retrieved by participants experiencing a natural TOT state.

A TOT experience is also a relatively stressful feeling, often accompanied by a sense of frustration, and has been shown to have an emotional dimension, that is, the "phenomenal TOT" (Schwartz et al., 2000). Benzodiazepines act as anxiolytic drugs and have anti-conflict effects (Harvey, 1980; Kleven & Koek, 1999; Vanover, Robledo, Huber, & Carter, 1999). Consequently they may have an effect that attenuates the stressful, phenomenal aspect of the cognitive conflict elicited by a TOT state. As a result, we suggest that drugged participants would honestly provide the persistent alternate as a convenient answer to the question asked and would not feel that they were on the verge of recognizing the correct answer. According to this interpretation lorazepam should reduce the phenomenological TOT experiences (Massin-Kraus et al., 2002) while at the same time increasing the number of retrieval failures, as a result of incorrect reporting of persistent alternates.

To confirm this hypothesis, we investigated the effects of lorazepam on TOT states by using the drug as a pharmacological tool that should allow us to gain some insight into this phenomenon. We wanted to show that, in some cases, the phenomenological TOT does not occur until after participants have found out that their retrieval was inaccurate. Thus, we shall argue in favour of a TOT model that distinguishes between cognitive and metacognitive (phenomenological) aspects of the TOT process (for a different view, see Taylor & MacKay, 2003).

4.2 Evidence for the TOT Model

Before exploring experimentally the effects of lorazepam on TOT states, we analysed unpublished data from Bacon et al. (1998) and Massin-Krauss et al. (2002) to examine a number of memory and metamemory features that might consitute additional cues in support of our hypothesis about the effects of lorazepam on TOT states.

First, we observed that lorazepam participants are able to experience the phenomenology of TOT in some error trials (data from Massin-Krauss et al., 2002), but to a lesser extent than the placebo participants. In that particular experiment, in the stressful situation of forced report recall of semantic memory, participants under lorazepam (0.038 mg/kg) reported an average of 2.3 TOT states (out of a total of 120 questions), which was a significantly lower rate of TOT states than that of the placebo participants ($M=4.9$, $p=0.037$). This observation lends weight to the hypothesis that the anxiolytic effect of lorazepam might have an effect on the number of TOT states reported.

Another of our aims was to determine whether the incorrect recalls provided by the participants under lorazepam were similar to the persistent alternates found in TOT studies. To that end, we examined the nature of the incorrect recalls provided by the participants in the Bacon et al. (1998) study. The drug had no effect on the mean number of recall answers given, but increased the number of incorrect answers. The commission errors were analyzed for the lorazepam 0.038 mg/kg group according to four criteria: (a) semantic substitution (oenologist instead of wine waiter); (b) phonological or semantico-phonological substitution, for example, *faines* (beechnuts) instead of *fanons* (whalebones); (c) perseverative errors (the answer was a word used in the question); and (d) commission errors with no apparent link or invented words. The majority of errors were semantic, with 237 in the lorazepam group and 118 in the placebo group. The proportion of semantic errors was 76% for the lorazepam participants and 80% for the placebo participants. This difference was not statistically significant, $t(22)=1.3$, $p=0.19$. Phonological substitution counted for only 4% of the errors for lorazepam and 6% for the placebos, with no significant difference as a function of the treatment, $t(22)=1.4$, $p=0.18$. Perseverative error scores differed as function of the treatment for the commission errors. Specifically, participants under lorazepam had a higher perseverative error score (16%) than placebo-treated participants (9%), $t(22)=2.8$, $p<0.02$. Finally, there were some errors that were without any apparent link. Lorazepam-treated participants made no more errors in this category than placebos (4 and 5%, respectively), $t(22)=0.9$, $p=0.40$, and the overall rate was very low. So, taken as a whole, these results show that participants under lorazepam make commission errors similar to the persistent alternates observed in the TOT literature (Burke et al., 1991; Harley & Bown, 1998).

Also, in an item-by-item analysis, the total number of questions (out of the 120 questions asked) that produced at least one incorrect recall answer was higher in the lorazepam group than in the placebo group: taken together, the 12 participants in the lorazepam (0.038 mg/kg) group made commission errors out of a set of 105 questions and gave 235 different wrong answers to this set of questions, whereas the 12 placebo participants gave only 106 different wrong answers to 67 questions taken from the entire set of 120 questions they were asked to answer. So more questions were likely to elicit recall errors from lorazepam participants, and the range of possible incorrect answers was more diverse. This shows that under lorazepam the questions lead to the retrieval of information relevant to the target (Koriat, 1995), and that lorazepam participants do not inhibit incorrect answers as and when they are retrieved. It must be borne in mind, however, that the individual number of recall answers given by each participant did not vary with the lorazepam intake. Consequently, lorazepam may exert a cognitive disinhibition, prompting participants sometimes to provide an incorrect rather than correct answer, but not a behavioral disinhibition, which would have caused all of them to provide more recall answers than the placebo participants. The observation that participants under the effect of lorazepam are also more sensitive to the Moses effect and more often make partial matching also argues in favour of an impaired semantic treatment of the question (Izaute et al., 2004).

Finally, we explored the possibility that the overestimation in confidence judgments of incorrect recalls observed in lorazepam participants was a kind of "ghost memory", similar to what happens to people who have lost an arm and still have sensations in the missing arm (i.e., the phantom limb syndrome; see Flor, 2002; Melzack, Coderre, Katz, & Vaccarino, 2001). If that were the case, confidence judgments would be based on information about the permanent, usual, and undrugged state of participants rather than their current drugged state. Thus, a high confidence judgment may be considered to concord with the usual permanent state of the participant (knowing the answer), but unadapted to his/her actual temporarily amnesic state. To explore this possibility, three participants from the Bacon et al. (1998) study who had received lorazepam were re-tested in their "natural" state 2 years later, together with one placebo participant. We compared their performance, as well as the accuracy of the judgments predicting their performance. From the drugged state to the normal state, the proportion of incorrect recalls dropped by more than 10% for the three participants who had first taken the lorazepam (respectively from 28 to 17%; 36 to 24% and 42 to 26%). The performance of the participant who had received a placebo for the first examination remained relatively stable across time (respectively 11 and 9% incorrect recalls). However, the mean confidence levels elicited by the participants who had previously been under the effect of lorazepam were in the range of 75–92 and varied by only six points (on a scale of 100) between the two test phases. The mean confidence levels of the placebo participant varied across a similar range (from 93 to 84) from the first to the second trial. The small sample presented here merely provides a few clues about the cognitive and metacognitive processes at work, but the fact that the memory performance of the three lorazepam participants was better when they were re-tested in an undrugged state, whereas that of the placebo participant was unchanged, seems to confirm that semantic memory is genuinely impaired by lorazepam. However, the confidence levels attributed to the recall answers by all the four participants remained relatively constant when tested either under lorazepam or under placebo.

What occurred with semantic knowledge under the effect of lorazepam may have been similar to the pain felt with the Phantom limb syndrome, "where the perception of pain does not simply involve a moment-to-moment analysis of afferent noxious input, but rather involves a dynamic process that is influenced by the effects of past experiences. Sensory stimuli act on neural systems that have been modified by past inputs, and the behavioral output is significantly influenced by the 'memory' of these prior events." (Melzack et al., 2001, p. 157). Participants under lorazepam could have been "influenced by the effects of past experiences", when they had easy access to the presently missing item. Consequently, they could have attributed to the transitory incorrect recall the same high confidence that they would usually have attributed to the correct answer that is momentarily not available because of the effect of the drug. Their behavioral output when rating their confidence would still rely on past inputs and their "memory of prior events", and this in turn would explain why their confidence levels under the drug or the placebo were the same. This lends support to the general idea of ghost memories; participants were basing their judgments on how their memory usually worked, not how it works under lorazepam. However, the present observation confirms the temporary nature of this retrograde impairment of semantic memory, as the ex-lorazepam participants performed better once the drug had been eliminated from their body, whereas the placebo participants' performance remained stable across time.

To summarize the additional analyses of previous experiments, the temporary impairment of semantic memory induced by lorazepam was confirmed in a general knowledge task, as was the preserved general access to knowledge about the topic of each question. Moreover, when under the effect of lorazepam, participants made commission errors that were semantically related to the target and more perseverative errors than the placebo participants.

4.3 The TOT Model as an Explanation for the Lorazepam-Induced Impairment of Semantic Memory

Of particular interest as regards the TOT model is that stating that an unrecalled target is on the tip of one's tongue implies at the very least that the target is known, and that recall is eminent to occur very soon.

In the next experiment, we investigated the possibility that participants under lorazepam could, for some items, be in a state of retrieval failure (i.e., temporary inaccessibility of a known item) and could retrieve a persistent alternate. However, they may not spontaneously experience phenomenological TOT, which would have told them that the correct target is a different word, on the verge of being retrieved. The persistent alternate would be given as the correct answer and attributed the same high degree of confidence they would usually attribute to the correct and otherwise known item. On being informed, however, that their response alternate is not correct, the lorazepam participants might then experience TOT states in respect of some of those items, just as normal participants do.

Thus, the aim of the present study was to see if participants under lorazepam experience more phenomenological TOT states after commission errors than control participants. Given that lorazepam participants made more commission errors (Bacon et al., 1998; Izaute et al., 2004; Massin-Krauss et al., 2002), we predicted that the general cognitive process of memory search is slowed down by the drug and, therefore, that the participants under lorazepam should manifest more commission TOT states than placebo participants (Brown, 1991). The subsequent resolution ability (i.e., recovery of the correct answer after a TOT experience) was also investigated. The literature shows that the recognition of TOT targets is usually good (Schwartz, 2002b). We conjectured that the phenomenological TOT states should also predict recognition here. It was hypothesized that retrieval performance in a recognition task of the commission TOT states should be equivalent to the performance of placebo participants because in previous studies lorazepam has not affected recognition performance, only recall performance (Bacon et al., 1998).

5 Experimental Ways to Capture a Particular TOT State

Diary studies have allowed for the collection of some information about the occurrence of a TOT state in a natural setting. Brown and McNeil (1966) were the first researchers to design an experimental paradigm for inducing TOT states in a controlled setting. The TOT states were precipitated by presenting students with definitions of low frequency English words and asking them to recall the words. Since then, several researchers have focused on this question. Schwartz et al. (2000) devised an experimental paradigm that seemed highly interesting for the study of lorazepam, especially as previous research tended to focus only on omission errors (Koriat, 1993). Using the paradigm devised by Schwartz et al. (2000) it is also possible to explore a TOT state occurring after the participant has provided an incorrect recall; this was called “commission TOT state”. In the procedure developed by Schwartz et al. (2000) participants were presented with general knowledge questions, and those who were unable to recall the target word were asked whether they were experiencing a TOT. The TOT states were assessed after both omission and commission errors. In addition, after a commision error, participants were informed that their response was incorrect and subsequently asked whether they were now experiencing a TOT. It was observed that in some cases, a phenomenological TOT could occur once a person was made aware that her/his first response was incorrect. Moreover, in the Schwartz et al. (2000) study the commission TOT state had the same general properties as the omission TOT state. In particular, following a comission TOT state, participants were more likely to retrieve the correct target later on than when they did not experience the phenomenological TOT. This is exactly what we expected to happen, that is, eventual retrieval of the correct target word would occur more frequently under the effet of the amnesic drug lorazepam. In the following study, we were keen to see whether this effect would be exaggerated in lorazepam participants.

6 The Experiment

The experimental procedure was based on Schwartz et al. (2000). For a complete description see Bacon et al. (2007).

6.1 Stimuli

The stimuli were 100 general knowledge questions. In the recognition task, participants were offered five possible answers, including the correct one. Except from the 100 questions, 20 unanswerable questions were also presented; most of them were taken from Schwartz et al. (2000; e.g., “For which country the monetary unit is the jaque?”). These 20 questions sounded plausible but had no correct answer (e.g., no country has a monetary unit called the jaque).

6.2 Participants and Experimental Design

Participants in the study were 30 healthy, French native-speaking students from Strasbourg University. They were pseudo-randomly assigned (on the basis of age, weight, and general knowledge) to one of two parallel groups, that is, a placebo group ($n = 15$) and a lorazepam 0.038 mg/kg group ($n = 15$), taking into account their general knowledge as evaluated by the Information and Vocabulary subtests of the Wechsler Adult Intelligence Scale–Revised (WAIS-R; Wechsler, 1987). The two groups were not significantly different in terms of age, $t(28) = 0.07$, *ns*, of weight, $t(28) = 0.05$, *ns*, or of pre-drug general knowledge as assessed using the Information subtest, $t(28) = 0.61$, *ns*, and Vocabulary subtest, $t(28) = 0.87$, *ns*. Informed written consent was obtained from all volunteers before they embarked on the study, which was approved by the Faculty Ethics Committee.

The drug capsule was given orally in a double-blind procedure. Each participant was tested individually in the presence of an experimenter. The questions were displayed on the computer screen one at a time. Participants were given an explanation of the term “tip-of-the-tongue”. It was drawn to their attention that they should not confuse a TOT experience with a very strong feeling of knowing. They were also informed that the TOT experience is relatively rare, and that they might not experience it at all in the course of this experiment. These instructions were given to avoid the risk of an artifactual TOT state – Widner et al. (1996) suggest that participants might sometimes express TOT states just to avoid appearing uneducated in front of the experimenter. Participants were asked to give the answer aloud or to say “I don’t know”. If they indicated they did not know the answer (omission errors) or provided an incorrect response (commission errors), they were asked if they were in a TOT state. In the case of answerable questions, the questions were displayed a second time, and each participant then made a feeling-of-knowing judgment, as a prediction of successful recognition of the correct answer from among a total of five answers. Finally, the participants completed a recognition test in case of answerable questions.

At the end of the study participants were required to rate their sedation state using a set of 16 visual analogue scales (Norris, 1971). Overall, sedation scores were higher for the lorazepam group ($M = 36.6$; $SD = 14.1$) than for the placebo group ($M = 23.8$; $SD = 13.3$), $t(28) = 2.5$, $p < 0.05$. Pearson correlations were also calculated between the sedation score and the memory and metamemory performance levels. No significant correlation was found in either group between self-ratings of sedation and recall performance and mean feeling-of-knowing results.

7 Results

7.1 Memory Performance

Memory performance scores (see Table 5.2) confirmed the previous observations (Bacon et al., 1998), since the mean proportion of total answers in the recall phase was not significantly different between the two groups, $t(28) = 0.2$, $p = 0.86$. Also, lorazepam participants’ ratio of commission errors was higher than that of placebo participants, $t(28) = 2.3$, $p < 0.05$. Lorazepam participants did not give significantly more answers ($M = 3.4$, $SD = 1.80$) than the placebo participants ($M = 2.5$, $SD = 1.86$) to the unanswerable questions, $t(28) = 1.2$ $p = 0.26$. The recognition performance of the two groups was not significantly different, $t(28)=1.2$, $p=0.23$. Thus, lorazepam impaired semantic memory performance only when participants had to recall the correct anwer.

7.1.1 Occurrence of TOT States

Of the 30 participants, seven (three from the lorazepam group and four from the placebo group) did not produce any of the two types of TOT. The lorazepam participants experienced more commission TOT states than placebo ones (see Table 5.3). However, the individual TOT percentages were similar in both groups, as there was no difference between the proportion of TOT states produced after an omission error, $t(27)=0.26$, $p=0.80$, or between the proportion of TOT states produced after a commission error in placebo groups and lorazepam, $t(25)=0.2$, $p=0.86$. Thus, lorazepam participants had more commission TOT states because they made more commission errors. The analysis of the commission TOT states showed that the nature of the errors was similar in both groups, with most of them being semantically related to the target word. The proportion of semantically

Table 5.2 Mean (and *SD*) proportions of answers in the free recall and recognition tests in the placebo and lorazepam groups (adapted from Bacon et al., 2007)

	Group	
	Placebo	Lorazepam
	M (*SD*)	*M* (*SD*)
Free recall test		
Total of answers	0.76 (0.08)	0.76. (0.10)
Correct answers	0.79 (0.09)	0.70 (0.12)*
Commission errors	0.21 (0.09)	0.30 (0.12)*
Recognition test		
Correct answers	0.84 (0.05)	0.81 (0.06)
Commission errors	0.16 (0.05)	0.09 (0.06)

*Significant difference at $p<0.05$

Table 5.3 Frequencies of omission and commission TOT states, of semantically related commission TOT states, and means (and *SD*) of individual proportions of the respective TOT states in the placebo and lorazepam groups

	Placebo group	Lorazepam group
Frequencies		
Total of TOT states	166	184
Omission TOT states	117	108
Commission TOT states	49	76
Semantically related commision TOT states	48	72
Means (and SD)		
Individual proportion of TOT states		
Omission TOT states	0.33 (0.18)	0.32 (0.15)
Commission TOT states	0.24 (0.19)	0.25 (0.17)

related commission TOT states to overall number of commission TOT states was 0.98 and 0.95 for the placebo and lorazepan groups, respectively with no significant difference between them, $t(21)=1.0$, $p=0.33$.

7.1.2 Resolution of TOT States

"Resolution" is the likelihood that a TOT state was followed by subsequent correct recognition of the target answer. For placebo participants, the resolution rate in the case of a TOT state ($M=0.83$, $SD=0.10$) was significantly better than after a non-TOT state ($M=0.60$, $SD=0.06$), $t(13)=8.3$, $p<0.01$. For lorazepam participants, resolution in the case of a TOT state ($M=0.73$, $SD=0.10$) was only marginally better than that of a non-TOT state ($M=0.59$, $SD=0.11$), $t(13)=2.1$, $p=0.054$. The difference between the TOT resolution rate of placebo and lorazepam participants was not significant, $t(27)=1.3$, $p=0.19$.

7.1.3 Metamemory Characteristics of the TOT and Non-TOT States

The mean feeling-of-knowing judgments did not differ between the placebo and lorazepam participants, $t(28)=0.26$, $p=0.80$, *ns*. The mean feeling-of-knowing judgments were significantly higher after a TOT state than after a non-TOT state, $F(1,27)=158.6$, $p<0.001$. There was no difference between the placebo and lorazepam group, $F(1,28)=0.004$, $p=0.95$. The predictive accuracy of TOT states on recognition, computed with the gamma correlations (Nelson, 1984), was preserved by the drug as the gamma correlations between TOT states and recognition were not significantly different for the placebo and lorazepam participants, $t(26)=1.2$, *ns*. Similarly, the predictive value of feeling-of-knowing judgments on recognition was not curtailed by lorazepam, $t(28)=0.2$, *ns*. However, the predictive value of TOT states on feeling-of-knowing judgments was significantly higher in the placebo group, $t(27)=2.3$, $p<0.05$. This means the lorazepam participants suffered from an impaired relationship between the two forms of knowledge monitoring (see Table 5.4).

Table 5.4 Means (and *SD*) of feeling-of-knowing judgments for answers after TOT and non-TOT states and gamma correlations in the placebo and lorazepam groups

	Placebo group	Lorazepam group
	M (*SD*)	*M* (*SD*)
FOK judgments in TOT states	75.6 (14.8)	68.1 (17.1)
FOK judgments in non-TOT states	38.6 (17.4)	46.1 (18.7)
Gamma correlations		
TOT and recognition	0.55	0.37
Feeling of knowing and recognition	0.36	0.35
TOT and feeling of knowing	0.86	0.67*

*Significant difference at $p<0.05$

8 Discussion

The present study was undertaken in order to investigate whether the higher number of incorrect recalls by participants under the effect of the amnesic drug lorazepam in a general knowledge task could be attributed partly to the fact that they experience more recall failures that can be identified as specific kinds of TOT, that is, commission TOT state (Schwartz et al., 2000). It was hypothesized that participants under the effect of the amnesic drug lorazepam would experience more often a cognitive TOT state (i.e., the failure of the process to retrieve a known word), which becomes a phenomenological TOT state (i.e., the strong feeling that a particular word is on the verge of being retrieved) only after they became aware of the retrieval failure. We also wanted to confirm that the cognitive and the phenomenological TOT states can be dissociated. Finally, given that a TOT state reveals a conflict between the cognitive and metacognitive levels, we suspected that the anxiolytic and anticonflict effects of lorazepam may be partly responsible for the mechanisms and occurrences of commission TOT.

Thanks to further analysis of findings by Bacon et al. (1998) and Massin-Krauss et al. (2002), we have confirmed some preliminary conditions for the workability of the model of TOT state as an explanation for the pattern of semantic memory/metamemory impairment induced by lorazepam.

The research undertaken to confirm the hypothesis produced conclusive results. As observed in the previous experiments, both lorazepam and placebo participants gave the same number of recall answers to a set of general knowledge questions. However, the lorazepam participants made more recall errors and experienced more TOT states following retrieval errors than placebo participants, whereas resolution of the TOT state (the ability to recognize the correct answer eventually) was unimpaired. The group of participants having received lorazepam reported 29 more cases of commission TOT states than the placebo participants. The eventual analysis of the memory and metamemory characteristics of these commission TOT states revealed that the commission TOT states experienced under lorazepam were similar in all respects to those experienced under placebo and in everyday life. Consequently, the difference induced by lorazepam in respect of commission TOT state is only quantitative, that is, commission TOT state is more frequent under the effects of the drug than under a placebo. So, it would seem that the impaired recall performance of participants under lorazepam could indeed be partly due to dissociation between the phenomenology and cognitive process of a TOT state, that is, the participant would be in the cognitive state of a TOT, which implies that she/he knows the correct target but that this target is temporarily inaccessible. In addition, a persistent alternate would come to mind, but unlike what occurs in the case of normal participants experiencing a TOT state, the cognitive aspect of the situation would not be accompanied by the feelings characteristic of a TOT. Because there is no phenomenological TOT occurring alongside the retrieval failure, the persistent alternates are produced as answers and end up being recorded as commission errors. This seemed to be confirmed insofar as lorazepam participants, overall, experienced more commission TOT states, and their recognition ability was preserved.

8.1 The Mixed Effects of Lorazepam on Semantic Memory and TOT State

The benzodiazepine lorazepam drug does not radically disturb semantic processes. Semantic memory, when evaluated with tests of verbal fluency, remains largely unaltered under the effect of benzodiazepines, suggesting that overall accessibility to the semantic store is largely unaffected by lorazepam. The questionnaire used in the present study differs in many respects from the sentence verification and fluency tasks, in particular because it requires participants to give their own individual answers to general knowledge questions. The profile of the lorazepam group's performance, characterised by a preserved number of answers in the recall task and low percentage of correct answers in the recall task, indicates that the drug does not impair performance by reducing accessibility to information. The present observations could explain the coexistence of semantic memory impairment, as observed in general knowledge tasks, with the preserved performance of lorazepam-treated participants in verbal fluency tasks. Some of the authors who have used fluency tasks also observed a slowing down of the reaction time (Brown, Brown, & Bowes, 1983; Green et al., 1996; Vermeeren et al., 1995). In verbal fluency tasks, the slower-than-normal retrieval process brought about by the drug seems not to curtail its efficiency. This was not the case for general knowledge questions which are more demanding and require the retrieval of a single correct answer.

On the other hand, the cognitive TOT state is usually regarded as a slowing down of the normal retrieval phenomenon (Brown, 1991). Under the effect of benzodiazepines a cognitive TOT state seems to occur more often than in healthy individuals, and the wide range of different recall errors is also an argument in favor of the possibility that lorazepam participants remain stuck in one of the preliminary stages of lexical search (Miozzo & Caramazza, 1997). As most of the errors were semantically related to the target, preserved accessibility to the semantic store would allow items belonging to a general category to be provided correctly, but not necessarily the single correct answer corresponding to the question pointer.

The monitoring failure in the case of commission TOT state is the inability to detect the temporary inaccessibility of the correct target answer. Koriat (1998) argued that "the key to illusion of knowing must lie not only in the inaccessibility of the correct target, but also in the inflated accessibility of contaminating clues that cannot be readily discredited" (p. 27). This suggests that the failure to spontaneously experience the phenomenological TOT in the case of commission TOT states could also inflate confidence with respect to the persistent alternate. However, after the commission error is revealed, the partial information may serve again to trigger a TOT experience. Schwartz and Smith (1997) observed that participants used the products of retrieval as a source of information for phenomenological TOT states. This would help to explain the emergence of TOT phenomenology after the participant has been told his/her response is incorrect. Some additional cues from the current experiments with lorazepam also lend support to this hypothesis, that is, greater accessibility to related information under the effect of lorazepam may be inferred from the wide range of different commission errors produced under lorazepam.

8.2 The Anxiolytic Effect of Lorazepam on Phenomenological TOT State

When people experience a commission TOT, they do not feel the phenomenological TOT, that is, the anxiety and the conflict. The benzodiazepine lorazepam is an anxiolytic drug with well known anticonflict effects that alleviates emotions (Harvey, 1980; Kleven & Koek, 1999; Vanover et al., 1999). It seems likely that the anxiolytic effect of lorazepam has eliminated the conflict resulting in a dissociation between the phenomenology and cognitive component of a TOT state. When under the effect of lorazepam, people are more likely not to be aware of the emotional conflict between the persistent alternate and missing correct answer. Consequently, they are also more likely to report the persistent alternate with greater frequency. However, when it is brought to their attention that they are wrong, the state of retrieval failure becomes identifiable, triggering the phenomenological TOT state. Thus, in a sense the lorazepam masks the emotional state created by the TOT conflict, allowing more commission errors to be made, but then subsequently producing more TOT states. According to what happened under lorazepam, consciousness does not necessarily mirror the process under way, and it is possible that with commission TOT states we have experienced an additional type of TOT resolution, namely emotional resolution. Indeed, in a commission TOT state there is no conflict because there is no emotion. However, the price to be paid for this absence of conflict is a memory failure, since participants give the persistent alternate as the genuine answer. So, in fact, this would be a counterproductive resolution of the TOT conflict in commission TOT states. However, feedback allows the participants, either under placebo or lorazepam, to experience the phenomenology and eventually to retrieve the correct answer. This means, therefore, that identifying the existence of a cognitive conflict seems to be necessary for the cognitive problem to be correctly solved. Yet, in daily life we do not necessarily have the good fortune of receiving feedback about what we say, and this may be problematic for people using benzodiazepine as drug to alleviate their anxiety, particularly as it seems there is no complete tolerance to the drug's amnesic effects.

9 Conclusion

In summary, as predicted, the lorazepam-treated participants experienced TOT after an incorrect recall more often than placebo participants. However, their ability to resolve TOT states (to find the correct answer eventually) was preserved in a subsequent recognition task, and the cognitive and metacognitive characteristics of the TOT state were also preserved by the drug. The impairment caused by lorazepam in a general knowledge task to assess semantic memory might therefore be partly the result of a greater sensitivity to a very common memory error, the cognitive TOT state, probably because lorazepam-treated participants spend longer than normal participants in a very preliminary state of memory search. Participants under lorazepam seem to experience dissociation between the phenomenology and cognitive process of the TOT states. This peculiar means of conflict resolution may be interpreted in light of the drug's anxiolytic effect.

Cognitive commission TOT state seems to be real and plausible entity corresponding to a particular cognitive and metacognitive state. In commission TOT states, resolution of the TOT conflict seems to be emotional, involving suppression of the phenomenological feelings, but at the cost of an incorrect answer. The anxiolytic and anticonflict effect of benzodiazepines seems to play a part in the more frequent occurrence of these specific memory blocks. Retrieval of the phenomenological TOT seems necessary to overcome the block created by persistent alternates in an appropriate manner. The use of lorazepam allowed us to further our understanding of the possible mechanisms of the TOT experience. Of interest as regards lorazepam is that we experimentally increased the number and diversity of persistent alternates retrieved while keeping correct recognition constant; this could provide a good tool for psychologists and linguists keen to study the effect of persistent alternates on the TOT process. To precipitate TOT states in healthy people, researchers have provided the participants with words that are potentially plausible persistent alternates (Smith, 1994). With lorazepam, we caused the drugged participants to come up with their own natural, "endogenous" persistent alternates that may be of interest for further investigating the TOT processes. Insofar as the retrograde drug-induced impairment of semantic memory is temporary and reversible following acute administration of lorazepam, further investigations into the exact nature of these semantic failures might be of interest for gaining a better understanding of the memory deficits that may occur in both healthy participants and patients suffering from organic amnesia.

We are conscious of the fact that many of our arguments are quite speculative and need further investigation. Moreover, it is possible that different amnesic drugs or different pathological conditions (Matison, Mayeux, Rosen, & Fahn, 1982) could lead the person to become stuck in a different step of the retrieval process, and further research is needed to explore that possibility. The effects of lorazepam on semantic memory may not be generalized to include the other molecules of the benzodiazepine family, because the literature shows that the patterns of cognitive impairments induced by benzodiazepines may vary greatly from one molecule to the next (Curran, 1999). In particular, it seems that lorazepam is very different in a number of respects, especially as regards the effects on cognition (Giersch et al., 2010). To the best of our knowledge, there is only one study of another molecule from the same family, triazolam, that has been conducted using a general knowledge task (Mintzer et al., 2010), and in that study the researchers observed no effect of triazolam on recall ability. So it would be interesting to explore the effects of the other molecules in the family.

In the meantime, it seems obvious that lorazepam induces other semantic deficits. In the study presented here, lorazepam also induced more perseverative errors, and the Moses illusion paradigm enabled us to reveal another kind of subtle, highly specific and reversible impairment that may often go unnoticed in everyday situations (Izaute et al., 2004). Studies into the long-term effects of benzodiazepines on cognitive functions suggest that tolerance to the memory impairments caused by benzodiazepines never fully develops (Barker, Greenwood Jackson, & Crowe, 2004; Stewart, 2005). Consequently, these specific semantic impairments may severely compromise the normal conduct of day-to-day activities for the vast numbers of chronic lorazepam users throughout the world.

Acknowledgements This work was supported by Inserm and by a *Cognitique* "Impairments and recovery of cognitive functions" grant from the French Ministry of Research. The author wishes to thank Marie Izaute, Bennett Schwartz, and Laurence Paire-Ficout for their valuable contribution to the TOT study.

References

Allen, D., Curran, H. V., & Lader, M. (1993). The effects of single doses of CL 284, 846, lorazepam, and placebo on psychomotor and memory function in normal male volunteers. *European Journal Clinical Pharmacology, 45*, 313–320.

Bacon, E., Danion, J. M., Kauffman-Muller, F., Schelstraete, M. A., Bruant, A., Sellal, F., & Grange, D. (1998). Confidence level and feeling of knowing for episodic and semantic memory: An investigation of lorazepam effects on metamemory. *Psychopharmacology, 138*, 318–325.

Bacon E., Schwartz B., Paire-FicoutM., & Izaute M. (2007). Dissociation between the cognitive process and the phenomenological experience of TOT: Effect of the anxiolytic drug lorazepam on TOT states. *Conscious Cognition, 16*, 360–373.

Bäckman L., Lipinska B. (1993). Monitoring of general knowledge: Evidence for preservation in early Alzheimer's disease (1993). *Neuropsychologia*, 31, 335–345.

Barker, M. J., Greenwood, K. M., Jackson, M., & Crowe, S. F. (2004). Cognitive effects of long-term benzodiazepine use: A meta-analysis. *CNS Drugs, 18*, 37–48.

Beracochea, D. (2006). Anterograde and retrograde effects of benzodiazepines on memory. *ScientificWorld Journal, 16*, 1460–1465.

Biedermann, B., Ruh, N., Nickels, L., & Coltheart, M. (2008). Information retrieval in tip of the tongue states: New data and methodological advances. *Journal of Psycholinguist Research, 37*, 171–198.

Brennen, T., Vikan, A., & Dybdahl, R. (2007). Are tip-of-the-tongue states universal? Evidence from the speakers of an unwritten language. *Memory, 15*, 167–176.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin,109*, 204–223.

Brown, J., Brown, M. W., & Bowes, J. B. (1983). Effects of lorazepam on rate of forgetting, on retrieval from semantic memory and dexterity. *Neuropsychologia, 21*, 501–512.

Brown, R., & McNeill D. (1966). The tip of the tongue experience. *Psychological Bulletin, 1090*, 339–347.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30*, 542–579.

Curran, H. V. (1991). Benzodiazepines, memory and mood: A review. *Psychopharmacology, 105*, 1–8.

Curran, H. V. (1999). Effects of anxiolytics on memory. *Human Psychopharmacology, 14,* 72–79.
Danion, J. M. (1994). Drugs as tools for investigating memory. *European Neuropsychopharmacology, 4*, 179–180.
Duka, T., Curran, H. V., Rusted, J. M., & Weingartner, H. J. (1996). Perspectives on cognitive psychopharmacology research. *Behavioral Pharmacology, 7*, 401–410.
Erickson, T. D., & Mattson, M. E. (1981). From word to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behaviour, 20*, 540–551.
File, S. E., Sharma, R., & Shaffer, J. (1992). Is lorazepam-induced amnesia specific to the type of memory or to the task used to assess it? *Journal of Psychopharmacology, 61*, 76–80.
Flor, H. (2002). Phantom-limb pain: Characteristics, causes, and treatment. *Lancet Neurology, 3*, 182–189.
Fluck, E., File, S. E., Springett, J., Kopelman, M. D., Rees, J., & Orgill, J. (1998). Does the sedation resulting from sleep deprivation and lorazepam cause similar cognitive deficits? *Pharmacology Biochemistry Behavior, 59*, 909–915.
Giersch, A., Boucart, M., Elliott, M., & Vidailhet, P. (2010). Atypical behavioural effects of lorazepam: Clues to the design of novel therapies? *Pharmacology and therapeutics, 126*(1), 94–108.
Green, J. F., McElholm, A., & King, D. J. (1996). A comparison of the sedative and amnestic effects of chlorpromazine and lorazepam. *Psychopharmacology, 128*, 67–73.
Harley, T. A., & Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology, 89*, 151–174.
Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Education and Psychology, 56*, 208–216.
Harvey, S. C. (1980). Hypnotics and sedatives. In A. Goodman Gilman, L. S. Goodman, & A. Gilman (Eds.), *The pharmacological basis of therapeutics* (pp. 339–375). New York: Macmillan.
Izaute, M., & Bacon, E. (2005). Specific effects of an amnesic drug: Effect of lorazepam on study time allocation and on judgment of learning. *Neuropsychopharmacology, 30*, 196–204.
Izaute, M., Paire-Ficout, L., & Bacon, E. (2004). Benzodiazepines and semantic memory: Effects of lorazepam on the Moses illusion. *Psychopharmacology, 172*, 309–315.
Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989a). Memory and metamemory: Comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology 17*, 3–11.
Janowsky, J. S., Shimamura, A. P, & Squire, L. R. (1989b). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia, 27*, 1043–1056.
Jones, G. V., & Langford, S. (1987). Phonological blocking in the tip-of-the tongue state. *Cognition, 26*, 115–22.
Kleven, M. S., & Koek, W. (1999). Effects of benzodiazepine agonists on punished responding in pigeons and their relationship with clinical doses in humans. *Psychopharmacology, 141*, 206–212.
Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*, 609–639.
Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further Evidence for the Accessibility Model. *Journal of Experimental Psychology: General, 124*, 311–333.
Koriat, A. (1998). Illusions of knowing: The link between knowledge and metaknowledge. In V. Y. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Metacognition: Cognitive and social dimensions* (pp. 16–34). London: SAGE.
Koriat, A., & Goldsmith, M. (1996). Memory in naturalistic and laboratory contexts: Distinguish the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123*, 297–315.
Maril, A., Simons, J. S., Weaver, J. J., & Schacter, D. L. (2005). Graded recall success: An event-related fMRI comparison of tip of the tongue and feeling of knowing. *NeuroImage, 24*, 1130–1138.
Massin-Krauss, M., Bacon, E., & Danion, J. M. (2002). Effects of the benzodiazepine lorazepam on monitoring and control processes in semantic memory. *Consciousness and Cognition, 11*, 123–137.

Matison, R., Mayeux, R., Rosen, J., & Fahn, S. (1982). "Tip-of-the-tongue" phenomenon in Parkinson disease. *Neurology, 32*, 567–570.

Melzack, R., Coderre, T. J., Katz, J., & Vaccarino, A. L. (2001). Central neuroplasticity and pathological pain. *Annual reports of New York Academy of Sciences, 933*, 157–174.

Mintzer, M., Kleykamp, B., & Griffiths, R. (2010). Dose effects of triazolam and scopolamine on metamemory. *Experimental and Clinical Psychopharmacology, 18*(1), 17–31

Miozzo, M., & Caramazza, A. (1997). Retrieval of lexical-syntactic features in tip-of-the-tongue states. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1410–1423.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95*, 109–133.

Nelson, T. O., McSpadden, M., Fromme, K., & Marlatt, G. A. (1986). Effects of alcohol intoxication on metamemory and on retrieval from long-term memory. *Journal of Experimental Psychology: General, 115*, 247–254.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

Norris, H. (1971). The action of sedatives on brain stem occulomotor system in man. *Neuropharmacology, 10*, 181–191.

Pappas, B. A., Sunderland T., Weingartner, H. M., Vitiello, B., Martinson, H., & Putnam, K. (1992). Alzheimer's disease and feeling-of-knowing for knowledge and episodic memory. *The Journal of Gerontology. Series B: Psychological Sciences, 47*, 159–164.

Reason, J., & Lucas, D. (1984). Using cognitive diaries to investigate naturally occuring memory blocks. In J. E. Harris & P. E. Morris (Eds.), *Everyday memory, actions, and absentmindedness* (pp. 53–70). London: Academic.

Reder, L. M., & Kusbit, G. W. (1991). Locus of the Moses illusion: Imperfect encoding, retrieval, or match? *Journal of Memory and Language, 30*, 385–406.

Schacter, D. L. (1999). The seeven sins of memory. Insight from psychology and cognitive neuroscience. *American Psychologist, 54*, 182–203.

Schwartz, B. L. (1998). Illusory tip-of-the-tongue states. *Memory, 6*, 623–642.

Schwartz, B. L. (1999). Sparkling at the end of the tongue: The etiology of tip-of-the-tongue phenomenology. *Psychonomic Bulletin Review, 6*, 379–93.

Schwartz, B. L. (2001). The relation of tip-of-the-tongue states and retrieval time. *Memory and Cognition, 229*, 117–126.

Schwartz, B. L. (2002a). The phenomenology of naturally-occurring tip-of-the-tongue states: A diary study. In S. P. Shohov (Ed.), *Advances in Psychology Research* (Vol. 8, pp. 71–84). New York: Nova Sciences.

Schwartz, B. L. (2002b). Tip-of-the-tongue states. *Phenomenology, mechanism, and lexical retrieval*. Mahwah, NJ: Erlbaum.

Schwartz, B. L. (2008). Working memory load differentially affects tip-of-the-tongue states and feeling-of-knowing judgments. *Memory Cognition, 36*, 9–19.

Schwartz, B. L., & Smith, S. M. (1997). The retrieval of related information influences tip-of-the-tongue states. *Journal of Memory and Language, 36*, 68–86.

Schwartz, B. L., Travis, D. M., Castro, A. M., & Smith, M. S. (2000). The phenomenology of real and illusory tip-of-the-tongue states. *Memory and Cognition, 28*, 18–27.

Shimamura, A. P., & Squire, L. R. (1986). Memory and metamemory: A study of the feeling-of-knowing phenomenon in amnesic patients. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 452–460.

Shimamura, A. P., & Squire, L. R. (1988). Long-term memory in amnesia: Cued recall, recognition memory, and confidence ratings. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 763–770.

Shimamura, A. P. (1994). The neuropsychology of metacognition. In: J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 253–276). Cambridge, MA: MIT Press.

Smith, S. M. (1994). Frustrated feelings of imminent recall: On the tip-of-the tongue. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 27–46). Cambridge, MA: MIT Press.

Stewart, S. A. (2005). The effects of benzodiazepines on cognition. *Journal Clinical Psychiatry, 66*, 9–13.

Taylor, J. K., & MacKay, D. G. (2003). Tip-of-the-tongue phenomena: Goldmine or can of worms. Review of Tip-of-the-tongue states: Phenomenology, mechanism, and lexical retrieval. *American Journal of Psychology, 116*, 291–298.

Vanover, K.E., Robledo, S., Huber, M., & Carter, R. B. (1999). Pharmacological evaluation of a modified conflict procedure: Punished drinking in non-water-deprived rats. *Psychopharmacology, 145*, 333–341.

Vermeeren, A., Jackson, J. L., Muntjewerff, N. D., Quint, P. J,. Harrison, E. M., & O'Hanlon, J. F. (1995). Comparison of acute alprazolam (0.25, 0.50 and 1.0 mg) effects versus those of lorazepam 2 mg and placebo on memory in healthy volunteers using laboratory and telephone tests. *Psychopharmacology, 118*, 1–9.

Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science, 8*, 314–317.

Wechsler, D. (1987). *Wechsler memory scale – Revised manual.* San Antonio, TX: Psychological Corporation.

Widner, R. L., Smith, S. M., & Graziano, W. G. (1996). The effects of demand characteristics on the reporting of tip-of-the-tongue states and feeling-of-knowing states. *American Journal of Psychology, 109*, 525–538.

Chapter 6
Prospective Memory Failure and the Metacognitive Experience of "Blank in the Mind"

Anastasia Efklides and Alexandra Touroutoglou

1 Introduction

Metamemory research has a long history in cognitive psychology. Specifically, since the 1960s, a great number of studies have investigated experiences related to retrospective memory failures, such as the tip-of-the-tongue (TOT) experience and the feeling of knowing judgements (Brown & McNeil, 1966; Koriat, 2000). However, whereas there is a lot of metamemory research on retrospective memory failures, there is less research on metacognitive experiences that denote failures in prospective memory (PM) (Sugimori & Kusumi, 2009). According to Meacham and Leiman (1982; see also Brandimonte, Einstein, & McDaniel, 1996), prospective memory is the process of remembering to carry out an action in the future, for example, taking one's medicine at a particular hour. Prospective memory comprises an intention (what to do) and a cue that signifies the condition for enacting the intention. The cue could either be a stimulus in the environment (event-based PM) or an internal stimulus such as the time elapsed (time-based PM) or the outcome of an activity (activity-based PM).

In the present chapter, we report a study on metacognitive experiences related to PM failures, and more specifically, on the "blank in the mind" (BIM) experience (Moraitou & Efklides, 2009). The BIM experience is defined as a sudden awareness of having no content in conscious awareness. In the case of PM, this entails that the person feels that s/he has lost track of the intention or the cue for initiating one's PM action. In an attempt to assess the effect of ongoing task demands on PM performance (cf. Einstein, Smith, McDaniel, & Shaw, 1997) and relevant metacognitive experiences, we examined BIM and other experiences related to PM failures (e.g., awareness of having committed an omission or commission error, or TOT) under conditions of different task demands.

A. Efklides (✉)
School of Psychology, Aristotle University of Thessaloniki,
541 24, Thessaloniki, Greece
e-mail: efklides@psy.auth.gr

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_6, © Springer Science+Business Media, LLC 2010

In what follows, we shall firstly refer to metacognitive experiences related to PM and PM failures, and we shall distinguish BIM from other PM-related metacognitive experiences. Then we shall briefly present theories that can explain PM failures and their potential for differentially predicting the emergence of the various PM-related metacognitive experiences. Then the study will be presented and the findings will be discussed in light of their implications for metamemory theory and research.

1.1 BIM and Other Metacognitive Experiences Related to Prospective Memory (PM)

According to Koriat, Ben-Zur, and Druch (1991) and Koriat, Ben-Zur, and Sheffer (1988) there are two types of monitoring processes involved in PM: output monitoring and input monitoring. Output monitoring refers to the monitoring of the outcome of one's action (e.g., that the action was successfully carried out) whereas input monitoring refers to the monitoring of the stimulus (cue) that should initiate the enactment of an intention. Accordingly, the two types of monitoring give rise to different metacognitive experiences. For example, a deficiency in output monitoring may give rise to uncertainty as to whether one has performed the purported action (Koriat et al., 1991), leading to repetition of the action, or uncertainty as to whether one has responded correctly or incorrectly. The former experience often accompanies omission errors (Sugimori & Kusumi, 2009).

Input monitoring may involve searching of the environment for the presence of the PM cue or probing one's memory for the availability of the PM cue. Remembering the PM cue is crucial for the execution of the PM action, because failure to retrieve the cue while there is memory of the intention nullifies the ability for the enactment of the intention. According to Einstein and McDaniel (2005), cue retrieval can be automatic, and hence, the recognition of the cue spontaneously pops up into mind. Cue retrieval can also be an effortful process in the sense that the person rehearses the cue in working memory and monitors the environment for its presence. The phenomenological characteristic of the experience of cue retrieval in the latter case is one of "search". According to Meier, Zimmermann, and Perrig (2006), automatic cue retrieval is typically accompanied by a "pop up" experience. In their study, Meier et al. (2006) asked participants to report their experiences at the end of each successful trial and found that "pop up" experiences increased after priming of the cue. Their study also showed that when context-specificity was induced by informing participants beforehand that a specific block of trials would include the PM cue, participants reported a "search" rather than a "pop up" experience. Apparently, context-specificity increased strategical input-monitoring and hence the search experience (Einstein & McDaniel, 2005).

Besides the "pop up" or "search" experience, input monitoring may lead to another metacognitive experience, namely, the BIM experience (Moraitou & Efklides, 2009). There are various occasions when people experience such a blank in the mind. Broadbent, Cooper, Fitzgerald, and Parkes (1982) considered BIM as a cognitive failure due to attention or memory lapses (see also Reason & Mycielska, 1982). BIM is present in younger as well as in older people, and differs from the experience of lack of memory information (knowledge) on a topic (Moraitou & Efklides, 2009). Depressed or anxious people as well as schizophrenics tend to report such BIM experiences (Chapman, 1966; Watts, MacLeod, & Morris, 1988) but BIM cannot be reduced to a pathological state. The person experiencing BIM has the strong feeling that the information was in memory but it cannot be retrieved at that moment despite one's efforts. In the case of PM, BIM can be experienced with respect to the intention (e.g., what did I want to do?) or the cue for the enactment of the intention (e.g., what is the cue I am looking for?).

A typical occasion in which BIM is experienced is when a course of action starts; for example, one starts preparing a meal and because some activity cannot be performed (e.g., because an ingredient is missing) a sub-goal is set, namely to go get the ingredient from a cupboard in another room. The activity initiated by the sub-goal is partly executed, that is, the person moves to the other room and there realizes that s/he does not remember what s/he wanted to do by going there. This is often followed by posing questions to oneself, such as "What was I about to do next?" or "Why did I come here?".

Thus, contrary to the experiences related to output monitoring deficiency (e.g., uncertainty), people having the BIM experience do not have doubts whether a correct response was made to a PM cue; instead, they are aware of their failure to retrieve the intention or the PM cue related to a goal or sub-goal. More specifically, BIM is conceptualized as the metacognitive experience that accompanies a failure to maintain in memory the intention that should initiate a secondary PM task in the context of another main task or the cue that should initiate the PM activity. BIM is likely not associated with the encoding of the cue or of the intention because the cue or intention was initially encoded and can, later on, be retrieved from memory. The specific characteristic of BIM is that consciousness temporarily seems to have no content (blank) because the initially encoded information was not maintained active in memory to guide the action. Yet, one feels that it is in memory and can be retrieved if one goes back to the initial conditions in which the intention or the cue was established. What is missing in conscious awareness is the memory of the cue or of the intention, which is necessary either for the monitoring of the environment for the presence of the cue or for executing the intended action. Another possibility is that part of the intended activity was automatically (or mindlessly) executed and the failed memory of the initial intention impedes output monitoring and initiation of the next part of the intended activity.

Finally, another experience that denotes retrospective memory failure but can be also present in the context of PM is the tip-of-the-tongue (TOT) experience. Specifically, the TOT experience regards recalling part of the stimulus but not the whole of it. This is often the case in situations with multiple target cues, in which cue interference may occur (Koriat, Pearlman-Avnion, & Ben-Zur, 1998). The TOT experience informs that some part of the target cue is accessible but not the whole cue and at the same time there is a feeling that cue retrieval is imminent (Metcalfe, 2000; see also Chap. 5, this volume). Hence, phenomenologically, TOT differs from

BIM both in terms of the content of conscious awareness (there is information in TOT but not in BIM) and the feeling of imminent retrieval (present in TOT but not in BIM). In fact, in BIM there is a clear feeling that the missing information cannot be retrieved at that moment despite one's efforts and therefore the secondary action has to be cancelled until the intention or the cue is reinstalled, for example, by going back to the original activity, so that the need experienced there reinstalls the missing intention or cue. Thus, BIM is associated with awareness of inaccessibility of the cue but with no interference of one cue over another, that is, no presence of persistent alternate that blocks the accessibility of the targeted information. Therefore, there should be no association between TOT and BIM.

1.2 Theories of PM Failures and Metacognitive Experiences

1.2.1 The Monitoring Deficiency Hypothesis

In general, PM failures are examined within four different theoretical frameworks. According to the monitoring deficiency hypothesis proposed by Koriat et al. (1991), PM failures are due to a deficiency in the process of input or output monitoring. Marsh, Hicks, Hancock, and Munsayac (2002) also emphasize output monitoring deficiency as the main factor responsible for repetitions and omissions in PM. In terms of metacognitive experiences, the monitoring deficiency hypothesis does not make any predictions about the exact subjective experiences that might occur when there is input or output monitoring failure except for uncertainty as to whether one successfully performed the intended action. BIM or TOT experiences, on the contrary, denote that input monitoring is operational and informs on memory failures.

1.2.2 Preparatory Attention and Memory Theory (PAM)

This theory of PM failures (Loft, Kearney, & Remington, 2008; Smith, 2003) posits that cue detection requires resources. Attentional resources are always needed for PM performance to be successful. This is evident in studies that demonstrate the task interference effect, according to which PM task demands slow down the main (ongoing) task performance even in no cue trials (Hicks, Marsh, & Cook, 2005). This implies that input monitoring per se is not the critical factor in PM failures. Rather it is the attention allocation policy at the early stage of cue encoding that is important, because it is formed by taking into account the PM task difficulty.

From a metamemory point of view the PAM theory implies that if the main (ongoing) task demands on attentional or working memory resources are increased, then there will be less resources available for the PM task; consequently, PM-related information (e.g., the intended action or cue) might not be sustained in memory, thus leading to BIM experience. On the other hand, if there is interference of information related to the ongoing PM task (alternate cues for the PM task), then TOT would be experienced. However, to the best of our knowledge, this assumption has not been tested in the past. Specifically, Marsh and Hicks (1998) showed that tasks making demands on executive control (or working memory) increased the PM failures. Einstein et al. (1997) studied the influence of increased ongoing task demands at the encoding and retrieval stage and found that adding at either stage a digit-monitoring task to the word-rating task (ongoing task) reduced significantly the PM performance. Furthermore, Einstein, McDaniel, Williford, Pagan, and Dismukes (2003) induced an additional task during a delay period in the PM task and showed that it increased the probability for making PM errors. However, Marsh, Cook, and Hicks (2005) showed that PM performance under demanding situations, such as divided attention conditions, is reduced only when the process needed to detect the cue is similar to the processing required to perform the ongoing task. This implies that resources are required for preventing possible interference between the cue detection processes such as input monitoring processes and the processes involved in the execution of the ongoing task. If interference does occur, then TOT should be experienced.

On the other hand, Sugimori and Kusumi (2009) showed that limited attentional resources influence not only PM performance but also PM judgements, such as old/new judgments and output monitoring. Participants in a limited attentional resources condition tended to guess more if the correct cue had been presented to them (input monitoring) and if they had performed the required action (output monitoring). And this was found even in the case in which participants had successfully performed the PM action. This implies that under limited-resources conditions omission or commission errors would occur because of input or output monitoring deficiency. Moreover, output monitoring would inform the person of the possible PM failure and the omission or commission errors.

1.2.3 The Multiprocess Theory of PM

While the PAM theory posits that capacity-demanding processes are always present in a PM task, McDaniel and Einstein (2000) argue that PM performance relies on multiple processes, including automatic ones. According to this view, people rely on both strategical monitoring and automatic retrieval processes of environmental cues, depending on the circumstances. Individual differences and specific task conditions, such as the distinctiveness of the target, the ongoing main task, and the PM task characteristics determine the extent to which individuals will allocate attentional resources to the process of cue detection or rely on automatic retrieval processes (Einstein et al., 2005). Einstein and McDaniel (2005) have provided evidence for automatic retrieval, confirming the multiprocess hypothesis. At the metamemory level, Meier et al.’s (2006) findings on search and “pop up” experiences provide support for the multiprocess theory. However, no exact prediction can be made regarding other kinds of metacognitive experiences in PM conditions, such as BIM or TOT.

1.2.4 Momentary Lapses of Intention

According to Craik and Kerr (1996), momentary lapses of intention (MLI) may be another condition for eliciting omission errors in PM. In MLIs, there is failure to maintain the cue-action schema in an active or easily accessible state, that is, the intention temporarily falls below awareness (Craik & Kerr, 1996). MLIs were observed in West and Craik's (1999) study that showed fluctuations of the PM performance over the course of an event-based PM task. This implies that MLIs could lead either to input monitoring failure, leading to omission errors without the person being aware of such errors, or to awareness of BIM as a failure to maintain the cue-action schema active. However, the MLI theory does not make any specific prediction as to whether BIM would occur. In any case, this theory does not predict that increased task demands are a necessary condition for the occurrence of MLIs and consequently of BIM.

To sum up, PM research has not extensively investigated the various metacognitive experiences that may occur during a PM task. The analyses of the effects of increased PM task demands on working memory resources suggested that both BIM and TOT could occur under such conditions but TOT should be related to awareness of interference of the processing of the ongoing task with the processing of the PM task, whereas BIM should not. Moreover, BIM could occur in cases of MLIs, which implies that it should be present even in tasks with lower task demands. In any case, there should be no correlation between BIM and TOT.

1.3 The Present Study

The aim of the present study was to investigate the awareness people have of BIM and other metacognitive experiences that are related to PM failures such as TOT and awareness of omission and commission errors. Also, to test the hypothesis that increased PM task demands have a negative effect on performance on the PM task and an increase of the self-reported frequency of metacognitive experiences related to PM failures. Specifically, the assumption was that the higher the demands of the ongoing task on working memory (WM), the higher the probability that there will not be sufficient resources for the maintenance of the PM intention or the cue that should trigger the related action. This would increase the PM failures, or the time for carrying out the PM action, as well the awareness of BIM. Moreover, when the main task involves processes that may interfere with cue detection, then TOT should be experienced. However, if BIM is related to MLIs, then task demands should not make a difference on the awareness of BIM.

The above assumptions were tested with a computerized PM task. The design of the PM task, however, did not conform to the standard PM paradigm (Einstein & McDaniel, 2005) in which both the cue (e.g., a sound) and the response that should follow it are defined at the beginning of the task. Then another task (ongoing or main task) is taking place and at certain points the cue appears and the person has to execute the predefined response to it. In this paradigm, the intention and the cue are established once at the beginning of the task and do not change during the execution of the main task. Therefore, if one commits an omission error, this does not entail that a BIM or TOT experience would occur because omission could be due to cue monitoring failure for any reason. To be able to create conditions for the elicitation of BIM, and in order to distinguish it from the TOT experience, we designed a task that could give rise to both experiences. Thus, the PM cue changed from one block of trials to the next but the same type of cue (e.g., number) was used in all blocks. This could create either BIM experience, if the cue was not maintained in memory, or TOT experience if part but not the exact cue was recalled, due probably to interference of the memory of the previous cue(s) to the current one.

The context of the PM task was a reading comprehension task that preceded the main and the PM task and offered the database from which the participants would retrieve information when asked during the PM task. The main task required carrying out arithmetic operations whereas the PM task required pressing the Spacebar key when a predefined cue was present; pressing the Spacebar then led to questions on the reading comprehension task that had to be answered. PM task demands on working memory were manipulated by including an *n*-back task that was embedded in the main task.

1.3.1 Hypotheses

The following hypotheses were formulated: (a) *PM performance.* Increasing the demands of the main task on WM should increase the PM failures, that is, not pressing the spacebar key when the cue is present (omission error) or pressing the spacebar when the cue is not present (commission error) as well as the time needed for performing the PM task (Hypothesis 1). (b) *Metacognitive experiences.* Increasing the demands on WM should increase the awareness of PM failures as denoted in ratings of BIM, TOT, and other PM-related metacognitive experiences such as awareness of omission or commission errors. However, if BIM is related to MLIs then no increase of BIM ratings should be found in the presence of increased WM demands (Hypothesis 2a). Moreover, the ratings of BIM and other metacognitive experiences regarding the various PM failures should correlate with performance and reaction time on the PM task (Hypothesis 2b). (c) *Relations between metacognitive experiences*. Since the assumed mechanisms underlying the formation of the various PM-related metacognitive experiences differ, the correlations between BIM and TOT, and between them and the awareness of omission and commission errors should be low (Hypothesis 3).

2 Method

2.1 Design

The design of the study was 2(type of PM task: event-based vs. activity-based) ×2(WM load: no *n*-back vs. *n*-back task) between subjects factorial. Thus, the four groups of the study were as follows: (a) event-based PM task (PM-event group);

(b) activity-based PM task (PM-activity group); (c) event-based PM task with an embedded *n*-back task in the main task (PM-event *n*-back group), and (d) activity-based PM task with an embedded *n*-back task in the main task (PM-activity *n*-back group). The reading comprehension and the main task (arithmetic operations) were the same in all groups. What differed was the cue for executing the intention. Also, the *n*-back task was the same in the two groups that received it. After the completion of the PM task, participants responded to a series of self-report measures tapping metamemory awareness of PM performance and PM failures. Before starting the experiments, participants were administered two working memory tasks and the Stroop task that tested inhibitory control. The four groups did not differ with respect to working memory capacity of central executive and episodic buffer, $F(3, 105)=3.641$, $p>0.05$ and $F(3, 105)=2.190$, $p>0.05$, respectively, nor in inhibitory control (Stroop-task performance), $F(3, 105)<1$, *ns*. The WM and the inhibitory control measures were included as control because differences in resources might have an effect on the experimental treatments.

2.2 Participants

There were 110 psychology students in the four groups. They participated voluntarily in the study. The PM-event group comprised 32 students (6 males and 26 females; $M=21.63$ years, $SD=4.26$); the PM-activity group comprised 27 students (9 male and 18 females; $M=20.37$ years, $SD=1.52$); the PM-event *n*-back group comprised 23 students (4 males and 19 females; $M=19.74$ years, $SD=2.82$); and the PM-activity *n*-back group comprised 28 students (6 males and 22 females; $M=18.50$ years, $SD=0.88$). There were no significant differences as regards age and gender representation in the four groups.

2.3 Apparatus

An Intel PC with standard keyboard and a 17-in SVGA monitor was used for task presentation and response registration. Programming of tasks was completed with E-prime (Schneider, Eschman, & Zuccolotto, 2002) which allowed recording of response accuracy (0 or 1) and logging of actual response as well as of time (in milliseconds).

2.4 Tasks

2.4.1 Reading Comprehension Task

Participants were instructed to read a text because later on, they would have to answer to a set of multiple-choice questions on it. The text was 659 words long on a cognitive psychology topic that was unfamiliar to the participants. Questions were presented one at a time during the PM task, upon participants' pressing the spacebar when the cue appeared. No feedback on the correctness of response was provided.

2.4.2 Main Task

The main task included a series of arithmetic operations (e.g., 56×4, $47 + 39$) that were presented to participants one at a time. Participants were instructed to carry out the arithmetic operations and write the outcome on the blank space shown on the screen. One of the outcome numbers represented the cue for carrying out the secondary task (PM task), that is, pressing the SPACE key, in order to respond to questions related to the reading comprehension task. The key to proceed to the next arithmetic operation was ENTER.

For all groups there were six blocks of arithmetic operations, and six PM occasions, one at each block. The cue was given at the beginning of a block of arithmetic operations items. The blocks varied in terms of the number of items included. Thus, there were blocks ranging from 1 to 7 items, respectively. The position of the cue varied from block to block. There was no feedback on the correctness of the arithmetic operations or on pressing the Space or Enter key.

2.4.3 PM Task

PM-Event Group

In this group, participants pressed SPACE as soon as they had come across a predefined number, that is, the cue (e.g., 15). This number was the outcome of one of the arithmetic operations. This number changed from block to block.

PM-Activity Group

In this group, participants pressed SPACE after solving a predefined number of arithmetic operations (specified each time by the experimenter at the beginning of a block of arithmetic items; for example, "Press SPACE after the fifth item"). Thus, they had to rely on their own monitoring of the number of arithmetic operations they had carried out. This explicit monitoring of the number of items executed (based on internal counting) was expected to increase demands on WM compared to the event-based PM task, in which there could be automatic retrieval of the cue.

PM-Event *n*-Back and PM-Activity *n*-Back Groups

In both of these groups, participants had to carry out the main task and at the same time perform an embedded working memory *n*-back task. The embedded *n*-back task required participants to remember the last digit of the first number of the two involved in each arithmetic operation (e.g., digit 5 of the number 25 in the operation: 25 + 3 = 28). After a series of arithmetic operations (not known to the participants) and without prior notice, participants were asked to repeat the last digit of the two operations they had carried out immediately before, in the same order in which the respective items had been executed.

Scoring

There were the following scores: (a) correct pressing of SPACE (i.e., correct response when the PM cue was present); (b) omission errors (not pressing the SPACE key when it should be pressed); (c) commission errors (pressing SPACE instead of the correct ENTER), and (d) Reaction time needed for the motor response, that is, for writing the outcome of the arithmetic operations and pressing the SPACE or the ENTER key, respectively.

2.4.4 Metacognitive Experiences

Upon completion of the six blocks of items, the following questions appeared on the screen in order of their numbering. Responses were given on a five-point Likert-type scale, measuring the perceived frequency of having done or having experienced what was stated in the item. The scale ranged from 1 (never) to 5 (all the time).

(a) Correct pressing of SPACE

Q1 (Output monitoring): "How many times did you correctly press the SPACE key?"

(b) PM failures

Q2 (Interference of the previous cue with the current cue): "How many times did a previous cue interfere with the current one?"
Q3 (Failure to retrieve the current cue although one remembered the previous cue; possible BIM experience): "How many times did you remember a previous cue but failed to remember the current one?"
Q6 (Failure to retrieve the predefined cue; possible BIM experience): "How many times did you fail to remember the specific cue for pressing the SPACE key?"
Q13 (Failure to retrieve which key should be pressed; possible BIM experience): "How many times did you fail to remember what you had to press: The SPACE or the ENTER key?"
Q8 (BIM experience): "How many times did you experience a blank-in-the-mind state?"
Q10 (Remembering only part of the number that served as cue but not the whole number; possible TOT experience): "How many times could you recall one or more digits of the cue-number but failed to recall the whole number?"

Q11 (TOT experience): "How many times did you experience a tip-of-the-tongue state?"
Q4 (Awareness of omission errors): "How many times did you fail to press the SPACE key?"
Q7 (Failure to press the SPACE key; omission error/output monitoring): "How many times did you recognize the cue but failed to press the SPACE key?"
Q5 (Awareness of commission errors): "How many times did you fail to press the ENTER key and pressed SPACE instead?"
Q9 (Uncertainty as to whether the cue had been presented; failure of input monitoring): "How many times did you feel uncertain as to whether you had come across the cue?"
Q12 (Uncertainty about the execution of the intended action; failure of output monitoring): "How many times did you feel uncertain as to whether you had pressed the SPACE or the ENTER key?"

2.5 Procedure

Participants were tested in two phases, each lasting approximately 20 min. During the first phase, participants were tested with a non-computerized Working Memory test battery. In the second phase, participants were first asked to complete a computerized form of the Stroop task (word-reading and color-naming task) loaded in the E-prime and afterwards completed the PM task. Upon completion of the task, they were asked to fill in the ME questionnaire.

3 Results

The data were submitted to a series of 2(type of PM task: event-based vs. activity-based) ×2(WM load: no *n*-back vs. *n*-back task) ANOVAs. The descriptives of performance and RT are given in Table 6.1. The significant *F* effects and interactions are shown in Tables 6.2 and 6.3.

3.1 PM Task Performance

3.1.1 Accuracy of Response and Errors

WM load (*n*-back vs no *n*-back) on the main task rather than type of task (event-based vs. activity-based) affected the accuracy of response to the PM task, that is, correctly pressing the SPACE key. The same was true for the omission errors but

Table 6.1 Means and standard deviations of PM performance and reaction time (RT) in ms as a function of type of task and WM load

	Group							
	PM-event		PM-activity		PM-event *n*-back		PM-activity *n*-back	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Correct response to PM task	4.56	1.31	4.48	1.55	4.30	1.42	3.61	1.44
Omissions	1.44	1.31	1.52	1.55	1.70	1.43	2.39	1.44
Commissions	0.65	1.31	0.78	1.18	0.13	0.55	0.54	0.74
RT of correct PM response	14,272.92	8,219.45	20,155.66	11,065.48	24,000.39	13,775.02	15,535.11	8,490.36
RT of omission errors	4,575.48	4,216.28	6,600.96	7,075.52	9,804.69	7,640.24	12,662.39	9,644.82
RT of commission errors	6,009.61	14,098.83	7,318.55	12,264.92	1,031.17	2,556.63	3,689.00	5,901.75

Table 6.2 Significant *F* values in the ANOVAs applied on the PM performance measures

	F	*df*	Sig.	Partial η^2
Correct response to the PM task				
WM load effect	3.99	1, 101	0.048	0.03
Omission errors				
WM load effect	3.99	1, 101	0.048	0.03
RT of correct PM response				
Type of task × WM load effect	12.26	1, 101	0.001	0.01
RT of omission errors				
WM load effect	15.11	1, 101	0.000	0.13
RT of commission errors				
WM load effect	4.78	1, 101	0.031	0.04

Table 6.3 Means and standard deviations of metacognitive experiences as a function of type of task and WM load

	Group							
	PM-event		PM-activity		PM-event *n*-back		PM-activity *n*-back	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Q1	3.47	1.27	3.11	1.18	3.52	1.12	2.50	1.10
Q2	1.66	0.70	2.19	0.88	2.22	1.04	2.86	0.89
Q3	1.91	0.73	2.26	0.90	2.17	0.78	2.57	1.25
Q4	1.88	0.87	2.19	0.92	1.91	0.95	3.03	1.10
Q5	1.28	0.52	1.48	0.85	1.26	0.62	1.96	1.10
Q6	2.13	0.94	2.15	1.09	2.04	0.64	3.07	0.89
Q7	1.88	1.16	2.04	1.12	1.43	0.66	2.64	1.31
Q8	2.34	0.83	2.48	0.97	2.83	1.03	2.53	1.03
Q9	1.78	0.66	2.07	0.83	2.35	1.11	3.21	0.99
Q10	1.44	0.67	1.59	0.79	2.39	1.34	2.57	1.17
Q11	2.22	1.18	2.07	0.87	2.22	1.167	2.78	1.10
Q12	1.38	0.79	1.96	1.19	1.61	0.72	2.32	1.27
Q13	1.66	0.79	2.04	1.05	1.61	0.89	2.71	1.15

Note: Q1 = correct pressing of SPACE (output monitoring); Q2 = interference of the previous cue with the current cue; Q3 = failure to retrieve the current cue although one remembered the previous cue (possible BIM experience); Q4 = awareness of omission errors; Q5 = awareness of commission errors; Q6 = failure to retrieve the predefined cue (possible BIM experience); Q7 = failure to press the SPACE key (omission error/output monitoring); Q8 = BIM experience; Q9 = uncertainty as to whether the cue had been presented (failure of output monitoring); Q10 = remembering only part of the number that served as cue but not the whole number (possible TOT experience); Q11 = TOT experience; Q12 = uncertainty about the execution of the intended action (failure of output monitoring); Q13 = failure to retrieve which key should be pressed (possible BIM experience)

not the commission errors. The *n*-back groups (i.e., high WM load) had lower accuracy than did the low WM load (no *n*-back task) groups. These findings partly support Hypothesis 1 because WM load did not have an effect on commission errors in the PM task. Moreover, the results suggest that event-based and activity-based PM tasks did not differ as to their effect on accuracy of response or PM failures.

3.1.2 Reaction Time of Correct PM Response

The main effects of type of task and WM load were not significant in the case of RT to correct PM response. However, a significant interaction of type of task with WM load was observed, although the effect size was small. The interaction effect was analyzed using simple effects analysis. Results showed that the interaction was due to the PM-event group having longer reaction time in high WM load than in the low WM load condition. On the contrary, the reaction time in the activity-based groups was not affected by the WM load manipulation.

3.1.3 Reaction Time of Omission Errors

In agreement with Hypothesis 1, a significant main effect of WM load on reaction time of omission errors was found. Reaction time was longer in the high WM load groups than in the low WM load groups. No main effect of type of task or interaction between type of task and WM load were found. This suggests that the *n*-back task increased demands on WM during the main task and left fewer resources for the monitoring of the PM cue, thus, leading to omission errors. However, the increased RT in the omission errors suggests that there was an effort for cue monitoring albeit not always successful, possibly due to interference of the main task to the PM task.

3.1.4 Reaction Time of Commission Errors

A significant main effect of WM load was also found in the reaction time of commission errors. Contrary to Hypothesis 1, and to the findings regarding the reaction time of omission errors, reaction time of commission errors was longer in the low WM load groups than in the high WM load groups. This suggests that the source of commission errors is different from the source of omission errors. It is likely that in the low WM conditions, participants rehearsed the PM response required (i.e., SPACE) and this created a set for the SPACE even in the absence of the cue. This rehearsing, on the other hand, slowed down reaction time. However, when there was increased WM load no such rehearsing of response was possible and this led to shorter RT for commission errors. Finally, no main effect of type of task or interaction of type of task and WM load was observed.

To sum up the findings regarding PM performance, WM load rather than type of task increased the PM errors and the respective reaction time for the correct response. WM load increased omission errors and the related reaction times, but did not affect the number of commission errors; moreover, the reaction time for commission errors was shorter in the high WM load conditions than the low WM load ones.

3.2 Metacognitive Experiences

Means (and standard deviations) of metacognitive experiences related to PM correct response and PM failures as a function of type of task and WM load are given in Table 6.3. The significant *F*s are given in Table 6.4.

Table 6.4 Significant *F* values in the ANOVAs applied on metacognitive experiences

	F	*df*	Sig.	Partial η^2
Q1				
Type of task effect	9.28	1, 106	0.003	0.08
Q2				
Type of task effect	12.17	1, 106	0.001	0.10
WM load effect	13.55	1, 106	0.000	0.13
Q3				
Type of task effect	4.30	1, 106	0.040	0.03
Q4				
Type of task effect	15.00	1, 106	0.000	0.12
WM load effect	5.77	1, 106	0.018	0.05
Type of task × WM load effect	4.82	1, 106	0.030	0.04
Q5				
Type of task effect	8.55	1, 106	0.004	0.07
Q6				
Type of task effect	8.85	1, 106	0.004	0.07
WM load effect	5.68	1, 106	0.019	0.05
Type of task × WM load effect	8.09	1, 106	0.005	0.07
Q7				
Type of task effect	10.33	1, 106	0.002	0.09
Q9				
Type of task effect	11.33	1, 106	0.001	0.09
WM load effect	24.56	1, 106	0.000	0.19
Q10				
WM load effect	25.15	1, 106	0.000	0.19
Q12				
Type of task effect	10.85	1, 106	0.001	0.09
Q13				
Type of task effect	15.66	1, 106	0.000	0.13

Note: Q1 = correct pressing of SPACE (output monitoring); Q2 = interference of the previous cue with the current cue; Q3 = failure to retrieve the current cue although one remembered the previous cue (possible BIM experience); Q4 = awareness of omission errors; Q5 = awareness of commission errors; Q6 = failure to retrieve the predefined cue (possible BIM experience); Q7 = failure to press the SPACE key (omission error/output monitoring); Q8 = BIM experience; Q9 = uncertainty as to whether the cue had been presented (failure of output monitoring); Q10 = remembering only part of the number that served as cue but not the whole number (possible TOT experience); Q11 = TOT experience; Q12 = uncertainty about the execution of the intended action (failure of output monitoring); Q13 = failure to retrieve which key should be pressed (possible BIM experience)

3.2.1 Awareness of Correct Response

The main effect of type of task was significant on the ratings of *correct pressing of SPACE key (Q1)*, with the PM-event groups giving higher ratings of correct response than the PM-activity groups. It is worth noting that this effect is contrary to the accuracy of performance data (see above) that showed no effect of the type of task. It seems that explicit monitoring of the number of items processed in the PM-activity group prevented monitoring of the PM response and consequently lowered ratings of correct response.

3.2.2 Awareness of PM Failures

The main effect of both the type of task and WM load were significant on the self-reported *interference of the previous cue with the current cue* (Q2). The ratings on interference of the previous cue with the current cue were higher in the activity-based tasks than in the event-based tasks. Also the ratings were higher in the high WM load groups than in low WM load groups. Therefore, both explicit monitoring of the number of items processed and WM load created conditions for perceived cue interference.

Awareness of *failure to retrieve the predefined cue* (Q6; possible BIM experience) was affected by type of task and WM load. The interaction between type of task and WM load was also significant, showing that the WM load increased the reported failure to retrieve the predefined cue only in the activity-based tasks. More specifically, the PM-activity *n*-back group gave higher ratings than any other group. Therefore, the explicit monitoring of the number of items processed in the activity-based tasks, along with increased WM load increased perceived memory loss of the predefined cue. Whether this cue retrieval failure was taking the form of BIM is not clear.

The type of task also affected the self-reported *failure to retrieve the current but not the previous cue* (Q3; possible BIM experience) with the PM-activity groups having higher ratings than the PM-event groups. The main effect of WM load and the interaction between type of task and WM load were not significant.

A main effect of the type of task was also found on the reported *failure to retrieve the specific key to be pressed* (Q13; possible BIM experience). PM-activity groups had higher ratings than the PM-event groups. The main effect of WM load and the interaction of type of task and WM load were not significant.

It is worth noting that no main effect (or interaction) of type of task and WM load was found on the item explicitly asking for the presence of the *BIM experience* (Q8). Therefore, although self-reports of explicit BIM were not influenced by type of task or WM load, awareness of failures that could be related to BIM were influenced mainly by the type of task.

With respect to input monitoring failure (Q9), both type of task and WM load was significant. The self-reported *uncertainty as to whether the person had come across the cue* was higher in the PM-activity groups than in the PM-event groups and in the high WM load groups than in the low WM load groups. The interaction of type of task and WM load was not significant.

With respect to TOT experience, a main effect of WM load was found on self-reports of *remembering only a part of the cue* (Q10; possible TOT experience). Results showed that high WM load groups had higher ratings than did low WM load groups. The type of task and the interaction of type of task and WM load were not significant.

However, neither the main effect nor the interaction of the type of task and WM load were significant on the item explicitly asking for *TOT experience* (Q11), as also shown in the question on BIM.

The main effects of type of task and WM load were significant for the self-reported *omission errors* (Q4) as well as for the self-reported *failure to press the SPACE key* (Q7). The interaction of type of task and WM load was significant for both experiences. The results showed that WM load increased the ratings on omission errors and the ratings on failure to press the SPACE key in the PM-activity groups only.

A main effect of type of task was found on the ratings of *commission errors* (Q5) and *uncertainty about the execution of the intended action* (Q12; failure of output monitoring). The ratings of commission errors and of uncertainty were higher in the PM-activity groups compared to the PM-event groups. The main effect of WM load and the interaction between type of task and WM load were not significant.

In sum, the PM-activity groups gave higher ratings on metacognitive experiences related to PM failures than did the PM-event groups. Moreover, possible BIM experiences were found to be higher in the PM-activity groups than in the WM load groups, as predicted in Hypothesis 2a. Yet, the ratings on the explicit question on BIM experience were found to be similar across the four groups. In case of possible TOT, the effect of WM load was critical although no such effect was found for the question on explicit TOT (Q11). Furthermore, the type of task affected the ratings on metacognitive experiences denoting failure of input and output monitoring as well as commission errors, whereas both type of task and WM load affected the ratings of omission errors.

3.3 *Intercorrelations Between the Various Measures*

The intercorrelations between the ratings of the metacognitive experiences, overall, were low and nonsignificant contrary to Hypothesis 2b. Because of the large number of correlations computed, only the largest correlations will be presented ($p<0.01$). Moreover, the intercorrelations were computed within each group in order to identify possible interaction effects of the type of task and WM load factors.

3.3.1 Awareness of Correct Response and Actual PM Failures

The ratings on Correct Pressing of the SPACE key (Q1) correlated negatively with the number of the omission errors actually made ($r=-0.71$, $p<0.01$) but not with the number of commission errors actually made in the PM task. Yet, this correlation was found in the PM-activity *n*-back group, only.

3.3.2 Awareness of PM Failures and Actual PM Failures

Awareness of Omission Errors (Q4) and Commission Errors (Q5)

A positive correlation was found between the ratings of omission errors and the omission errors actually made in the PM-activity *n*-back group ($r = 0.71$, $p < 0.01$). Moreover, in the PM-event group, the ratings of omission errors correlated positively with the RT of omission errors ($r = 0.73$, $p < 0.01$). These findings suggest that participants were aware of their omission errors. However, this was not the case for commission errors – the correlation between ratings of commission errors and commission errors actually made in the PM task was not significant.

Failure to Retrieve the Specific Key to be Pressed (Q13)

In the PM-event *n*-back group, a positive correlation was observed between the ratings on failure to retrieve the specific key to be pressed and the commission errors actually made ($r = 0.78$, $p < 0.01$) as well as between the reported failure to retrieve the specific key to be pressed and the RT of commission errors ($r = 0.71$, $p < 0.01$).

Explicit BIM (Q8) and TOT (Q11)

No significant relationship was observed between ratings of explicit BIM and TOT on the one hand, and PM failures (actual omission or commission errors on the PM task), on the other. Also, there was no significant correlation between explicit BIM and TOT ratings and RT of actual omission or commission errors.

Finally, there was no correlation between the explicit BIM and TOT ratings, as well as between the ratings on the various metacognitive experiences, as predicted in Hypothesis 3.

4 Discussion

The aim of our study was to investigate the effect of PM task demands on metacognitive experiences such as BIM, TOT, awareness of interference of cues, awareness of omission and commission errors, as well as input and output monitoring failures. The results showed the following:

4.1 Performance

Accuracy of response, as well as omission errors in the PM task, was found to be affected by WM load. This finding is in agreement with that of Einstein et al. (2003) who found increased PM errors when an additional task was introduced during the delay period. This is also compatible with the PAM theory (Loft et al., 2008; Smith, 2003) according to which PM task performance is based on the attentional resources available. An example of this WM load effect is the situation in which a number of activities have to be carried out in a limited time frame and we forget to do some of them. This assumption is further supported by the RT data. High WM load increased the RT in the PM-event-*n*-back group in the case of accurate responses and RT in both *n*-back groups in the case of omission errors. These findings are in line with research on the effects of increased task demands on PM performance (Einstein et al., 2003; Marsh et al., 2005).

However, the WM effect did not apply to commission errors, which suggests that different processes might be involved in omission and commission errors. Contrary to RT in accurate PM response or omission errors, RT in commission errors was longer in the low than the high WM load condition, regardless of type of task. This finding cannot be explained by PAM theory. It seems that at the low WM conditions participants shifted attention from the cue to the response, thus committing commission error, but this focusing on response likely increased RT because there was an output monitoring process in order. This interpretation is supported by the finding that the number of commission errors made and RT of commission errors correlated with the ratings on the metacognitive questions denoting failure to retrieve the specific key to be pressed (Q13). However, this correlation was only found in the PM-event *n*-back group. This may denote that failure of output monitoring was more manifest in the case of high WM load and not in low WM load. Taken together, these findings suggest that commission and omission errors may be explained by different factors. Further research in prospective memory should focus on the mechanisms underlying the two types of PM failures.

4.2 Metacognitive Experiences

When it comes to the effects of type of task and WM demands on the responses to the metamemory questions, the most striking finding is that there was a systematic effect of type of task, although no such effect was found at the performance level (both response and RT). The WM effects were only found in the case of ratings on cue interference, on failure to retrieve the predefined cue (possibly due to interference), on failure of input monitoring, possible TOT, and on awareness of omission errors. In all these cases, except for the possible TOT (Q10), the WM effect coexisted with the type of task effect. This joint effect of WM and type of task suggests that the participants experienced the activity-based PM task as more demanding, particularly in the *n*-back condition. This was also evidenced at the performance level, although this trend did not reach statistical significance. Indeed, inspection of Table 6.1 suggests that the PM-activity-*n*-back group had the lowest correct response score and the highest score on omission errors than the other three groups. Therefore, the metacognitive data seem to reflect a tendency already present in performance scores. In any case, for ratings of TOT, interference, and omission errors WM load is a critical factor for both performance and metacognitive awareness.

On the contrary, the possible BIM experiences, awareness of commission errors and output monitoring failure were associated with the type of task (particularly the PM-activity tasks) and not with WM load. This suggests no cue interference in the case of BIM, but rather a failure to maintain in WM the cue or the key to be pressed when attention was devoted to internal counting of the number of arithmetic operations performed. Failure to maintain the intention (i.e., key to be pressed) seems to be involved in the commission errors and output monitoring failure as well. These findings lend support to the assumption that different processes are involved in omission and commission errors (as also suggested by performance data). They also imply that participants do differentiate between BIM and TOT experiences. The correlational findings also suggest that people have the ability to distinguish different types of PM-relevant experiences. In agreement with Hypothesis 3, no correlations were found between the various PM-related experiences, suggesting different mechanisms for output and input monitoring experiences (Koriat et al., 1998).

Of particular interest to this study was the awareness of BIM or TOT states. Looking at the mean ratings to the questions explicitly asking about these two metacognitive experiences (see Table 6.4: Q8 and Q11) it is clear that the participants reported that they had such states a few (or some) times (Ms > 2). However, these ratings were not influenced by either the type of task or the WM load. It is likely that other factors influenced the responses to these two items. One such factor could be that participants did not understand the questions the way they were conceptualized in the present study. In favor of this interpretation is the finding that these two questions did not correlate with others tapping the same or similar states (i.e., possible BIM or possible TOT). Another plausible explanation is that BIM is related to MLIs and the latter are independent of task demands or characteristics. It could be that MLIs are associated with worry or anxiety independently of the tasks themselves. TOT, on the other hand, is probably produced by interference of the main with the PM task, and again this interference was present in all conditions. More research is needed to determine which explanation is valid.

Despite the lack of clear evidence on the presence and the mechanism underlying the occurrence of BIM, drawing on the evidence of possible BIM and TOT states, it can be concluded that BIM is related to non-maintenance of information critical for input monitoring whereas TOT to task interference, particularly when there are high WM demands. TOT in essence jeopardizes input monitoring, and this may lead to omission errors.

4.3 Limitations of the Study

The present study had two limitations: first, the lack of a clear definition of BIM and TOT to the participants, when they responded in the metamemroy questions and, second, the post hoc character of the metamemory responses. In the latter case, memory processes might interfere and compromise the accuracy of the ratings given. This may explain the lack of substantial correlations between the metamemory responses and performance data. Yet, the evidence is that there was a relationship between omission errors and respective ratings whereas no such relationship existed in the case of commission errors and the respective self-reports. Ratings, however, on the question regarding failure to retrieve the specific key to be pressed correlated with commission errors, denoting that participants were aware of a possible reason that led to commission errors.

Overall, this study offers important data with regard to metacognitive experiences in PM and opens the way for the study of BIM. We found that participants had various PM-related metacognitive experiences, including experiences possibly related to BIM. We also demonstrated that increased PM task demands on WM increased PM failures and awareness of these failures. These results bear resemblance to the study of Sugimori and Kusumi (2009), which showed that limited attentional resources may affect both PM performance and PM judgements related to output monitoring. However, our data tap input monitoring and different kinds of metacognitive experiences. This kind of data is important for a deeper understanding of the mechanisms underlying PM performance and its related subjective experiences.

References

Brandimonte, M., Einstein, G. O., & McDaniel, M. A. (Eds.) (1996). *Prospective memory: Theory and applications*. Mahwah, NJ: Erlbaum.

Broadbent, D. E., Cooper, P. F., Fitzgerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21*(1), 1–16.

Brown, R., & McNeill, D. (1966). The "tip-of-the-tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*, 325–337.

Chapman, J. (1966). The early symptoms of schizophrenia. *British Journal of Psychiatry, 112*, 225–251.

Craik, F. I. M., & Kerr, S. A. (1996). Commentary: Prospective memory, aging, and lapses of intention. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 227–237). Mahwah, NJ: Erlbaum.

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science, 14*(6), 286–290.

Einstein, G. O., McDaniel, M. A., Williford, C. L., Pagan, J. L., & Dismukes, R. K. (2003). Forgetting of intentions in demanding situations is rapid. *Journal of Experimental Psychology: Applied, 9*(3), 147–162.

Einstein, G. O., Smith, R. E., McDaniel, M. A., & Shaw, P. (1997). Aging and prospective memory: The influence of increased task demands at encoding and retrieval. *Psychology and Aging, 12*(3), 479–488.

Einstein, G. O., Thomas, R., Mayfield, S., Shank, H., McDaniel, M. A., & Morrisette, N. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology: General, 134*(3), 327–342.

Hicks, J. L., Marsh, R. L., & Cook, G. I. (2005). Task interference in time-based, event-based, and dual intention prospective memory conditions. *Journal of Memory and Language, 53*(3), 430–444.

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition, 9*(2), 149–171.

Koriat, A., Ben-Zur, H., & Druch, A. (1991). The contextualization of input and output events in memory. *Psychological Research, 53*(3), 260–270.

Koriat, A., Ben-Zur, H., & Sheffer, D. (1988). Telling the same story twice: Output monitoring and age. *Journal of Memory and Language, 27*(1), 23–39.

Koriat, A., Pearlman-Avnion, S., & Ben-Zur, H. (1998). The subjective organization of input and output events in memory. *Psychological Research, 61*(4), 295–307.

Loft, S., Kearney, R., & Remington, R. (2008). Is task interference in event-based prospective memory dependent on cue presentation? *Memory and Cognition, 36*(1), 139–148.

Marsh, R. L., Cook, G. I., & Hicks, J. L. (2005). On the relationship between effort toward an ongoing task and cue detection in event-based prospective memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 31*(1), 68–75.

Marsh, R. L., & Hicks, J. L. (1998). Event-based prospective memory and executive control of working memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 24*(2), 336–349.

Marsh, R. L., Hicks, J. L., Hancock, T. W., & Munsayac, K. (2002). Investigating the output monitoring component of event-based prospective memory performance. *Memory and Cognition, 30*(2), 302–311.

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14*, S127–S144.

Meacham, J. A., & Leiman, B. (1982). Remembering to perform future actions. In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 327–336). San Francisco: Freeman.

Meier, B., Zimmermann, T. D., & Perrig, W. J. (2006). Retrieval experience in prospective memory: Strategic monitoring and spontaneous retrieval. *Memory, 14*(7), 872–889.

Metcalfe, J. (2000). Feelings and judgments of knowing: Is there a special noetic state? *Consciousness and Cognition, 9*(2), 178–186.

Moraitou, D., & Efklides, A. (2009). The Blank in the Mind questionnaire (BIMQ). *European Journal of Psychological Assessment, 25*(2), 115–122.

Reason, J., & Mycielska, K. (1982). *Absent-minded?* Englewood Cliffs, NJ: Prentice-Hall.

Schneider, W., Eschman, A., & & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh, PA: Psychology Software Tools Inc.

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning Memory and Cognition, 29*(3), 347–361.

Sugimori, E., & Kusumi, T. (2009). Limiting attentional resources influences performance and output monitoring of an event-based prospective memory task. *European Journal of Cognitive Psychology, 21*(1), 112–128.

Watts, F. N., MacLeod, A. K., & Morris, L. (1988). Associations between phenomenal and objective aspects of concentration problems in depressed patients. *The British Journal of Psychology, 79*, 241–250.

West, R., & Craik, F. I. M. (1999). Age-related decline in prospective memory: The roles of cue accessibility and cue sensitivity. *Psychology and Aging, 14*(2), 264–272.

Chapter 7
Metamemory in Schizophrenia: Monitoring or Control Deficit?

Marie Izaute and Elisabeth Bacon

1 Introduction

Schizophrenia is a common disease with a lifetime risk of about 1%. It is a chronic mental disease that appears at a young age (15–25 years) and can be found in all sectors of society. This mental disease is present in all regions of the globe and affects both men and women. Schizophrenia is characterized by a dissociation of thought and behavior and a fragmentation of consciousness. It has an extremely rich symptomatological pattern which may be observed in the form of behavioral problems, language and thought disorders, perceptual difficulties, affectivity-related problems and cognitive deficits (Bruder, Wexler, Sage, Gil, & Gorman, 2004). Overall, these symptoms can be subdivided into three categories: (a) positive or psychotic symptoms (hallucinations, deliria, etc.); (b) negative symptoms or deficiencies (limited volition, flattened affectivity, social withdrawal, etc.); and (c) cognitive deficits.

Schizophrenia has been linked to a wide range of cognitive deficits (Fioravanti, Carlone, Vitale, Cinti, & Clare, 2005; Green, 1996). These deficits constitute core symptoms of the pathology and should be included in the diagnostic criteria for schizophrenia (Lewis, 2004). Indeed, there is evidence suggesting that virtually all cognitive functions are impaired in patients with schizophrenia (Aleman, Hijman, de Haan, & Kahn, 1999; Heinrichs & Zakzanis, 1998), including executive functions (Bryson, Whelahan, & Bell, 2001), as well as attention and memory (Dickinson, Iannone, Wilk, & Gold, 2004; Moritz, Woodward, & Rodriguez-Raecke, 2006). In particular, with respect to executive functions, poor performance has been observed on tests such as the Wisconsin Card Sorting Test (Morice, 1990) and other traditional executive functions tasks such as the Verbal Fluency Task (Liddle & Morris, 1991).

M. Izaute (✉)
Social and Cognitive Psychology Laboratory (LAPSCO UMR CNRS 6024), Clermont University, Blaise Pascal University, 34 avenue Carnot, 63037 Clermont-Ferrand, France
e-mail: marie.izaute@univ-bpclermont.fr

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_7, © Springer Science+Business Media, LLC 2010

However, it must be emphasized that not all cognitive functions are impaired to the same degree and several studies have shown that memory functions are disproportionately impaired (Aleman et al., 1999; Driesen et al., 2008; Kraus & Keefe, 2007; Ranganath, Minzenberg, & Ragland, 2008). For instance, memory deficits observed in patients with schizophrenia primarily relate to explicit (declarative) memory, whereas implicit (non-declarative) memory seems to be relatively preserved (Danion, Meulemans, Kauffmann-Muller, & Vermaat, 2001b; Sponheim, Steele, & McGuire, 2004). It has been shown that patients' memory deficits are not caused by their treatment, the chronic nature of the disease, or by attendance at an institution (Huron & Danion, 2002). Memory deficits are typically evident in tasks assessing episodic memory, that is, memory for personal events. This has been consistently demonstrated in free-recall tasks (Koh & Peterson, 1978), cued-recall tasks (Schwartz, Rosse, & Deutsch, 1993) and, to a lesser degree, recognition tasks (Aleman et al., 1999; Calev, 1984a, b). There is also consistent evidence that the episodic memory deficit is related to an impairment of both encoding and retrieval processes.

Moreover, studies have shown that memory deficits in schizophrenia seem to be predictive of patients' income, satisfaction with daily activities, difficulties in everyday life (Green, 1996; Mohamed et al., 2008) and general health (Fujii, Wylie, & Nathan, 2004). Also, a number of neuroimaging studies support the view that high-level memory functions (proverb comprehension and inferring non-literal intentions behind interlocutors' use of proverbs) are impaired in schizophrenia (Thoma et al., 2009). In a review of neuroimaging studies conducted while schizophrenia patients were performing executive tasks, Ranganath et al. (2008) found that strategic memory encoding is linked to executive functions and that both are impaired in schizophrenia.

Moreover, schizophrenia is often defined as pathology of consciousness and associated with lack of self-awareness (Danion, Rizzo, & Bruant, 1999; Osatuke, Ciesla, Kasckow, Zisook, & Mohamed, 2008). In addition to cognitive deficits and lack of self-awareness, patients also manifest deficits in awareness of their memory capacity. Thus, as suggested by Bacon, Danion, Kauffmann-Muller, and Bruant (2001), the concept of metamemory, which refers to people's awareness of their own memory capacity and control of related behavior (Flavell, 1979), is of interest for a better understanding of the cognitive behavior of schizophrenia patients.

The present chapter on metamemory functioning in schizophrenia is organized as follows. First, consciousness and metacognition in schizophrenia is discussed with particular emphasis on metamemory. Both monitoring and control in metamemory are examined. Monitoring refers to an individual's subjective assessment of his or her own memory capacity and knowledge; within the broad category of memory monitoring two types of metacognitive experiences (Efklides, 2008; Flavell, 1979) will be discussed, one related to retrieval (i.e., feeling of knowing; Koriat, 1993) and the other to learning (i.e., judgment of learning; Nelson & Dunlosky, 1991). Control refers to the regulation of one's memory or behavior. Second, empirical findings from two studies will be presented suggesting that schizophrenia patients have preserved monitoring but impaired control of memory.

Finally, in the last part of the review, the interconnections or contradictions between empirical findings in various studies will be discussed, as well as attempts to reconcile the entire body of data.

2 Consciousness, Metacognition, and Schizophrenia

Schizophrenia patients' cognitive deficits are often accompanied by a disruption of their consciousness. Huron and Danion (2002) have described schizophrenia as pathology of states of consciousness. This notion is central to our understanding of this disease which is characterized, in particular, by a loss of the feeling of internal unity; some patients, for example, have the feeling that someone else is directing their behavior or thoughts. For Frith (1989), the symptoms of schizophrenia (hallucinations and delusions) can be interpreted as the result of a defect in the mechanism that controls and limits the contents of consciousness. The positive symptoms observed in schizophrenia seem to be related to a dysfunctioning of the awareness of one's own actions. Dapratti et al. (1997) have suggested that one possible explanation for the dysfunctioning of conscious awareness might lie in an inability to attribute meaning to external events. More precisely, Frith (1992) claims that schizophrenia patients fail to attribute elements arising from their long-term memory to themselves and are unable to differentiate their thoughts from "intentions" which result from external stimuli.

Metacognition has been seen as a key factor contributing to functional outcomes in schizophrenia (Koren et al., 2006). Metacognition involves, besides other facets (Efklides, 2008), our knowledge of how we perceive, remember, think, and act – in other words, what we know about what we know (Metcalfe & Shimamura, 1994; Nelson, 1996). Metacognitive knowledge is explicit knowledge about our cognitive strengths and weaknesses. There is evidence that some aspects of metacognitive knowledge are impaired in patients with schizophrenia (Bacon & Huet, 2005; Lysaker et al., 2005) whereas others are preserved. Some researchers further suggest that positive symptoms are likely to be associated with metacognitive beliefs (Laroi & Van der Linden, 2005; Morrison, Haddock, & Tarrier, 1995).

Specifically, according to Laroi and Van der Linden (2005, p. 1426), "metacognitive beliefs are beliefs that are linked to the interpretation, selection and execution of particular thought processes. These may include beliefs about thought processes (e.g., 'I do not trust my memory'), the advantages and disadvantages of various types of thinking (e.g., 'I need to worry in order to work well', 'I could make myself sick with worrying') and beliefs about the content of thoughts (e.g., 'It is bad to think certain thoughts')". Laroi and Van der Linden (2005) assessed metacognitive beliefs and hallucinations/delusions in a sample of 296 nonclinical participants. They used different scales to assess hallucination and delusion proneness and employed the Metacognition Questionnaire (MCQ; Cartwright-Hatton & Wells, 1997) to assess metacognitive beliefs. Their results showed that metacognitive beliefs were the best predictors of proneness to hallucinations. This finding is consistent with other research showing the same association in clinical samples of schizophrenia patients (Baker & Morrison, 1998; Lobban, Haddock, Kinderman, & Wells, 2002).

Impaired self-referential source memory also represents a specific cognitive deficit in schizophrenia. Specifically, schizophrenia patients exhibit a difficulty in recognizing that they themselves are responsible for the words produced during an earlier sentence completion task. The ability to remember that the self is the source of responses to a sentence completion task is strongly related to basic social cognitive processes such as face recognition and emotion identification. In the respective studies, the participants were asked to produce responses in alternation with the experimenter (Moritz & Woodward, 2006; Moritz, Woodward, & Ruff, 2003; Moritz, Woodward, Withman, & Cuttler, 2005). During a subsequent recognition test, the participants had to judge whether the presented items were study items or new items and then make a confidence judgment. Schizophrenia patients displayed increased confidence for their memory errors and a reduced level of confidence for their correct answers. This pattern of results is also detectable in first-episode schizophrenia (Moritz et al., 2006). A number of experiments, including those conducted by Huron et al. (1995) and Danion et al. (1999), based on the source recognition paradigm, have shown that schizophrenia patients make significantly more errors than controls when asked to state whether it was themselves or the experimenter who performed a given task.

Danion et al. (1999) have also shown that during a memory recognition task, the conscious recollection process which characterizes autonoetic consciousness (defined as the capacity to mentally relive a passed event; Tulving, 1985) is disrupted in schizophrenia. In contrast, another form of consciousness which is based on a feeling of familiarity (without conscious recollection) seems to be preserved in such patients. This other state of consciousness is referred to as noetic consciousness. Danion et al. (1999) hypothesized that it is autonoetic consciousness that permits the process referred to as "cognitive binding" (i.e., the ability to establish links between the different elements of an event) which is necessary for the construction of a coherent global representation of the world in which we live and which appears to be deficient in patients with schizophrenia (Burglen et al., 2004; Waters, Maybery, Badcock, & Michie, 2004).

2.1 Metamemory and Schizophrenia

The study of metamemory permits an experimental approach to the disorders of consciousness related to memory processes. Although metamemory research has flourished since the time when John Flavell first introduced the term "metamemory" (Flavell, 1971), most of the studies have been devoted to the understanding of normal cognitive and metacognitive functioning. More specifically, metamemory refers to the subjective awareness of one's memory capacity and control of the related cognitive behavior.

Flavell (1979) distinguished between metamemory knowledge and metamemory awareness, that is, metacognitive experiences. Metamemory knowledge comprises beliefs about memory irrespective of any particular task context; it refers to everyday life (e.g., "I have a good memory for numbers"). Metamemory awareness refers to the ability to monitor and control how relevant information is processed as a function of the goals and requirements of the task at hand. It can be thought of as a regulatory system which influences both memory encoding and retrieval. Previous studies conducted in normal participants have confirmed that monitoring and control influence memory performance (Koriat & Goldsmith, 1996; Nelson, 1993; Nelson, Dunlosky, Graf, & Narens, 1994; Nelson & Narens, 1990).

Monitoring of memory refers to the person's awareness of subjective experiences related to his or her memory. Monitoring of memory is expressed as metamemory judgments such as judgment of learning (JOL) at the end of the acquisition phase, and feeling of knowing (FOK) or confidence level (CL) at the phase of retrieval. When people fail to retrieve a target answer, they can at least say whether they have the feeling of knowing (or not knowing) the missing response, that is, that they can recognize it when they come across it. Monitoring also relates to the feelings and judgments the individuals have while performing a task (Efklides, 2006), for example, "This poem is difficult to learn". These feelings and judgments about our memory capacity allow the strategic regulation of learning behaviors. The control of memory refers to strategy use, allocation of study time, decision to produce an answer or abstain, continue memory search or spend more time searching for the known information, etc. For example, we spend more time learning something that we considered to be difficult to master rather than something we believed to be easy (Son & Metcalfe, 2000).

In studies investigating metamemory, the experiments are typically based on the calculation of the correspondence between the accuracy of an answer and its metamemory rating (Nelson & Narens, 1990). The rationale and methods for using these indexes were reviewed by Nelson (1984). It has been suggested that a non-parametric measure of association, the gamma correlation, provides a good summary index of FOK performance. The gamma index is a measure of association developed by Goodman and Kruskal (1954) which allows researchers to compare the correct predictions to the incorrect predictions. This index ranges from −1 to +1, with large positive values corresponding to a strong association between memory performance and metamemory judgments, whereas negative values show an inverse relationship.

A number of arguments proposed in the literature suggest that the memory deficit in schizophrenia may be linked to metamemory processes. First, certain studies indicate that schizophrenia patients present a deficit in the self-initiation of strategy use (Bonner-Jackson, Haut, Csernansky, & Barch, 2005; Christensen, Girard, Benjamin, & Vidailhet, 2006). Second, Kircher, Koch, and Stottmeister (2007) observed that although schizophrenia patients exhibited the same memory performance as control participants, they failed to evaluate their memory correctly. The authors concluded that these patients suffer from a metamemory deficit which is independent of memory performance.

2.1.1 Monitoring in Schizophrenia

It has repeatedly been observed that the retrospective confidence judgments for incorrect answers expressed by schizophrenia patients are higher than those of matched healthy participants (Danion, Gokalsing, Robert, Massin-Krauss, & Bacon, 2001a; Moritz et al., 2003; Moritz et al., 2005; Moritz & Woodward, 2006). Some studies have also shown that schizophrenia patients make lower prospective metamemory judgments than healthy participants (Bacon et al., 2001; Souchay, Bacon, & Danion, 2006). Bacon et al. (2001) attempted to investigate FOK and CL judgments in patients with schizophrenia using a general knowledge task to assess semantic memory. They found that whereas CL and FOK accuracy (the gamma correlation) did not significantly differ between the schizophrenia and control groups, FOK ratings were significantly lower in the group of schizophrenia patients. Correct answers associated with very low FOK ratings were observed more frequently in the schizophrenic than in the control participants. In another study assessing episodic memory and FOK, participants had to learn sentences which were eventually presented without the last word in the recall phase. The results indicated that for healthy control participants the FOK accuracy (the gamma correlation) was reliably nonzero, whereas, for the schizophrenia patients, this index was not statistically different from zero indicating a very weak association between recognition performance and FOK judgment for the patients (Souchay et al., 2006).

2.1.2 Control in Schizophrenia

Danion et al. (2001a) presented their participants (schizophrenia and healthy participants) with a general knowledge task in which they had to give answers under forced-report instructions and then evaluate their level of confidence concerning the correctness of the response. They were then asked to answer the same question under free-report instructions with or without a monetary incentive. This procedure made it possible to measure the accuracy of monitoring (the extent to which confidence judgments appropriately assess the correctness of answers) and the sensitivity of control (the extent to which providing or withholding responses is sensitive to confidence judgments). The results revealed impaired control sensitivity in schizophrenia patients whose confidence judgments were seen to be less compatible with the decision criteria they adopted in order to produce an answer.

3 Empirical Evidence on Metamemory in Schizophrenia

In what follows, two studies will be presented that assessed schizophrenia patients prospective judgments related to episodic memory. The first study (for a complete description, see Bacon & Izaute, 2009) is on FOK, a metamemory judgment that is expressed at the time of retrieval; this study examines the processes underlying FOK ratings. The second study (Bacon, Izaute, & Danion, 2007) examines JOL, which is expressed at the time of learning; by studying the strategic regulation of learning in schizophrenia patients, it will explore the relationship between monitoring and control.

3.1 The Basis for FOK Ratings in Schizophrenia Patients

The study examined schizophrenia patients' FOK in episodic memory tasks within the framework of the accessibility model (for a complete description see Bacon & Izaute, 2009). When people fail to retrieve the information they are looking for, they may nevertheless still be able to retrieve certain information related to the desired but currently inaccessible target. According to the accessibility model proposed by Koriat (1993, 1995, 1997), FOK may be related to the products of the retrieval process itself. When we search our memories for a desired target, a variety of incomplete details come to mind, including fragments of the target, semantic attributes, and episodic information pertaining to the target. The main assumption concerning the basis of FOK judgments embodied in this model is that "the cues for FOK reside in the products of the retrieval process itself" (Koriat, 1993). The accessibility hypothesis developed by Koriat suggests that a large amount of information is activated early in the search process, that is, before the target has been fully retrieved. The model states that the tendency to produce a high or low FOK depends on the overall amount of partial information elicited by the question.

3.1.1 Method

In the study there were 21 chronic clinically stable schizophrenia outpatients and 21 healthy participants (control group). Schizophrenia diagnoses were based on DSM-IV. The healthy participants were matched with the patients for age, gender, and educational level.

The task consisted of 40 tetragrams, that is, four-letter consonant strings (e.g., QVMJ) and was adapted from Koriat (1993). Koriat (1993) had used an ingenious experimental paradigm that made it possible to control the partial information corresponding to a given target answer and to establish a relation between the accessibility of the partial information and the metamemory rating, on the one hand, and the predictive accuracy of the FOK on the other. The total target answer to be learned consisted of nonsense letter strings of four consonants, each letter being partial information for the whole tetragram. With such a paradigm, Koriat (1993) could observe that participants base their FOK on the amount of partial information related to the target that they are able to retrieve. Each target string appeared for 3 s on the screen of a computer with the letters arranged vertically. The participants were asked to read the letters aloud and to learn them. After a short retention interval, they were asked to recall the letters of the tetragram, the order of recall being unimportant. They then estimated the probability of recognizing the correct tetragram (i.e., report their FOK) among eight displayed options by clicking on a 100-mm visual analog scale presented on the screen. Ratings of FOK were defined as ranging from 0% (will definitely not recognize) on the left to 100% (will definitely recognize) on the right. The recognition test was then presented, and the same procedure was repeated for the other tetragrams.

3.1.2 Results and Discussion

Metamemory Ratings

The recall of partial information was calculated on an individual basis. For each case of partial information recalled (four letters to zero letter), the mean FOK ratings were calculated. The FOK ratings were only significantly lower in schizophrenia patients than healthy participants in the cases where one letter was recalled (see Fig. 7.1).

An examination of the FOK ratings for each group considered separately showed that these increased linearly with the amount of partial information retrieved for both schizophrenia patients and healthy participants. In other words, FOK ratings were significantly higher when four letters were retrieved than when three were retrieved, and also when three letters were retrieved rather than just two. Finally, the ratings were significantly higher when two letters were retrieved rather than one. Only in the case of the schizophrenia patients was there no difference in the FOK ratings depending on whether one or no letters were retrieved.

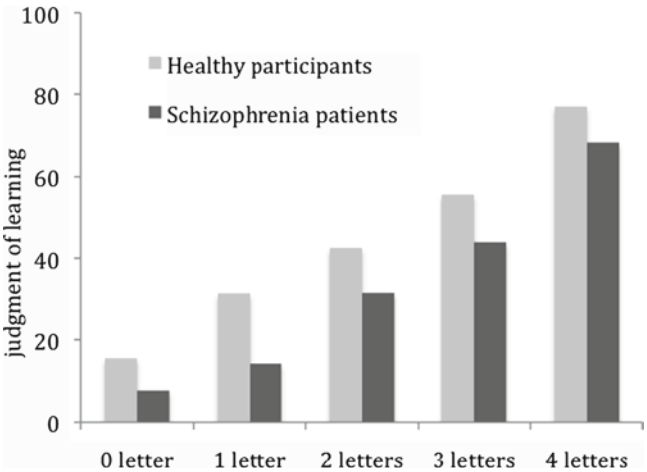


Fig. 7.1 FOK ratings for each group separately as a function of the amount of partial information recalled

Predictive Accuracy of the FOK Ratings and of Partial Recall (Relative Resolution)

Goodmann-Kruskal's gamma coefficient (Nelson, 1984) helps cast some light on the participants' ability to discriminate between correct and incorrect answers after making item-by-item judgments and provides a way of comparing correct and incorrect predictions. This gamma correlation was high for all the participants, thus suggesting that high FOK ratings were followed by correct recognitions and low FOKs by incorrect recognitions. We also calculated the gamma correlations between the amount of partial information retrieved for each tetragram and the FOK. The high correlation observed means that when a participant recalled a large amount of partial information, she/he also reported a strong feeling of being able to recognize the whole tetragram. Finally, we calculated a gamma correlation between the amount of partial information recalled and recognition performance. The high gamma correlation observed reflects a strong relationship between the retrieval of partial information and ultimate recognition performance. The results revealed no reliable differences between the group of schizophrenia patients and the group of healthy participants for any of the correlations.

3.1.3 Discussion

Taken together, the results showed that the schizophrenia patients achieved a lower global recall performance, and in parallel they also produced lower mean FOK ratings, as predicted by the accessibility model (Koriat, 1993). In addition, the present study demonstrated that, for a given memory target, a similar level of recall of related partial information was accompanied by similar FOK ratings in schizophrenia patients as in healthy control participants. Moreover, for both groups of schizophrenia patients and healthy participants, the FOK estimates increased linearly with the amount of partial information retrieved. The products of memory retrieval were predictive of both their accurate insight related to memory and their subsequent memory performance. This means that the relationship between the subjective prediction of knowledge and the products of memory retrieval observed by Koriat (1993) in healthy participants is preserved in schizophrenia patients. Despite memory impairments, patients were capable of relying on the products of memory retrieval to monitor accurately their judgments of what they do or do not know.

3.2 Strategic Regulation of Learning and JOL Accuracy in Schizophrenia Patients

Metamemory functioning during the acquisition of information has been explored in schizophrenia by Bacon et al. (2007). In typical metamemory experiments designed to study episodic memory encoding, participants are usually told to memorize pairs of words in order to recall the target word when they will subsequently be presented with the cue word (Nelson & Narens, 1990). They are then asked to make JOL predictions regarding the likelihood of recalling the target word during the eventual test (Nelson & Dunlosky, 1991). Memory control is sometime assessed using a measure of study time allocation. In healthy participants, the allocation of study time is compatible with the JOLs. For example, in the absence of time pressure, they spend more time studying difficult items than easy ones (Son & Metcalfe, 2000).

The aim of the second study (for a complete description, see Bacon et al., 2007) presented in this chapter was to assess the respective contributions of monitoring and control processes to the strategic regulation of episodic memory. The frequency of presentation of the to-be-learned items was varied, and memory monitoring and control were assessed using measures of JOLs and of study-time allocation, respectively (procedure adapted from Moulin, Perfect, & Jones, 2000). It was predicted that because patients with schizophrenia experience difficulties in taking into account contextual cues related to the learning episode as a result of working memory impairments or long-term memory impairments, or both (Bazin, Perruchet, Hardy-Bayle, & Feline, 2000; Cohen & Servan-Schreiber, 1992; Danion et al., 1999; Gras-Vincendon et al., 1994; Moritz et al., 2003; Rizzo, Danion, Van der Linden, & Grangé, 1996), they should produce lower JOLs than control participants. These patients may also allocate study time inappropriately when exposed to repeated items. This prediction is based on previous evidence that schizophrenia is associated with a problem in using subjective experience to regulate control behavior (Danion et al., 2001a).

3.2.1 Method

In the study there were 19 chronic clinically stable schizophrenia outpatients and 19 healthy participants (control group). Diagnoses were based on DSM-IV. The healthy participants were matched with the patients for age, gender, and educational level.

The items to be learned consisted of 30 weakly associated word-pairs adapted from Ferrand and Alario (1998) word-pair database. One third of the pairs were presented once, one third was presented twice and the last third was presented three times during the learning session.

A computerized version of the tasks was used. Participants were tested individually in the presence of the experimenter. The word-pairs appeared one by one on the screen and the participants were required to read them aloud. They were told that they were in control of the duration of presentation of each pair, that is, of their allocation of study time. The participants were instructed to study the words until they felt they had maximized their chance of remembering the second word of the pair later when they would be required to recall it in response to the presentation of the first word.

After the learning session, a 4 min retention interval was spent carrying out a nonverbal distracting task. Then, the first half of each word-pair (cue) was presented without the second half (target). Participants were asked to indicate their delayed JOL ratings (Nelson & Dunlosky, 1991) for their capacity to recall the target word when presented with the cue word, via a 100-mm visual analogue scale displayed on the screen. The participants produced their estimates by clicking with the mouse on the corresponding bar. Immediately after this, there was a recall phase in which the participants had to either provide the target word or say "I don't know".

3.2.2 Results and Discussion

JOL Ratings

A 2(group) × 3(Repetition) repeated measures ANOVA was conducted on JOL ratings for all the items. The results revealed a main effect of group, with healthy participants having significantly higher JOL ratings (M=65.8%, SD=22.8) than schizophrenia patients (M=35.4%, SD=21.8%), $F(1, 36)=33.4$, $p<0.001$. The main effect of repetition was also significant, $F(2, 72)=46.2$, $p<0.001$, with the JOL ratings being greater for items that were presented twice (M=53.5%, SD=25.7%), than those presented only once (M=34.8%, SD=23.1%), $F(1, 72)=38.0$, $p<0.001$, and greater also for three (M=63.5%, SD=24.2%), than for two item presentations, $F(1, 72)=11.0$, $p<0.01$. The Group × Repetition interaction was not significant, $F(2, 72)=0.3$, $p=0.73$.

Predictive Value of JOL Ratings on Recall

The predictive value of JOL ratings during the recall phase was assessed using the Goodman-Kruskall gamma correlation. The patients' gamma coefficients were slightly, but not significantly, lower ($\gamma=0.90$) than those of the control participants ($\gamma=0.95$), $t(34)=2.2$, $p=0.15$. The high gamma coefficient values indicated a close agreement between the metamemory judgments and the true memory performance.

Allocation of Study Time

A 2(group) ×3(Repetition) repeated measures ANOVA was conducted on study time (in second) of the items. The main effect of repetition on the allocation of study time was significant, $F(2, 72) = 15.7$, $p < 0.001$. The main effect of group was not, $F(1, 36) = 0.1$, $p = 0.78$. The Group × Repetition interaction was significant, $F(2, 72) = 3.2$, $p < 0.05$, that is, healthy participants spent less time ($M = 6.0$, $SD = 2.3$) exploring word-pairs on their second presentation than on their first ($M = 6.8$, $SD = 1.6$), $F(1, 72) = 13.0$, $p < 0.001$; they also spent less time ($M = 5.5$,

$SD = 2.6$) on the third presentation of a pair than on the second, $F(1, 72) = 4.3$, $p < 0.05$. In contrast, the amount of time allocated by patients with schizophrenia was less sensitive to repetition, that is, there were no significant differences in the time spent on each word-pair either between the first ($M = 6.5$, $SD = 1.8$) and the second presentation ($M = 6.3$, $SD = 2.3$), $F(1, 72) = 1.4$, $p = 0.25$, or between the second and the third presentation ($M = 6.1$, $SD = 2.4$), $F(1, 72) = 1.0$, $p = 0.32$. In addition, they did not spend the maximum learning time available.

3.2.3 Discussion

The aim of this study was to gain a better understanding of how patients with schizophrenia strategically process an extrinsic cue, that is, how they manage the repetitions during learning sessions, and to study their accuracy in monitoring their knowledge at the end of an encoding procedure. Repeated items were better recalled by both groups, but the memory performance of schizophrenia patients was always lower than that of healthy participants. Patients' patterns of behavior were abnormal when considering the study time allocated as a function of repetition, as patients did not adapt the study time allocated to each item in response to the frequency of its presentation. In addition, none of the patients reported using efficient strategies to help memorize target items. However, our results also show that patients' JOLs were lower than those of healthy participants but remained sensitive to item repetition. Their predictive value for memory accuracy and for the decision to respond or abstain at the recall phase was not significantly different from those observed in the healthy participants. These results argue in favor of impaired strategic regulation of episodic memory encoding in schizophrenia. They also confirm that schizophrenia induces dissociation between the preserved monitoring and the impaired control processes involved in the encoding phase of a memory task.

4 General Discussion

4.1 Memory Monitoring

It has been observed that schizophrenia patients produce lower prospective metamemory judgments than the healthy participants, whether in tasks assessing episodic memory (Bacon et al., 2007; Bacon & Izaute, 2009; Souchay et al., 2006) or semantic memory (Bacon et al., 2001). Despite these lower FOK ratings, however, patients were in general still able to differentiate between what they know and what they do not know since the predictive accuracy of their monitoring for future recall or recognition was most of the time preserved (Bacon et al., 2001, 2007; Bacon & Izaute, 2009).

According to the accessibility model (Koriat, 1993, 1995; Koriat, Levy-Sadot, Edry, & De Marcas, 2003), metamemory judgments are inferential in nature. To account for the observation that patients with schizophrenia underestimate their future ability to recall an item at retrieval time, when they are required to express their FOK of non-recalled items, some authors suggested that patients have access to only a reduced number of contextual cues relevant to the target items (Bacon et al., 2001). A reduction of the retrievability of contextual cues has indeed been evidenced in the first study presented here (Bacon & Izaute, 2009). However, the relationship between partial information retrieval, FOK ratings, and recognition ability remained remarkably preserved in schizophrenia patients.

A reduction of the retrieval of contextual cues could also account for the evidence reported in the second study presented in this chapter, that is, patients with schizophrenia displayed lower JOLs than their matched healthy participants, the patients being less confident of their future performance level than the healthy participants. Koriat (1997) has demonstrated that when producing JOLs, healthy participants do not directly monitor the strength of the memory trace for the target item. Instead, they monitor a variety of contextual cues which are predictive of their subsequent memory performance. The global quantity of contextual cues determines the magnitude of the JOL. These contextual cues may take the form of intrinsic, extrinsic and mnemonic cues. Intrinsic cues relate to the characteristics of an item such as its perceived relative difficulty. Extrinsic cues relate, for example, to the encoding operations employed by the participants or to the conditions under which encoding was performed, such as the number of times an item could be studied, as in the second experiment presented here. The number of extrinsic cues may be reduced because patients with schizophrenia, who exhibit marked episodic memory impairments, may forget that they have already seen some pairs during the course of the study.

This type of contextual memory deficit has already been reported in schizophrenia (Gras-Vincendon et al., 1994) and may be related to working memory impairments, long-term memory impairments, or both (Bazin et al., 2000; Cohen & Servan-Schreiber, 1992; Moritz et al., 2003; Rizzo et al., 1996). In addition, the quantity of extrinsic cues relating to encoding operations may be reduced in patients with schizophrenia. Indeed, none of the patients in these studies reported using efficient strategies to help them remember the word-pairs, such as making associations between target words and personal autobiographical recollections. The observation that schizophrenia is accompanied by an impairment of strategic processing during encoding provides further support for this interpretation (Danion et al., 1999; Huron & Danion, 2002; Huron et al., 1995; Koh et al., 1973; Traupmann, 1975). Regarding JOL, the contextual memory deficit and the impairment of strategic processing are two mechanisms that may have induced a reduction of the global quantity of extrinsic cues related to certain items and leading to a reduced JOL estimate. Consequently, it would be of interest in a future study to investigate whether immediate JOLs, which are less dependent on contextual memory than delayed JOLs, are also impaired in patients with schizophrenia.

Mnemonic cues are subjective and internal (Koriat, 1997). In the case of metacognitive judgments such as FOK, JOL, or retrospective confidence, several mnemonic cues have been considered, including the accessibility of pertinent information, the ease with which information comes to mind, cue familiarity, the ease of processing of a presented item, the memory of its ease of acquisition, and the memory for the outcome of previous recall attempts. Each of these internal cues can support a heuristic for predicting future recall. In contrast to extrinsic cues (such as number of presentations of an item) it is plausible that mnemonic cues (such as the familiarity of the items) are preserved in patients. Indeed, the evidence is that memory recognition, based on familiarity, is intact in patients with schizophrenia (Danion et al., 1999; Huron & Danion, 2002; Huron et al., 1995). A possible explanation of the observation that patients' JOLs remained sensitive to item repetition is that the patients primarily evaluated the recallability of items on the basis of a non-analytical, implicit inference that took advantage of the preserved mnemonic cues. An adequate relationship between the retrieval of mnemonic cues and FOK ratings has been observed in schizophrenia patients in a task assessing semantic memory (Bacon & Izaute, 2008; Koriat, 1995). These observations argue for a preservation of the accessibility to contextual information as a basis for prospective metamemory ratings in schizophrenia patients, in spite of their memory impairments and of lower metamemory monitoring.

Other processes involved in schizophrenia patients' memory monitoring did indeed seem to be preserved. The patients' accuracy of metamemory judgments in predicting recall or recognition, as assessed by the gamma correlations, was good in both studies. The first study showed a preservation of the gamma correlation between FOK and memory accuracy, between FOK and partial information retrieval, as well as between partial information and recognition ability. The second study showed a preservation of the relationship between JOL ratings and recall performance, the gamma correlation being slightly but not significantly lower in the patients than in the healthy participants. This finding indicates that patients were as successful as healthy participants in subjectively assessing the correctness of their future answers. Also, the effect of item repetition on JOL ratings was similar in patients and in controls. Patients had lower memory performances but they exhibited the normal pattern of responses, with more frequently presented items being rated as more easily retrievable and being better recalled than pairs that were viewed only once. These observations suggest that the patients' memory monitoring remained sensitive to item repetition.

Thus, taken together, the data presented here (Bacon et al., 2007; Bacon & Izaute, 2009) do not support the idea that the reduced metamemory judgments of schizophrenia patients compared to controls can be explained in terms of diminished monitoring ability.

4.2 Memory Control

The control aspect of metamemory processes was explored in the second study, where participants controlled themselves the study time allocated to each item as a function of the repetition of the item and were free to give an answer or to abstain in the recall phase. It was observed that the repetition of items induced different behaviors in schizophrenia patients and healthy participants. The healthy participants spent less time studying a word-pair on the second and third presentation than on the first. Patients did not regulate their study time as a function of repetition to the same extent. Therefore, when assessing the effect of item repetition on JOLs and on study time, dissociation was revealed in patients between memory monitoring, which was preserved, and memory control, which was to a certain extent impaired. There are several, non-mutually exclusive, explanations for the finding that patients did not adapt the allocation of study time as a function of repetition in the same way as the healthy participants. First, as already postulated, it is possible that the patients with schizophrenia forgot that they had already seen some pairs during the course of the experiment. They could not therefore make use of explicit, deliberate memory control of the study time. However, it must be noted that they did not spent the maximum time available for learning, and they provided higher JOLs for items repeated two or three times.

Second, another explanation needs to be discussed with reference to a study revealing dissociation between memory monitoring and memory control in Alzheimer's patients (Moulin et al., 2000) – results in the latter study were strikingly different from those observed in the present study in patients with schizophrenia. Although, like the healthy participants, Alzheimer's patients allocated less time to the study of repeated items, their JOL ratings were insensitive to presentation frequency. It can, therefore, be concluded that memory control was intact in these Alzheimer's patients whereas memory monitoring was impaired. This type of dissociation suggests that the allocation of study time does not necessarily involve deliberate memory control processes, but may rather reflect an automatic response to item repetition (Moulin et al., 2000). It is possible that this type of automatic process is impaired in schizophrenia. Taken together, the results observed in Alzheimer's patients and patients with schizophrenia suggest the possibility of a double dissociation between the monitoring and control processes (i.e., impaired monitoring with preserved control for Alzheimer's patients and impaired control with preserved monitoring for schizophrenia patients during learning).

However, the schizophrenia patients' control of their behavior at the time of retrieval, that is, produce a response or abstain, was consistent with the output of their monitoring. The gamma correlation between JOL ratings and the production of an answer during recall was not significantly different between the two groups, thus indicating that the patients were able to remain faithful to their monitoring in order to control their recall behavior and thus decide whether or not to respond. The patients were also quicker to produce "I don't know" responses despite the fact that, like the healthy participants, they took the same amount of time to provide correct and incorrect answers. This type of behavior is also compatible with the monitoring ratings, that is, patients decided not to produce items to which they had previously attributed lower JOLs. However, it should be noted that, because patients spend less time searching for an answer than healthy participants before they decided to give up, the likelihood of an answer being retrieved was reduced and this may have contributed to the retrieval impairments observed in our patients with schizophrenia.

5 Conclusion

To conclude, the two studies showed preservation of the accuracy of prospective metamemory judgments in schizophrenia patients. The first study demonstrated that the accuracy of FOK, the judgments elicited at the time of retrieval regarding the future recallability of unrecalled items, is preserved in an episodic task. Evidence from the second study indicates that the accuracy of judgments elicited at the time of encoding, namely JOLs, is also relatively preserved in patients and that the correlation between metamemory judgment and memory accuracy is high. Hence, the accessibility model of metamemory as a basis for FOK ratings seems to explain schizophrenics' performance. Patients do indeed rely on the products of memory retrieval to rate their FOK just as healthy participants do. Schizophrenia patients, on the other hand, had lower prospective JOLs because their judgments were based on a small amount of contextual cues. Therefore, the primary deficit leading to underestimated judgments would be a memory deficit, that is, a deficit in memory for contextual information. This is in agreement with the general observation that schizophrenia patients present specific deficits in context processing.

The findings of the second study suggest that it is mostly the control aspect of metamemory that is impaired in schizophrenia patients. Regarding memory control, the extent to which the decision to provide an answer is affected by the confidence level has been shown to be impaired in a semantic memory task (Danion et al., 2001a) and in the Wisconsin Card Sorting Test (Koren et al., 2004). The use of an episodic memory task showed that the decision to provide an answer in the recall task is related to JOLs, whereas the allocation of study time as a function of an extrinsic cue (the repetition of learning) is not related to the monitoring.

Taken together, the data of the two studies suggest that schizophrenia is not associated with a global, non-specific impairment of metamemory processes, but that different patterns of impairments are observed depending on the type of memory task (semantic vs. episodic), the phase of the task (encoding vs. retrieval) and the type of instructions (more or less demanding for metamemory abilities). Nevertheless, beyond these differences, a common feature of the metamemory impairment observed in schizophrenia may be the dissociation between the metacognitive experience based on the monitoring of one's knowledge and the control of memory or behavior. This dissociation between conscious awareness and behavior has also been observed during encoding in an episodic memory task (second study) and during retrieval in a semantic memory task (Danion et al., 2001a). It could represent a shared mechanism underlying the impairment of the strategic regulation of memory functioning in schizophrenia. In the case of schizophrenia patients, the dissociation between preserved monitoring and spared memory control might be explained in terms of a lack of self-initiation of strategy (Medalia, Dorn, & Watras-Gans, 2000). This dissociation, which is reminiscent of Bleuler's idea that schizophrenia is characterized by a splitting of thought and action, could be a hallmark of the disease (Knoblich, Stottmeister, & Kircher, 2004).

Within this perspective, assessing memory and metamemory in patients with schizophrenia should help guide the choice of the care strategy that is most appropriate for these patients. This is all the more true since, according to Nelson, Stuart, Howard, and Crowley (1999), metamemory processes should be considered when deciding on the therapeutic care to be offered to these patients, in particular because they enable them to better control their thoughts, emotions and behavior. As a result, our observations concerning the ability of schizophrenia patients to evaluate their learning performances accurately might help guide patients toward modes of rehabilitation that focus more fully on ways of taking account of metamemory cues in order to optimize the employed strategies and the achieved recall performances (Carter et al., 2008; Dunlosky et al., 2007).

Acknowledgements The authors wish to thank Jean-Marie Danion for helpful discussions and Anastasia Efklides for her precious remarks and editorial work. The studies reported in this chapter were supported by CNRS, INSERM, the University Hospital of Strasbourg and the Blaise Pascal University of Clermont-Ferrand. The first study was also supported by a grant of the FondaMental Foundation, foundation of scientific cooperation.

References

Aleman, A., Hijman, R., de Haan, E. H., & Kahn, R. S. (1999). Memory impairment in schizophrenia: A meta-analysis. *American Journal of Psychiatry, 156*, 1358–1366.

Bacon, E., Danion, J. M., Kauffmann-Muller, F., & Bruant, A. (2001). Consciousness in schizophrenia: A metacognitive approach to semantic memory. *Consciousness and Cognition, 10*, 473–484.

Bacon, E., & Huet, N. (2005). Metamemory knowledge and beliefs in schizophrenia. *Schizophrenia Bulletin, 31*, 349.

Bacon, E., & Izaute, M. (2008). What do patients with schizophrenia know when they cannot remember? *Schizophrenia Research, 98*, 23–24.

Bacon, E., & Izaute, M. (2009). Metacognition in schizophrenia: Processes underlying patients' reflections on their own episodic memory. *Biological Psychiatry, 66*, 1031–1037.

Bacon, E., Izaute, M., & Danion, J. M. (2007). Preserved memory monitoring but impaired memory control during episodic encoding in patients with schizophrenia. *Journal of the International Neuropsychological Society, 2*, 219–227.

Baker, C. A., & Morrison, A. P. (1998). Cognitive processes in auditory hallucinations: Attributional biases and metacognition. *Psychological Medicine, 28*, 1199–1208.

Bazin, N., Perruchet, P., Hardy-Bayle, M. C., & Feline, A. (2000). Context-dependent information processing in patients with schizophrenia. *Schizophrenia Research, 45*, 93–101.

Bonner-Jackson, A., Haut, K., Csernansky, J. G., & Barch, D. (2005). The influence of encoding strategy on episodic memory and cortical activity in schizophrenia. *Biological Psychiatry, 58*(1), 47–55.

Bruder, G. E., Wexler, B. E., Sage, M. M., Gil, R. B., & Gorman, J. M. (2004). Verbal memory in schizophrenia: Additional evidence of subtypes having different cognitive deficits. *Schizophrenia Research, 68*(2–3), 137–147.

Bryson, G., Whelahan, H. A., & Bell, M. (2001). Memory an executive function impairments in deficit syndrome schizophrenia. *Psychiatry Research, 101*(1), 247–255.

Burglen, F., Marczewski, P., Mitchell, K. J., van der Linden, M., Johnson, M. K., Danion, J. M., & Salamé, P. (2004). Impaired performance in a working memory binding task in patients with schizophrenia. *Psychiatry Research, 125*(3), 247–255.

Calev, A. (1984a). Recall and recognition in chronic nondemented schizophrenics: Use of matched tasks. *Journal of Abnormal Psychology, 93*, 172–177.

Calev, A. (1984b). Recall and recognition in mildly disturbed schizophrenics: The use of matched tasks. *Psychology and Medicine, 14*, 425–429.

Carter, C. S., Barch, D. M., Buchanan, R. W., Bullmore, E., Krystal, J. H., Cohen, J., Geyer, M., Green, M., Nuechterlein, K. H., Robbins, T., Silverstein, S., Smith, E. E., Stauss, M., Wykes, T., & Heinssen, R. (2008). Identifying cognitive mechanisms targeted for treatment development in schizophrenia: An overview of the first meeting of the cognitive neuroscience treatment research to improve cognition in schizophrenia initiative. *Biological Psychiatry, 64*, 4–10.

Cartwright-Hatton, S., & Wells, A. (1997). Beliefs about worry and intrusions: The Meta-Cognitions Questionnaire and its correlates. *Journal of Anxiety Disorders, 11*(3), 279–296.

Christensen, B. K., Girard, T. A., Benjamin, A. S., & Vidailhet, P. (2006). Evidence for impaired mnemonic strategy use among patients with schizophrenia using the part-list cuing paradigm. *Schizophrenia Research, 85*, 1–11.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99*, 45–77.

Danion, J. M., Rizzo, L., & Bruant, A. (1999). Functional mechanisms underlying impaired recognition memory and conscious awareness in patients with schizophrenia. *Archives of General Psychiatry, 56*, 639–644.

Danion, J. M., Gokalsing, E., Robert, P., Massin-Krauss, M., & Bacon, E. (2001a). Defective relationship between subjective experience and behavior in schizophrenia. *American Journal of Psychiatry, 158*, 2064–2066.

Danion, J. M., Meulemans, T., Kauffmann-Muller, F., & Vermaat H. (2001b). Intact implicit learning in schizophrenia. *American Journal of Psychiatry. 158*(6), 944–948.

Daprati, E., Franck, N., Georgieff, N., Proust, J., Pacherie, E., Dalery, J., & Jeannerod, M. (1997). Looking for the agent: An investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition, 65*(1), 71–86.

Dickinson, D., Iannone, V. N., Wilk, C. M., & Gold, J. M. (2004). General and specific cognitive deficits in schizophrenia. *Biological Psychiatry, 55*, 826–833.

Driesen, N. R., Leung, H. G., Calhoun, V. D., Constable, R. T., Gueorguieva, R., Hoffman, R., Skudlarski, P., Goldman-Rakic, P. S., & Krystal, J. H. (2008). Impairment of working memory maintenance and response in schizophrenia: Functional magnetic resonance imaging evidence. *Biological Psychiatry, 64*(12), 1026–1034.

Dunlosky, J, Cavallini, E., Roth, H., McGuire, C. L., Vecchi, T., & Hertzog, C. (2007). Do self-monitoring interventions improve older adult learning? The *Journals of Gerontology: Series B. Psychological sciences and social sciences, 62*, 70–76.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*(4), 277–287.

Ferrand, L., & Alario, F. X. (1998). Normes d'association verbales pour 366 noms d'objets concrets [Verbal association norms for 366 concrete objects]. *L'Année Psychologique, 9*, 659–709.

Fioravanti, M., Carlone, O., Vitale, B., Cinti, M. E., & Clare, L. (2005). A meta-analysis of cognitive deficits in adults with a diagnosis of schizophrenia. *Neuropsychological Review, 15*(2), 73–95.

Flavell, J. H. (1971). First discussent's comments: What is memory development the development of? *Human development, 14*, 272–278.

Flavell, J. (1979). Metacognitive and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34*, 906–911.

Fujii, D. E., Wylie, A. M., & Nathan, J. H. (2004). Neurocognition and long-term prediction of quality of life in outpatients with severe and persistent mental illness. *Schizophrenia Research, 69*(1), 67–73.

Frith, C. D. (1989). The neural basis of hallucinations and delusions. *Comptes Rendus Biologies, 328*(2), 169–175.

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. London: Psychology Press.
Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross-classifications. *Journal of the American Statistical Association, 49*, 732–764.
Gras-Vincendon, A., Danion, J. M., Grange, D., Bilik, M., Willard-Schroeder, D., Sichel, J. P., & Singer, L. (1994). Explicit memory, repetition priming and cognitive skill learning in schizophrenia. *Schizophrenia Research, 13*, 117–126.
Green, M. F. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia? *American Journal of Psychiatry, 153*, 321–330.
Heinrichs, W. R., & Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. *Neuropsychology, 12*, 426–445.
Huron, C., & Danion, J. M. (2002). Impairment of constructive memory in schizophrenia. *International Clinical Psychopharmacology, 17*, 127–133.
Huron, C., Danion, J. M., Giacomoni, F., Grangé, D., Robert, P., & Rizzo, L. (1995). Impairment of recognition memory with, but not without, conscious recollection in schizophrenia. *American Journal of Psychiatry, 152*, 1737–1742.
Kircher, T. T. J., Koch, K., & Stottmeister, F. (2007). Metacognition and reflexivity in patients with schizophrenia. *Psychopathology, 40*, 254–260.
Knoblich, G., Stottmeister, F., & Kircher, T. (2004). Self-monitoring in patients with schizophrenia. *Psychological Medicine, 34*, 1561–1569.
Koh, S. D., Kayton, L., & Berry, R. (1973). Mnemonic organization in young nonpsychotic schizophrenics. *Journal of Abnormal Psychology, 81*, 299–310.
Koh, S. D., & Peterson, R. A. (1978). Encoding orientation and the remembering of schizophrenic young adults. *Journal of Abnormal Psychology, 87*, 303–313.
Koren, D., Seidman, L. J., Goldsmith, M., & Harvey, P. D. (2006). Real-world cognitive and metacognitive dysfunction in schizophrenia: A new approach for measuring (and remediating) more "right stuff". *Schizophrenia Bulletin, 32*(2), 310–326.
Koren, D., Seidman, L. J., Poyurovsky, M., Goldsmith, M., Viksman, P., Zichel, S., & Klein, E. (2004). The neuropsychological basis of insight in first-episode schizophrenia: A pilot metacognitive study. *Schizoprenia Research, 70*, 195–202.
Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*, 609–639.
Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124(3)*, 311–333.
Koriat, A. (1997). Monitoring one's knowledge during study: A cue utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, 349–370.
Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*(3), 490–517.
Koriat, A., Levy-Sadot, R., Edry, E., & De Marcas, S. (2003). What do we know about what we cannot remember? Accessing the semantic attributes of words that cannot be recalled. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 6*, 1095–1105.
Kraus, M. S., & Keefe, R. S. (2007). Cognition as an outcome measure in schizophrenia. *British Journal of Psychiatry, 50 (Supplement)*, 46–51.
Laroi, F., & Van der Linden, M. (2005). Metacognitions in proneness towards hallucinations and delusions. *Behaviour Research and Therapy, 43*, 1425–1441.
Lewis, R. (2004). Should cognitive deficit be a diagnostic criterion for schizophrenia? *Journal of Psychiatry and Neurosciences, 29*, 102–113.
Liddle, P. F., & Morris, D. L. (1991). Schizophrenic symptoms and frontal lobe performance. *British Journal of Psychiatry, 158*, 340–345.
Lobban, F., Haddock, G., Kinderman, P., & Wells, A. (2002). The role of metacognitive beliefs in auditory hallucinations. *Personality and Individual Differnces, 32*, 1351–1363.
Lysaker, P. H., Carcione, A., Dimaggio, G., Johannesen, J. K., Nicolò, G., Procacc, M., & Semerari, A. (2005). Metacognition amidst narratives of self and illness in schizophrenia: Associations with neurocognition, symptoms, insight and quality of life. *Acta Psychiatrica Scandinavica, 112*, 64–71.

Medalia, A., Dorn, H., & Watras-Gans, S.W. (2000). Treating probem-solving deficits on an acute care psychiatric inpatient unit. *Psychiatry Research*, 97(1), 79–88.
Metcalfe, J., & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. Cambridge, MA: The MIT Press.
Mohamed, S., Rosenheck, R., Swartz, M., Stroup, S., Lieberman, J. A., & Keefe, R. S. E. (2008). Relationship of cognition and psychopathology to functional impairment in schizophrenia. *American Journal of Psychiatry, 165*, 978–987.
Morice, R. (1990). Cognitive inflexibility and pre-frontal dysfunction in schizophrenia and mania. *British Journal of Psychiatry, 157*, 50–54.
Morrison, A. P., Haddock, G., & Tarrier, N. (1995). Intrusive thoughts and auditory hallucinations: A cognitive approach. *Behavioural and Cognitive Psychotherapy, 23*, 265–280.
Moritz, S., & Woodward, T. S. (2006). The contribution of metamemory deficits to schizophrenia. *Journal of Abnormal Psychology, 115*, 15–25.
Moritz, S., Woodward, T. S., & Rodriguez-Raecke, R. (2006). Patients with schizophrenia do not produce more false memories than controls but are more confident in them. *Psychological Medicine, 36*(5), 659–667.
Moritz, S., Woodward, T. S., & Ruff, C. C. (2003). Source monitoring and memory confidence in schizophrenia. *Psychological Medicine, 33*, 131–139.
Moritz, S., Woodward, T. S., Withman, C., & Cuttler, C. (2005). Confidence in errors as a possible basis for delusions in schizophrenia. *Journal of Nervous and Mental Disease, 193*, 9–16.
Moulin, C. J. A., Perfect, T. J., & Jones, R. W. (2000a). The effects of repetition on allocation of study time and judgements of learning in Alzheimer's disease. *Neuropsychologia, 38*, 748–756.
Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95*, 109–133.
Nelson, T. O. (1993). Judgments of learning and the allocation of study time. *Journal of Experimental Psychology: General, 122*, 269–273.
Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102–116.
Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect". *Psychological Science, 2*, 267–270.
Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study time during multitrial learning. *Psychological Science, 5*, 207–213.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation, 25*, 125–173.
Nelson, T. O., Stuart, R. B., Howard, C., & Crowley, M. (1999). Metacognition and clinical psychology: A preliminary framework for research and practice. *Clinical Psychology and Psychotherapy, 6*, 73–79.
Osatuke, K., Ciesla, J., Kasckow, J. W., Zisook, S., & Mohamed, S. (2008). Insight in schizophrenia: A review of etiological models and supporting research. *Comprehensive Psychiatry, 49*(1), 70–77.
Ranganath, C., Minzenberg, M. J., & Ragland, D. J. (2008). The cognitive neuroscience of memory function and dysfunction in schizophrenia. Biological Psychiat*ry, 64*, 18–25.
Rizzo, L., Danion, J. M., Van der Linden, M., & Grangé, D. (1996). Patients with schizophrenia remember that an event has occurred, but not when. *British Journal of Psychiatry, 168*, 427–431.
Schwartz, B. L., Rosse, R. B., & Deutsch, S. I. (1993). Limits of the processing view in accounting for dissociations among memory measures in a clinical population. *Memory and Cognition, 21*, 63–72.
Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 204–221.

Souchay, C., Bacon, E., & Danion, J. M. (2006). Metamemory in schizophrenia: An exploration of the feeling-of-knowing state. *Journal of Clinical and Experimental Neuropsychology, 28*(5), 828–840.

Sponheim, S. R., Steele, V. R., & McGuire, K. A. (2004). Verbal memory processes in schizophrenia patients and biological relatives of schizophrenia patients: Intact implicit memory, impaired explicit recollection. *Schizophrenia Research, 71*(2–3), 339–348.

Thoma, P., Hennecke, M., Mandok, T., Wähner, A., Brüne, M., Juckel, G., & Daum, I. (2009). Proverb comprehension impairments in schizophrenia are related to executive dysfunction. *Psychiatry Research*, 170(2–3), 132–139.

Traupmann, K. L. (1975). Effects of categorization and imagery on recognition and recall by process and reactive schizophrenics. *Journal of Abnormal Psychology, 84*, 307–314.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26*, 1–12.

Waters, F. A., Maybery, M. T., Badcock, J. C., & Michie, P. T. (2004). Context memory and binding in schizophrenia. *Schizophrenia Research, 68*(2–3), 119–125.

Chapter 8
The Realism in Children's Metacognitive Judgments of Their Episodic Memory Performance

Carl Martin Allwood

1 Introduction

This chapter deals with the metacognitive realism in children's confidence judgments of their own episodic memory performance. *Metacognitive realism* is sometimes called metacognitive accuracy and is defined as the extent to which a metacognitive judgment is veridical with respect to the asserted status of the learning or of the recall (Lichtenstein, Fischhoff, & Phillips, 1982). After some remarks on metacognition, the realism in metacognition and its measurement, the chapter focuses on some factors that can influence the realism in confidence judgments of the correctness of one's episodic recall. Two studies are then presented more in detail in order to deepen the analysis of some of the factors envisaged. The chapter also relates the level of realism in children's confidence judgments to that of adults.

Metacognition is usually described as our knowledge about our own cognition, including the use of this knowledge to regulate our own cognitive processes (Weinert & Kluwe, 1997). Sometimes it is also taken to include knowledge about other people's cognition (Allwood & Granhag, 1999; Allwood & Johansson, 2004; Jost, Kruglanski, & Nelson, 1998). The approach on metacognition taken in this chapter, just as in current research (Koriat, 2007), sees metacognition in a system perspective where metacognitive processes are integrated into the individual's other cognitive processes and where they are also affected by various social processes taking place outside the individual. For example, retelling the event to different persons after the experience and taking part of one's listeners' reactions to one's story is likely to influence both the correctness of later recalls of the event and one's confidence that the recall is correct.

C.M. Allwood (✉)
Department of Psychology, University of Gothenburg, Box 500,
SE-40530 Gothenburg, Sweden
e-mail: cma@psy.gu.se

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_8, © Springer Science+Business Media, LLC 2010

1.1 Realism in Confidence Judgments

Confidence judgments are judgments of the veridicality of one's own or other's memory reports. At a general level the *realism* in confidence judgments is a function of the relation between the correctness of performance (as measured against a socially approved standard) and the performer's confidence in the correctness of his or her performance. As pointed out by Leippe (1980), although there are factors that may influence both the correctness of the memory report and the person's confidence that the memory report is correct (such as good performance conditions in general), some factors may predominantly influence the level of correctness whereas others may foremost influence the level of confidence. For example, an individual's confidence level may partly be a function of the individual's general level of expression of confidence over time; that is, individuals tend to have a certain stability in the level of confidence they express (see, e.g., Jonsson & Allwood, 2003).

There are at least two aspects to the realism of metacognitive judgments (for a much more complete presentation of these issues, see Yates, 1994). The two aspects can be distinguished through the use of many measurements for an individual or a group of individuals. These are *calibration* and *discrimination*. The calibration aspect pertains to the relation between the *level of confidence* in the correctness of the memory recall and the *level of correctness* in the same memory recall. When these two levels coincide the judge is said to show perfect calibration. The other aspect, discrimination, pertains to the individual's ability to discriminate between *correct* and *incorrect* items by means of his or her confidence judgments. Each of these two aspects can be measured in different ways.

Appendix 8.1 shows some common measures used in the calibration tradition branch of metacognitive research. The measures shown are calibration, over-/underconfidence, resolution, and slope. The first two of these measures relate to the calibration aspect. Of the last two, resolution relates to the discrimination aspect. Slope measures the separation between confidence for correct and incorrect items and has an advantage in that it may be more intuitively easy to understand than resolution. It picks up discrimination ability but also some of the calibration aspect.

2 Factors Influencing the Realism in Confidence Judgments

A host of different factors can influence metacognitive realism in confidence judgments (see, e.g., Allwood & Granhag, 1999; Jonsson & Allwood, 2003; Klayman, Soll, Gonzáles-Vallejo, & Barlas, 1999). Table 8.1 presents some factors that can be assumed to influence the realism in confidence judgments in episodic memory recall.

Although not mentioned in Table 8.1, these factors do not preclude the influence of cognitive processing biases (Koriat, Lichtenstein, & Fischhoff, 1980) such as the confirmation bias, and methodological and statistical factors (Erev, Wallsten, & Budescu, 1994; Gigerenzer, Hoffrage, & Kleinbölting, 1991; Juslin, 1994) such as

Table 8.1 Factors that may influence the realism in children's confidence judgments of their episodic performance

Event
The individual's general level of expressed confidence
Intermediate events (communication with others, etc.)
Social situation in which memories are reported and confidence judgments are made
Memory question asked (type of question, content in question, e.g., central/peripheral aspect)
Measurement scale
Aspect of realism in confidence analysed (e.g., calibration or discrimination)

biased selection of items, on confidence judgments of episodic memory reports. Various cognitive processing biases and methodological and statistical factors have received researchers' attention, especially in the context of semantic memory and for adults, but are less well researched for episodic memory and children. However, they fall outside of the focus of the present chapter.

Initially it can be noted that various features of the to-be-remembered event and its context, such as visibility and other encoding conditions are likely to influence the correctness of the memory recall and thus also the realism of the confidence judgments (see, e.g., Leippe & Eisenstadt, 2007). However, this factor (Factor 1 in Table 8.1) is not elaborated in the present chapter. Furthermore, as noted above, individual stability in the general level of confidence judgments (Factor 2 in Table 8.1) as such may also influence the realism of confidence judgments but this also falls outside the scope of the present chapter.

2.1 Events Intervening Between the Original Event and the Memory Report

Different *events intervening* in the time between the original event and when the memory is reported and confidence judged may influence the realism of the confidence judgments. Importantly, as noted above, social conversations of various types, that is, talk with other persons, could have an influence. For example, when the child reports and discusses the event with his/her family, friends and other categories of persons such as the police or other parties in a forensic process, various memory consequences are likely to take place. One important consequence is that the memory of the event is actively rehearsed verbally, and this is likely to increase the correctness in future recall of the memory (Roediger & Karpicke, 2006).

However, given that the event is not just reported but also *discussed*, both correctness and confidence may be influenced (see, e.g., Marsh, 2007). Correctness might be influenced, for example, due to the fact that the conversation partners ask questions that express or imply erroneous assumptions that are then encoded in the person's memory in such a way that these assumptions or their implications cannot be distinguished from the original event.

In addition, previous research shows that confidence in the correctness of a statement may increase as an effect of reasserting it. This has been called the *reiteration effect* (Hertwig, Gigerenzer, & Hoffrage, 1997). Thus, each time a memory is reasserted, for example when retelling the event or answering questions about it, the child's confidence in the correctness of the memory report may increase.

The effect of intermediate discussions of the event on the rememberer's correctness and confidence was investigated in students by Sarwar, Allwood, and Innes-Ker (2010a). In that study repeated retellings of the events in the film (approaching simple *repetition* of the experienced event) were found to increase correctness, confidence and the realism (as measured by the calibration measure, see Appendix 8.1) in the confidence judgments of the correctness of the reported memories in the context of a later open free recall task. In contrast, multiple *discussions* of the event with (each time a new) other person reduced the effects of repetition in the later free recall task, that is, compared with a control condition, discussions did not significantly affect any of the mentioned measures. To these authors' knowledge no similar study has been made for children.

2.2 Social Aspects of the Memory Report Situation

As shown in Table 8.1, *the type of social situation* in which the child recalls and then confidence judges the correctness of the recalled information might also influence the realism of the confidence judgments. How the child experiences and understands the situation is important here. For example, situations vary with respect to the extent to which the child feels expected to only report correct information from memory. In fact, the child may be given an explicit instruction to only provide information that they are absolutely sure is correct, or, alternatively, to report anything that might be true even if they are not sure about it. For example, when testifying in court a child can in ordinary situations be assumed to attempt to only provide correct information and for this reason choose not to report memories about which they feel unsure about. On other occasions the child might apply a less stringent threshold for reporting memories; for example, in free time discussions when they want to impress their friends.

Koriat and Goldsmith (1996) presented a model for how confidence judgments are integrated in ordinary memory recall. The point of the model that is relevant in the present context is that it assumes that the rememberer uses confidence judgments to regulate which memories are reported. This is accomplished by the implementation of a variable threshold for how sure he or she wants to be that the reported memories are correct. Accordingly, when a person can choose what memories to report (i.e., using Koriat and Goldsmith's (1996) term, they have *free report option*) they can themselves attempt to regulate the assumed proportion of correct memories.

For example, Koriat, Goldsmith, Schneider, and Nakash-Dura (2001) analysed 7–12 year-old children's answers to specific questions with and without answer alternatives about a slide show that they had seen. The results showed that when the child could control what information to report, correctness tended to increase, and completeness decrease compared with when they had to/were pressed to answer. To sum up, various constraints in the social situation where the recall and confidence judgments take place can affect the level of both correctness and confidence.

2.3 The Memory Question Asked

The factors presented in Table 8.1 also bring attention to the importance of the memory question asked for the ensuing realism of the confidence judgments. For example, the degree of veracity of the information provided in the memory question is important. Previous research has shown that children have difficulties with misleading questions (Roebers, 2002; Roebers & Howie, 2003). For example, Roebers and Howie (2003) studied 8- and 10-year-old children's and adults' discrimination of correct and incorrect answers to memory questions on specific aspects of a short film clip by means of their confidence judgments. In this research unbiased and misleading questions (questions that suggested an incorrect answer) were compared. Misleading questions were more difficult for the children to handle. For *unbiased questions* all three age groups gave higher confidence judgments for correct answers than for incorrect answers. However, for *misleading questions* only adults gave higher confidence judgments for correct answers than for incorrect answers.

There are also a number of other important aspects to the question asked. First, the question asked can vary with respect to *how much information* it provides about the to-be-reported memory (for example, free recall or recognition questions). Everything else being equal, recognition questions usually promote higher correctness than open specific questions since they provide more cues to the answer.

Second, questions differ with respect to how broad the assigned *answer-area* is that the question allows. Everything else being equal, the broader the assigned answer-area the more report control the remembering person has. For example, in open free recall questions the person is given a general indication about which area to report on ("Tell me all that happened on the Monday afternoon…") and within this area it is up to the person's own discretion what exactly to report. For more specific questions such as "What was the colour of the girl's jacket?" the assigned answer-area is much smaller.

Third, the *giving of report option* can as such, at least partly, be manipulated independently of how much information is provided about the to-be-reported memory and how large the assigned answer-area is. For example, for a set of recognition questions, the person may be told, or not, that he or she can choose which of the questions he/she wants to answer.

A fourth and final aspect of the memory question that can affect the realism in confidence judgments is the *type of contents* asked for. For example, Sarwar et al. (2010b) found indications that central information, such as the culprit's features or actions, may not only be better remembered but also more realistically confidence judged than peripheral information, such as details not immediately relevant to the central action or to the actors. This may to a large extent be a function of how the person's attention is allocated during the encoding of the original event and of the person's prior knowledge about the contents (see, e.g., Christianson & Loftus, 1991; Ibabe & Sporer, 2004).

2.4 The Measurement Scale Used and the Aspect of Metacognitive Realism Analyzed

As shown in Table 8.1, it has also been suggested that the specific confidence scale used to measure children's confidence may affect their ability to give realistic confidence judgments (Roebers & Howie, 2003). This suggestion is supported by the findings reported by Tunney and Shanks (2003), that is, students who used a binary confidence scale showed better realism in their confidence judgments of fairly implicit knowledge compared with participants who used a continuous scale.

A number of different confidence scales have been used in previous research to measure children's confidence (e.g., Allwood, Granhag & Jonsson, 2006a; Dirkzwanger, 1996; Newman & Wick, 1987; Roebers, 2002). One reason for this variation is the controversy concerning how complex tasks and scales children can handle at different ages. For example, numerical scales might be more (or even, too) complex for younger children compared with scales using qualitative steps, such as "Very unsure", "Not so sure", "Neither unsure nor sure" "Pretty sure" and "Very sure" (used by, e.g., Roebers, 2002).

Roebers and Howie (2003) suggested that one reason why younger children (e.g., 8-year olds) might perform poorly on metacognitive tasks could be that they are tested with age-inappropriate scales, specifically, scales that are too complex. Other inappropriate features of scales, such as use of *smiley faces* with broader smiles for higher confidence levels were also remarked upon by these authors. Roebers (2007) suggested that scales with fewer scale steps, for example three, would be easier for young children to handle. (The issue of scale inappropriateness is further discussed below, in Sects. 3 and 6.)

Finally, the factors presented in Table 8.1 suggest that the aspect of metacognitive realism analyzed (e.g., the calibration or the discrimination aspect) is likely to influence the conclusions drawn about the level of realism in children's metacognitive performance. For example, as discussed below in Sect. 5, children and adults may differ for some tasks in the calibration aspect but not in the discrimination aspect.

Next, two studies are reviewed that investigated the importance of two of the factors presented in Table 8.1 for the resulting metacognitive realism. These factors are the measurement scale used when the participants give their confidence ratings and the memory question asked.

3 The Effect of the Confidence Scale Used on Children's Metacognitive Realism

Allwood, Granhag, and Jonsson (2006a) analyzed the effect of the confidence scale used to measure the children's confidence in their memories of an experienced event. The participants were 81 children (41 girls and 40 boys) aged 11–12 years from Grades 5 to 6, in schools located in a middle class area in southern Sweden.

Four confidence scales were investigated: (a) Numeric, (b) Picture, (c) Line, and (d) Verbal scale. These scales are shown in Appendix 8.2. The rationale for including the respective scales in the study was as follows. The Numeric scale is common in calibration research with adult participants and it was included in order to allow for comparison with results from such studies. Picture scales (smilies, etc.) are common in research with younger children (Roebers, 2002). A Picture scale was included in order to be able to compare the results from this scale with those from the Numeric scale. The use of the Line scale was inspired by results reported by Nilsson (1998, p. 97), who concluded that younger children's (6-year-olds and to some extent 10-year-olds) handling of probabilities might be influenced by "perceptual factors such as size, shape and colour." The Line scale was included in order to examine if the participants' confidence ratings would be improved by a scale that highlights spatial aspects. Finally, Teigen (2001) presented results for adults that showed that written probability phrases tended to differ from numerical probabilities. Given this, we included the Verbal scale since we wanted to examine if the inclusion of written probability statements would influence the level of children's confidence judgments. Based on a review of earlier research we predicted that no difference would be found between the scales with respect to the level of the confidence ratings, nor with respect to the realism in the confidence judgments.

3.1 Method

A between-subjects design was used, that is, the participants were randomly divided into four conditions (i.e., *numeric*, *picture*, *line* and *verbal*) and in each condition one of the four respective confidence scales was used.

3.1.1 Procedure

The participants first watched a videoclip (approximately 4 min long) showing the kidnapping of a woman by two men who pulled her into a car by force. After viewing the videotape, the participants were given a 10 min training session on probability assessments. In this training the participants were provided with general explanations about probability estimates. For example, it was explained that a scale value of 60% meant that in the long run 60% of the items they had confidence rated as

Table 8.2 Means (and *SD*) of correctness, confidence, calibration, over-/underconfidence, and resolution for the four scales (numeric, picture, line, and verbal)

	Scale			
	Numeric	Picture	Line	Verbal
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Correctness	0.59 (0.09)	0.58 (0.08)	0.56 (0.08)	0.57 (0.08)
Confidence	0.81 (0.09)	0.82 (0.08)	0.76 (0.10)	0.78 (0.07)
Calibration	0.10 (0.06)	0.10 (0.05)	0.09 (0.05)	0.09 (0.05)
Over-/underconfidence	0.22 (0.13)	0.24 (0.11)	0.20 (0.12)	0.22 (0.09)
Resolution	0.04 (0.02)	0.03 (0.02)	0.03 (0.02)	0.03 (0.02)

60% sure should be correct. In addition, the children solved practical examples, guided by the experimenter.

Next, each participant answered 44 two-alternative directed questions on various details in the video (forced choice). After each question, the participants immediately rated their confidence in the correctness of the answer on a rating scale ranging from 50% (guessing) to 100% (completely sure).

3.2 Results

The results for the various dependent measures are shown in Table 8.2 for each of the four confidence scales. Since the random chance for selecting the correct answer was 50%, the results for correctness (56–59%) shows that the questions were rather difficult. As expected, the results showed no significant differences between the four scales for correctness or for confidence.

The calibration aspect of realism in metacognition was shown by the specific measures calibration and over-/underconfidence. The discrimination aspect was shown by means of the resolution measure (see Appendix 8.1). Again, as expected, neither of these measures showed any significant differences between the four confidence scales.

4 The Effect of Question Type on Children's Metacognitive Realism

The study by Allwood, Innes-Ker, Homgren, and Fredin (2008) analyzed the effect on children's metacognitive realism of asking open free recall (henceforth called *free recall*) and specific directed questions (henceforth called *focused questions*) about an experienced event. We also analyzed the effect of repetition of answers on the realism in confidence judgments. We expected that both the children and the adults would show equal, and good, levels of metacognitive realism with respect to the over-/underconfidence measure for their free recall. We also expected poor metacognitive realism in the confidence judgments of both children and adults of the answers to the focused questions.

4.1 Method

The same procedure was used in two experiments. The only difference was that a different film clip was used as the to-be-remembered event. Information about the participants is shown in Table 8.3.

4.1.1 Procedure for Experiment 1 and 2

Each experiment had three sessions. In Session 1 the participants first saw a short 3–4 min videotape. In Experiment 1 the videotape was about a man who was looking for his lost dog in a park. The video in Experiment 2 was the same as the video used by Allwood, Granhag et al. (2006a) showing a kidnapping event. This video clip is likely to have been more complex than the video used in Experiment 1.

In Session 2, 1 week after Session 1, each participant was tested individually and all interviews in this session were audio-taped. First, the participants were asked to give a *free recall* of the events in the video clip they had watched 1 week before. The participant was asked to tell everything he or she could remember about the events in the video and the people in it, with as much detail as possible. Next the participant answered a questionnaire with 39 (Experiment 1) or 44 (Experiment 2) two-alternative forced-choice questions on specific details in the video (i.e., the focused questions). Two examples of questions used in Experiment 1 are "Was there a sandbox in the park shown in the film? Answer: (a) Yes, (b) No" and "How many swings were there in the park? Answer: (a) 3, (b) 5".

During the week between Session 2 and 3, the free recall data were segmented into elementary statements and further prepared for the confidence judgments in Session 3. When doing this segmentation the researchers followed the guidelines presented in Allwood, Ask, and Granhag (2005). A numeric confidence rating scale was placed under each elementary statement. The participants' answers to the focused questions were similarly prepared for the confidence rating in Session 3. For each question, a numeric confidence rating scale was placed immediately under the two answer alternatives (whereof the participant had selected one).

Table 8.3 Participants in Experiment 1 and 2 in the Allwood et al. (2008) study

8–9-year-olds	12–13-year-olds	Adults
Experiment 1		
31 (20 girls)	31 (15 girls)	32 (21 women); $M = 25$ years, Range = 19–56 years
Experiment 2		
43 (20 girls)	52 (24 girls)	38 (25 women); $M = 25$ years, Range = 18–46 years

Session 3 took place in the children's classrooms. First, the children were given a 10 min explanation about confidence ratings. This explanation included specific examples and a detailed explanation about what a confidence judgment is. Next, the participants first confidence rated each item in the questionnaire with the statements in their free recall and then their answers to the focused questions. The confidence rating scale for both types of ratings went from 0 (I am sure that the answer is wrong), via 50 (I guess) to 100 (I am sure that the answer is correct). The same procedure but with suitable adjustments, was used for the adults; for example, the explanation about confidence ratings was shorter.

4.2 Results Experiment 1

Table 8.4 shows some important results for the free recall task and for the focused questions task in Experiment 1. The results are reported in more detail in Allwood et al. (2008). The last column in Table 8.4 shows the outcome of planned contrasts between the three age groups. We first consider the results for the *free recall*.

First, as shown in Table 8.4, the analysis of the total number of statements recalled showed that the 8–9-year-olds and the 12–13-year-olds on average recalled significantly less statements than the adults. From youngest to oldest, the three age groups recalled 12.1, 12.3, and 15.6 statements, respectively. Moreover, both the correctness and the confidence levels were quite high. From 74 to 84% of the statements in each age group were located at the 100% confidence level and over 90% of these were correct. No age differences were found for correctness and the 12–13-year-olds tended to show the highest confidence.

We next look at the metacognitive measures. Here it is most noteworthy that the youngest age group had close to perfect realism in the over-/underconfidence measure.

Table 8.4 Experiment 1: Means (*SD*) and number of participants [*n*] in the three age groups for the free recall and for the focused questions

	8–9-year-olds	12–13-year-olds	Adults	
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*p*
Free recall				
Number of statements	12.1 (4.3)	12.3 (4.6)	15.6 (3.9)	1, 2,<3, $p<0.05$
Correctness	91.0% (10.9%)	88.3% (10.5%)	87.8% (11.1%)	*ns*
Confidence	90.2% (9.5%)	94.4% (6.1%)	92.1% (5.9%)	$p=0.08$
Over-/underconfidence	−0.005 (0.093)	0.065 (0.100)	0.042 (0.101)	1<2, $p<0.05$
Slope [*n*]	13.5 (21.7) [19]	10.5 (21.1) [23]	3.7 (22.3) [26]	*ns*
Focused questions				
Correctness	71.7% (6.7%)	74.8% (10.5%)	73.5% (6.6%)	*ns*
Confidence	78.4% (13.0%)	85.0% (8.5%)	74.1% (8.7%)	1<2>3, $p<0.05$
Over-/underconfidence	0.068 (0.131)	0.101 (0.110)	0.008 (0.085)	1, 2,>3, $p<0.05$
Slope [*n*]	7.8 (11.4) [31]	13.9 (10.6) [31]	12.8 (6.4) [32]	1, 2,>3, $p<0.05$

Note: 1 = 8–9-year-olds; 2 = 12–13-year-olds; 3 = Adults

However, the other age groups also showed quite good values. *Slope* was used as a measure of ability to use confidence judgments to separate correct from incorrect answers (see Appendix 8.1). However, as shown in sharp brackets in Table 8.4, only 19 (61%), 23 (74%), and 26 (81%) of the participants in each age group (starting with the youngest) had any errors at all. Still, it can be noted that no age differences were found.

The results for the *focused questions* are also shown in Table 8.4. Here the confidence judgments were much more evenly spread over the confidence scale, but with concentrations at the 50% and the 100% level. No age differences were found for correctness and the 12–13-year-olds showed the highest confidence. For the metacognitive measures it is noteworthy that the adults demonstrated next to perfect realism with respect to the degree of overconfidence in contrast to the two child groups who showed overconfidence. The youngest age groups showed poorer slope than the other two groups.

It can be noted that the 12–13-year-olds and the adults showed less overconfidence for the focused questions than similar groups in our previous research (e.g., Allwood, Granhag, et al., 2006a; Allwood, Granhag, & Johansson, 2003). Two possible reasons for this are, first, that the contents of the film had a rather simple and monotonous structure which may have made the task simpler. Second, *repetition* of some assertions in the focused questions that were already made in the free recall may have caused an increase in confidence for these items due to the reiteration effect (Hertwig et al., 1997). (Seen from a forensic point of view, the non-overlapping items, that is, the answers to the focused questions not already mentioned in the free recall, are the most interesting since they may provide new information compared with the witnesses' free recall.)

To study the effect of repetition in the focused questions, the data for the focused questions was analyzed again. This time items already mentioned by a participant in the free recall were not included for that participant. Over all participants, 14.2% of all focused questions overlapped with some content mentioned in the free recall (522 questions out of 3,666 questions). The interjudge reliability for this coding was 88% and there were no age differences in number of overlapping items.

Table 8.5 shows the results for the focused questions when the focused questions that overlapped with content mentioned in the free recall are excluded. As can be

Table 8.5 Experiment 1: Means (*SD*) and number of participants [*n*] in the three age groups for the focused questions when the questions overlapping between the free recall and the focused questions are excluded

	8–9-year-olds	12–13-year-olds	Adults	
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*p*
Correctness	58.2% (7.1%)	62.1% (9.1%)	59.4% (7.2%)	*ns*
Confidence	76.5% (12.8%)	83.3% (9.4%)	71.6% (9.0%)	1 < 2 > 3, $p = 0.05$
Over-/underconfidence	0.185 (0.122)	0.215 (0.130)	0.123 (0.099)	1, 2 > 3, $p < 0.05$
Slope [*n*]	1.3 (9.4) [31]	5.1 (10.2) [31]	1.9 (10.2) [32]	*ns*

Note: 1 = 8–9-year-olds; 2 = 12–13-year-olds; 3 = Adults

noted by comparing Tables 8.4 and 8.5 one effect of excluding overlapping questions was that the level of correctly answered questions decreased by about 13%. However, the level of the confidence judgments was fairly unaffected and thus, there was no clear sign of a reiteration effect in this fairly realistic test of this effect. Due to the decrease in correctness the level of overconfidence increased substantially by about 0.12. Finally, the slope measure decreased to nearly zero for all groups. These results also show that when the overlapping items are not included in the results for the focused questions there was a clear format difference between the free recall task and the focused questions task.

4.3 Results Experiment 2

First, the analysis of the total number of statements recalled showed that the 8–9-year-olds on average recalled significantly less statements (6.4) than both the 12–13-year-olds (9.8) and the adults (17.6). The latter two groups also differed in the number of recalled statements.

Table 8.6 shows that correctness was again high in the *free recall* task. For the 8–9-year-olds and the adults it was over 90%. Again a high percentage of the items (76–87%) were located at the 100% confidence level and over 90% of these were correct, except for the 12–13-year-olds who scored about 85%. Furthermore, the results showed that the 12–13-year-olds had significantly poorer correctness than the other two age groups, but still high at 84%.

For the metacognitive measures it is again striking that the youngest age group had close to perfect realism for the over-/underconfidence measure, but the adults

Table 8.6 Experiment 2: Means (*SD*) and number of participants [*n*] in the three age groups for the free recall and for the focused questions

	8–9-year-olds	12–13-year-olds	Adults	
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*p*
Free recall				
Number of statements	6.4 (3.6)	9.8 (3.8)	17.6 (9.9)	1, 2<3, 1<2 $p<0.05$
Correctness	92.5% (15.0%)	83.8% (12.4%)	90.7% (8.2%)	1>2, 2<3 $p<0.05$
Confidence	91.4% (13.1%)	93.4% (7.5%)	93.4% (7.5%)	*ns*
Over-/underconfidence	–0.004 (0.158)	0.096 (0.116)	0.034 (0.084)	1<2, 2>3, $p<0.05$
Slope [*n*]	11.7 (34.5) [12]	10.5 (20.5) [44]	28.3 (36.8) [28]	2<3, $p<0.05$
Focused questions				
Correctness	54.4% (8.7%)	56.7% (6.3%)	54.4% (7.0%)	*ns*
Confidence	73.5% (13.0%)	68.9% (10.1%)	72.0% (9.4%)	*ns*
Over-/underconfidence	0.194 (0.169)	0.122 (0.109)	0.176 (0.117)	1>2, $p<0.05$
Slope [*n*]	1.5 (7.7) [43]	4.1 (5.7) [52]	3.2 (7.2) [38]	*ns*

Note: 1 = 8–9-year-olds; 2 = 12–13-year-olds; 3 = Adults

also showed quite good realism. With respect to slope only 28, 84, and 74% of the participants in each age group (starting with the youngest) had any errors and thus were available for the analysis. In this experiment the adults had better slope than, especially, the 12–13-year-olds.

For the *focused questions* it should be noted that the adults showed higher overconfidence than in previous research (unclear why). This deviation from previous research can not be explained by the repetition of items that took place between the free recall and the focused questions since an analysis showed that over all participants, only 1.7% of all focused questions overlapped with some content mentioned in the free recall (103 of 5,779 items). This coding had 91% interjudge reliability and ANOVAs showed no difference from the results compared with when all items were included for the focused questions.

5 Comparison of the Realism in Children's and Adults' Confidence Judgments

Metacognitive ability is usually assumed to improve as children get older. For this reason it is of interest to see what empirical research shows with respect to this issue in the domain of confidence judgments of event memory. Most previous studies on the realism in children's confidence judgments have used some form of specific directed questions. Next, we will compare the results from children with the results from adults (students) in previous studies using the same film clip and the same confidence scale, and when the comparison concerns focused questions, also the same questions.

5.1 Overconfidence

The value for overconfidence for adults (students) in the relevant conditions in Allwood, Granhag, and Johansson (2003), Allwood, Knutsson, and Granhag (2006b), Granhag (1997) and Granhag, Strömwall, and Allwood (2000) varied between $M = 0.061$ and $M = 0.127$. In contrast, the overconfidence for the children in Allwood, Granhag, et al. (2006a) was on average 0.22 (see Table 8.2).

The study reviewed above by Allwood et al. (2008) is not strictly comparable with these studies for the focused questions since these questions in Allwood et al. (2008) were preceded by a free recall session. However, the results for the free recall in Experiment 2 in that study can be used for comparison since the same film was used. As reviewed above, it was shown that there was no significant difference between the children and the adults (the youngest children showed next to perfect realism for this measure).

These results suggest that before conclusions about developmental age differences are drawn in metacognitive research it is important to consider how well practiced the participants are at the specific task they are asked to perform. When the task can be assumed to be well-trained (for example regulating the veracity of the output from memory in response to a free recall instruction), children can be expected to perform well, often at the level of adults.

5.2 Discrimination and Separation Measures

Most of the above mentioned previous studies have also included measures of discrimination. Here a comparison of the results for children and adults show no great differences for the *resolution* measure (a measure of discrimination ability) when *focused questions* were used. The results generally show a resolution level of between 0.03 and 0.04 for both groups (Allwood et al., 2003; Allwood, Granhag, et al., 2006a; Allwood, Jonsson, & Granhag, 2005b; Allwood, Knutsson, et al., 2006b; Granhag, 1997).

The study by Allwood et al. (2008) used another measure, *slope* that in addition to separation as given in terms of scale values, also indicates discrimination. These results are less reliable for the free recall since the high performers were lost in the analyses due to their lack of incorrect items. The results for the focused questions in Experiment 1 favoured the adults but this effect disappeared when the items that overlapped between the free recall and the focused questions were eliminated from the analysis. Experiment 2 showed no significant differences between the age groups. In brief, summarized over both resolution and slope, the indications of age differences for the focused questions for the discrimination aspect are quite meagre. However, the finding by Roebers and Howie (2003), reviewed above, that for *misleading questions* only adults gave higher confidence judgments for correct answers than for incorrect answers, may be a sign of weakness in children's separation ability, as compared with adults'.

6 Level of Noise in Children's and Adults' Confidence Judgments

One possible explanation for the difference observed between children and adults in overconfidence is that children in the investigated ages are *generally less skilled* in using the confidence scale. If children are poor at handling the confidence scale, one might expect the error component in the children's confidence judgments to be affected. The size of the standard deviation for individual participants' confidence judgments can be taken to be an indicator of such an error component. Everything else being equal, when the standard deviation for an individual participant's confidence judgments is higher the error component can be assumed to be greater. In this context it is also relevant that Erev et al. (1994) argued convincingly that greater variability in confidence judgments is associated with greater overconfidence. Thus, taken together and in brief, the notion is that the children's greater overconfidence

Table 8.7 Mean within-subject standard deviations (and *SD*) for the confidence judgments for the three age groups for Experiment 1 and 2 in the Allwood et al. (2008) study

	8–9-year-olds	12–13-year-olds	Adults	
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*p*
Experiment 1				
Free recall	14.6 (11.9)	10.8 (11.1)	13.6 (9.3)	*ns*
Focused questions	22.1 (9.2)	18.5 (6.3)	20.0 (3.7)	*ns*
Focused questions when excluding overlapping questions	22.5 (8.7)	18.5 (6.4)	19.3 (3.7)	1 > 2, $p < 0.05$
Experiment 2				
Free recall	10.1 (12.3)	10.4 (11.1)	14.5 (12.7)	*ns*
Focused questions	21.8 (9.0)	20.6 (5.8)	19.4 (4.1)	*ns*

Note: 1 = 8–9-year-olds; 2 = 12–13-year-olds; 3 = Adults

for focussed questions compared with the adults in previous studies might be explained by the fact that children, due to their lesser skill in handling the confidence scale, have larger individual variability in their confidence judgments, compared with adults.

To explore this possibility Allwood, Granhag, et al. (2006a) computed the standard deviations for each child's confidence judgments in the different scale conditions. The average individual standard deviations in each condition were for the Numerical scale = 16.57, for the Picture scale = 14.64, for the Line scale = 15.82, and for the Verbal scale = 15.29. All these values are lower than the average within-subject standard deviation for the confidence judgments (17.69) in the most similar adult comparison group in our previous research (Allwood et al., 2003; Exp. 2, Phase 1). As shown in Table 8.7, similar analysis of the data reported in Allwood et al. (2008) also did not support this idea since there were no significant differences between the adults and any of the two child groups with respect to these standard deviations.

Thus, taken together, the data reported by Allwood, Granhag, et al. (2006a) and by Allwood et al. (2008) do not support the idea that the children's higher overconfidence compared with the adults can be explained by their poorer ability to handle the confidence rating task as such, at least not as indicated by the presence of larger individual standard deviation for their confidence judgments.

7 Discussion

This review of research on the realism in children's confidence judgments of their episodic memory performance has shown that the level of realism in confidence judgments is influenced by a number of different factors. However, the results reviewed also suggest that certain aspects of the confidence scale used may not be as important for the outcome in realism as previously speculated. The results reported by Allwood, Granhag, et al. (2006a) showed that the realism in children's confidence judgments for focused questions demonstrate a fair amount of stability for the four different types of confidence scales tested.

The review also showed that the results of the comparison of the realism in confidence judgments between children and adults are somewhat complex. For example, the outcome of this comparison depends on the type of memory question given and the aspect of realism investigated. Comparison of the results for children from 8 to 13 years of age and adults (students) for *focused questions* showed fairly clearly that children are more overconfident in their confidence judgments of their own episodic memory.

However, the research in Allwood et al. (2008) showed that the same comparison, when made for free recall performance, did not show any difference between children and adults. Here, if anything, there were indications of lesser overconfidence among the younger children (8–9-year-olds). This was demonstrated in two experiments which only varied in the complexity of the episodic event experienced. We had no formal measure of the complexity of the video clips shown but our strong intuitive impression is that the clip used in Experiment 1 was less complex than that shown in Experiment 2.

Furthermore, the aspect of realism considered makes a difference. The results for the discrimination aspect of metacognitive realism in confidence judgments in the research reviewed did not clearly support any differences between the ability of children in middle childhood and adults.

The study by Allwood et al. (2008) did not solve the question of what aspect of the questions asked that was most causally important for the reported metacognitive results. However, two, probably interacting, aspects appear important. First, there was a difference in the used question formats in the degree of report option allowed (Koriat & Goldsmith, 1996). The free recall instruction gave a large degree of report option, that is, control over what information to report (they assigned a larger answer-area). In contrast, the focused questions involved forced report and assigned a smaller answer-area. When the participants can control what to report (i.e., free recall) they may choose information that they are confident is correct (Koriat et al., 2001). As noted above, this assumes that the participants attempt to hold a high threshold for the level of correctly reported items in their memory report. Even the youngest children in Allwood et al. (2008) appear to have had sufficient training to be able to live up to the philosopher Grice's *quality maxim* for communication, which involves the notion "Try to make your contribution one that is true" (see Schwarz, 1996).

However, apart from the usefulness of report option, it can also be argued that the free recall task allowed access to memories that were easily available. In contrast, in the context of the focused questions, the children were forced to answer questions picked by someone else. On average this can be expected to have had the effect that the focused questions asked the children to provide information that was less accessible in memory compared to the free recall task. The poorer correctness level for these questions compared to the one achieved in the free-recall context in combination with the results for the overconfidence measure, suggests that it may be a harder task to provide realistic confidence judgments for items that are less accessible in memory. This point was well illustrated above in connection with the results for those of the focused questions that had not been spontaneously mentioned in the participants' free recall in Experiment 1 in the study by Allwood et al. (2008), that is, the non-overlapping questions.

Here it is also relevant to note that a well-known effect in calibration research called the *hard-easy effect* shows that harder questions (poorer correctness) tend to be associated with higher levels of over-/underconfidence (see, e.g., Juslin, Winman, & Olsson, 2000). Future research is needed to systematically sort out the impact of these factors by attempting to vary the correctness level, report option and question format independently of one another.

Finally, the issues of whether some of the scales used in metacognitive confidence research are inappropriate because they are too complex for younger age groups and whether the confidence judgment task as such is too complex for younger children will be discussed. Clearly, there is obviously an age at which a confidence judgment task of one's own memory report is too demanding. However, a number of arguments suggest that these may not be the most pertinent reasons why children of 8–10 years of age have often been found to show worse metacognitive realism than adults. Although children at this age may not understand all aspects of the probability concept, the same can be said of most adults.

Here it is also of interest to note that Schlottmann and Anderson (1994) reported that even 5-year-olds understood the probability concept better than was expected from previous research. The results reported by Allwood et al. (2008) that children of 8–9 years of age showed excellent performance on the overconfidence measure for open free recall, suggest that the confidence judgment task as such may not be too difficult for this age group. Finally, as discussed above, the within-subject standard deviations for the confidence judgments (an indicator of noise in the confidence judgments) did not differ between the four confidence scales investigated by Allwood, Granhag, et al. (2006a), nor between children and adults (Allwood et al., 2003, Exp. 2 phase 1; Allwood, Granhag, et al., 2006a; Allwood et al., 2008).

8 Appendix 8.1. Some Common Metacognitive Measures in the Calibration Research Tradition Referred to in the Text

Calibration measures the relation between the level of the confidence ratings and the correctness of the memory report. The following formula is used:

$$\text{Calibration} = 1/n\sum_{t=1}^{T} n_t (r_{tm} - c_t)^2$$

Here *n* is the total number of questions answered. *T* is the number of confidence classes used; for example if the confidence scale runs from 50% ("guessing") to 100% ("absolutely sure"), the following six confidence classes (T=6) may be used: 50–59, 60–69, 70–79, 80–89, 90–99, 100). c_t is the percent correct answers of all items in the confidence class r_t (e.g., 50–59). n_t is the number of times the confidence class r_t was used and r_{tm} is the mean of the confidence ratings in confidence class r_t.

Over-/underconfidence is computed like calibration, except that the differences are not squared. Over-/underconfidence shows if an individual is overconfident (positive value) or underconfident (negative value). Calibration and over-/under-confidence are perfect when their values are zero.

Resolution reflects the ability of the person to discriminate between two sets of answers, one correct and one incorrect. The formula is:

$$\text{Resolution} = 1/n\sum_{t=1}^{T} n_t (c_t - c)^2$$

Here, *c* is the percent of all items for which the correct answer was provided. A higher value on this measure reflects better resolution than a lower. These measures are better described in Lichtenstein, Fischoff, and Phillips (1982) and in Yates (1994).

Slope reflects the ability to separate correct from incorrect answers by means of one's confidence judgments. The formula is:

$$\begin{aligned}\text{Slope} = &(\text{Mean confidence for the correct answers}) \\ &-(\text{Mean confidence for the incorrect answers})\end{aligned}$$

9 Appendix 8.2. The Four Confidence Scales Used in the Study (Translated into English)

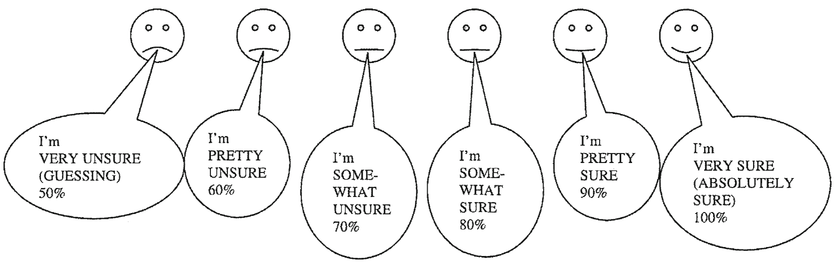
9.1 The Numeric Scale

What was the girl's hair color? (A) Red, (B) Black

I'm __________% sure that I answered the question correct.

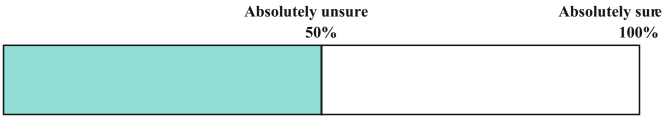
9.2 The Picture Scale

What was the girl's hair color? (A) Red, (B) Black



9.3 *The Line Scale*

What was the girl's hair color? (A) Red, (B) Black



9.4 *The Verbal Scale*

What was the girl's hair color? (A) Red, (B) Black

___ 50% Absolutely unsure (Correct 50 times of 100)
___ 60% Pretty unsure (Correct 60 times of 100)
___ 70% Somewhat unsure (Correct 70 times of 100)
___ 80% Somewhat sure (Correct 80 times of 100)
___ 90% Pretty sure (Correct 90 times of 100)
___ 100% Absolutely sure (Correct 100 times of 100)

References

Allwood, C. M., Ask, K., & Granhag, P. A. (2005). The cognitive interview: Effects on the realism in witnesses' confidence in their free recall. *Psychology, Crime & Law, 11*(2), 183–198.

Allwood, C. M., & Granhag, P. A. (1999). Feelings of confidence and the realism of confidence judgments in everyday life. In P. Juslin & H. Montgomery (Eds.), *Judgment and decision making: Neo-Brunswikian and process-tracing approaches* (pp. 123–146). Hillsdale, NJ: Erlbaum.

Allwood, C. M., Granhag, P. A., & Johansson, M. (2003). Increased realism in eyewitness confidence judgments: The effect of dyadic collaboration. *Applied Cognitive Psychology, 17,* 545–561.

Allwood, C. M., Granhag, P. A., & Jonsson A. C. (2006). Child witnesses' metamemory realism. *Scandinavian Journal of Psychology. 47*(6), 461–470.

Allwood, C. M., & Johansson, M. (2004). Actor-observer differences in realism in confidence and frequency judgments. *Acta Psychologica, 117*, 251–274.

Allwood, C. M., Innes-Ker, Å., Homgren, J., & Fredin, G. (2008). Accuracy and realism in confidence in children's and adults' event recall in response to free recall and focused questions. *Psychology, Crime & Law, 14,* 529–547

Allwood, C. M, Jonsson, A. C., & Granhag, P. A. (2005). The effects of source and type of feedback on child witnesses' meta-memory accuracy. *Applied Cognitive Psychology, 19,* 331- 344.

Allwood, C. M., Knutsson, J., & Granhag, P. A. (2006). Eyewitness under influence: How feedback affect the realism in confidence. *Psychology, Crime & Law,12,* 25–38.

Christianson, S.-Å., & Loftus, E. F. (1991). Remembering emotional events: The fate of detailed information. *Cognition & Emotion, 5*(2), 81–108.

Dirkzwager, A. (1996). Testing with personal probabilities: 11-year-olds can correctly estimate their personal probabilities. *Educational and Psychological Measurement, 56*, 957–971.

Erev, I., Wallsten, T. S., & Budescu, D. V. (1994). Simultaneous over- and underconfidence: The role of error in judgment processes. *Psychological Review, 101,* 519–527.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98,* 506–528.

Granhag, P. A. (1997). Realism in eyewitness confidence as a function of type of event witnessed and repeated recall. *Journal of Applied Psychology, 82,* 599–613.

Granhag, P. A., Strömwall, L., & Allwood, C. M. (2000). Effects of reiteration, hindsight bias, and memory on realism in eyewitnesses' confidence. *Applied Cognitive Psychology, 14,* 397–420.

Hertwig, R., Gigerenzer, G., & Hoffrage, U. (1997). The reiteration effect in hindsight bias. *Psychological Review, 104,* 194–202.

Ibabe, I., & Sporer, S. L. (2004). How you ask is what you get: On the influence of question form on accuracy and confidence. *Applied Cognitive Psychology, 18*(6), 711–726.

Jonsson, A.-C., & Allwood, C. M. (2003). Stability and variability in the realism of confidence judgments over time, content domain, and gender. *Personality and Individual Differences, 34,* 559–574.

Jost, J. T., Kruglanski, A. W., & Nelson, T. O. (1998). Social metacognition: An expansionist review. *Personality and Social Psychology Review, 2,* 137–154.

Juslin, P. (1994). The overconfidence phenomenon as a consequence of informal experimenter-guided selection of almanac items. *Organizational Behavior and Human Decision Processes, 57,* 226–246.

Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect. *Psychological Review, 107,* 384–396.

Klayman, J., Soll, J. B., Gonzáles-Vallejo, C., & Barlas, S. (1999). Overconfidence: It depends on how, what and whom you ask. *Organizational Behavior and Human Decision Processes, 79,* 216–247.

Koriat, A. (2007). Metacognition and consciousness. In P.D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289–325). Cambridge, UK: Cambridge University Press.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,* 490–517.

Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children's testimony: Can children control the accuracy of their memory reports? *Journal of Experimental Child Psychology, 79,* 405–437.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 107–118.

Leippe, M.R. (1980). Effects of integrative memorial and cognitive processes on the correspondence of eyewitness accuracy and confidence. *Law and Human Behavior, 4,* 261–274.

Leippe, M. R., & Eisenstadt, D. (2007). Eyewitness confidence and the confidence-accuracy relationship in memory for people. In R. C. Lindsay, D. F. Ross, J. Don Read, & M. P. Toglia (Eds.), *Handbook of eyewitness psychology, Volume 2, Memory for people* (pp. 377–425). Mahwah, NJ: Erlbaum.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgments under uncertainty: Heuristics and biases* (pp. 306–334). New York: Cambridge University Press.

Marsh, E. J. (2007). Retelling is not the same as recalling: Implications for memory. *Current Directions in Psychological Science, 16*(1), 16–20.

Newman, R. S., & Wick, P. L. (1987). Effect of age, skill, and performance feedback on children's judgments of confidence. *Journal of Educational Psychology, 79*, 115–119.

Nilsson, S.-E. (1998). *Children's decision making under conditions of certainty and risk.* Unpublished doctoral dissertation, Göteborg University, Sweden.

Roebers, C. M. (2002). Confidence judgments in children's and adults' event recall and suggestibility. *Developmental Psychology, 38,* 1052–1067.

Roebers, C. M. (2007). Favourable and unfavourable conditions for children's confidence judgments. *British Journal of Developmental Psychology, 25,* 109–134.
Roebers, C. M., & Howie, P. (2003). Confidence judgments in event recall: Developmental progression in the impact of question format. *Journal of Experimental Child Psychology, 85,* 352–371.
Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory. *Perspectives on Psychological Science, 1*(3), 181–210.
Sarwar, F., Allwood, C. M., & Innes-Ker, Å. (2010a). *Effects of communication with a non-witness on eyewitnesses' recall correctness and metacognitive realism.* Manuscript submitted for publication.
Sarwar, F., Allwood, C. M., & Innes-Ker, Å. (2010b). Effects of repeated recall and discussion on eyewitness' accuracy and meta-memory realism for different types of forensic information. Unpublished manuscript. Lund University and University of Gothenburg, Sweden.
Schlottmann, A., & Anderson, N. H. (1994). Children's judgments of expected value. *Developmental Psychology, 30,* 56–66.
Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods and the logic of conversation.* Hillsdale, NJ: Erlbaum.
Teigen, K. H. (2001). When equal chances = Good chances: Verbal probabilities and the equiprobability effect. *Organizational Behavior and Human Decision Processes, 85,* 77–108.
Tunney, R. J., & Shanks, D. R. (2003). Subjective measures of awareness and implicit cognition. *Memory & Cognition, 31,* 1060–1071.
Weinert, F. E., & Kluwe, R. H. (Eds.). (1997). *Metacognition, motivation and understanding.* Hillsdale, NJ: Erlbaum.
Yates, J. F. (1994). Subjective probability accuracy analysis. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 381–410). New York: Wiley.

Chapter 9
Cognitive Interruption as an Object of Metacognitive Monitoring: Feeling of Difficulty and Surprise

Alexandra Touroutoglou and Anastasia Efklides

1 Introduction

Problem solving is by definition related to lack of an immediate, ready-made response to a problem. This is often associated with experienced difficulty as Dewey (1910) had noted. Despite the importance of experienced difficulty as subjective indicator of the presence of a problem, research has mainly focused on task difficulty and less so on the feeling of difficulty (FOD) the person experiences, presumably because task difficulty can be objectively determined whereas FOD is a subjective state. In fact, the latter is related to objective task difficulty (Efklides, 2001, 2002; Efklides, Papadaki, Papantoniou, & Kiosseoglou, 1997, 1998; Efklides, Samara, & Petropoulou, 1999) but cannot be reduced to it, because FOD is also related to one's ability and self-concept (Efklides et al., 1997, 1998; Efklides & Tsiora, 2002).

Efklides (2001) has conceptualized FOD as the interface between the person and the task. Yet, it is not known what is the mechanism that underlies the formation of FOD, that is, what features of cognitive processing or which cognitive states give rise to it. One possible mechanism that may give rise to FOD is increased working memory load; another mechanism is the monitoring of cognitive interruption. In what follows, three experiments will be reported that aimed at elucidating the cognitive underpinnings of FOD. Before doing this, however, a definition of FOD will be given along with its conceptualization as metacognitive experience. Then two postulated sources of FOD, namely working memory load and cognitive interruption, will be presented. Their cognitive and affective sequelae will be also discussed. Finally, our research on the effects of working memory load and cognitive interruption on performance, reaction time (RT), FOD, and surprise (another affective concomitant of cognitive interruption) will be presented and discussed.

A. Efklides (✉)
School of Psychology, Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece
e-mail: efklides@psy.auth.gr

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_9,

1.1 Feeling of Difficulty as a Metacognitive Experience

Difficulty reflects the obstacles encountered in a problem-solving situation (Davidson, Deuser, & Sternberg, 1994). According to Herbert (1978, p. 540) "something is difficult when it is easy to fail in it", and according to Delignières (1998), difficulty is defined as a set of constraints that limit the possibilities of problem solving.

The term "feeling of difficulty" has been mainly used by Efklides (Efklides, 2001, 2006; Efklides et al., 1997, 1998; Efklides et al., 1999; Efklides & Petkaki, 2005). However, research on subjective difficulty goes back almost 30 years when Borg (1978) used the term "perceived difficulty" (see also Borg, Bratfisch, & Dornič, 1971) to denote the person's perception of task difficulty as distinct from objective task difficulty. Another related term used in metacognition research is "ease of learning" (Nelson, 1996; Nelson & Narens, 1994), which is used to denote a judgment about how easily one can learn or remember the material to be studied. This judgment is made before one starts learning, and is important for the allocation of study time (Leonesio & Nelson, 1990). Indeed, people seem to rely on the ease with which they process initial information to determine how much effort will be required in order to process subsequent information (Alter, Oppenheimer, Epley, & Eyre, 2007, Exp. 2). According to Pintrich, Wolters, and Baxter (2000) judgments of ease of learning are based on the person's metacognitive knowledge of the task and metacognitive knowledge of the self, that is, how well one did in the past on the same or a similar task. Perceived difficulty can also be considered as judgment based on the person's metacognitive knowledge of the task moderated by metacognitive knowledge of the self; however it is in the opposite direction than ease of learning.

More recently, Efklides (2001, 2002, 2005, 2006; Efklides et al., 1999) posited that FOD is a metacognitive experience that monitors cognitive processing as it takes place; 'metacognitive', because it is the outcome of the monitoring of online cognitive processing and conveys information about features of it such as lack of fluency, interruption, or conflict in response formation; 'feeling', because the information it conveys takes the form of a subjective inner state, a subjective experience that has a negative affective quality. Therefore, because of its experiential nature FOD differs from perception of difficulty or judgment of ease of learning, because the latter are metacognitive judgments based on metacognitive knowledge, that is, declarative memory knowledge of task and self vis-à-vis the current task (Flavell, 1979). Of course, if the person tries to make sense of his/her FOD, then metacognitive knowledge of task or self can be called in to support the analytic process of meaning making of one's experiential state.

According to Efklides (2006), FOD is a metacognitive experience manifested when processing cognitively demanding tasks. In such situations, there is lack of fluency in cognitive processing (Efklides, 2002). Research on metacognitive experiences such as feeling of knowing or feeling of familiarity strongly suggests that processing fluency is a powerful cue that contributes to their formation (Koriat, 2007, 2008). However, in the case of FOD, it is the lack of fluency that is critical rather than fluency. Indeed, Efklides (2002) showed that ratings of feeling of familiarity (that denotes fluency) loaded a different factor than ratings of FOD, and the relation between the two was negative. Moreover, estimate of effort (to be) exerted was directly and positively related to FOD, and indirectly with feeling of familiarity – through its relationship with FOD (Efklides, 2002; Efklides et al., 1999; Efklides, Kourkoulou, Mitsiou, & Ziliaskopoulou, 2006).

Furthermore, perception of difficulty and judgment of ease of learning are usually present before actual cognitive processing or learning begins. FOD, however, can occur at any point (before, during, after) of cognitive processing as long as an obstacle or an interruption occurs. The experienced difficulty may result from an inability to understand the task or represent the problem space, from an inability to figure out or plan the solution, from increased demands on resources, such as working memory, and finally, from conflicts or errors that occur and block the formation of response (Delignières, 1998; Efklides, 2001, 2002, 2005, 2006; Efklides et al., 1999).

Thus, the intensity of the FOD may change during problem solving (Efklides, 2002; Efklides et al., 1999). For example, a decrease in FOD may be observed as ideas and insights offer a resolution to the impasse experienced (Davidson et al., 1994). On the contrary, FOD increases when previously successful strategies fail to handle the available evidence and, thus, cognitive processing is interrupted (Mandler, 1975).

When cognitive processing is slowed down or interrupted, this gives rise to negative affect and effort exertion (Efklides, 2001, 2006). FOD is crucial because it informs whether additional effort is needed on one's part (Efklides, 2006). Evidence from psychophysiological research on fluency of cognitive processing strongly suggests that fluency is associated with positive affect (Winkielman & Cacioppo, 2001); lack of fluency is associated with negative affect and effort exertion, as indicated by nonverbal expressions, such as contraction of the corrugator muscle (Hrubes & Feldman, 2001; Stepper & Strack, 1993), and physiological activity, such as cardiovascular and electrodermal reactivity (Brinkmann & Gendolla, 2007; Gendolla, 1999; Gendolla & Krüsken, 2001; Pecchinenda & Smith, 1996).

However, FOD should not be confused with awareness of effort exertion – "estimate of effort" in Efklides's terminology (Efklides, 2001, 2002). FOD is a product of the monitoring of cognitive processing whereas effort is the product of control processes that involve motivational factors, available resources and monitoring of cognitive processing demands, as indicated by the FOD (Efklides et al., 2006). Thus, the estimate of effort captures the awareness of intensity of cognitive activity or the energy put to it but not the reason for it.

Finally, FOD triggers control decisions and strategy use. Alter et al. (2007) showed that metacognitive awareness of difficulty (caused by letter font) led to decreased reliance on quick responses and increased the frequency of more elaborate responses in persuasion, judgment, and syllogistic reasoning tasks – see also Efklides et al. (1999) for change of strategies in response to FOD.

To sum up, despite the importance of the FOD for problem solving, the mechanism that gives rise to this metacognitive experience is not well understood. Experimental work on metacognitive feelings, such as feeling of knowing or feeling of familiarity, has focused on the fluency with which stimulus or memory information is processed, but ignored other processing features, such as increased working memory demands, blockages, conflicts in response formation, or discrepancies and interruptions that induce lack of fluency. Most importantly, the relationship between lack of processing fluency and FOD has been shown in correlational research (Efklides, 2002), rather than being itself an object of experimental scrutiny. Hence, experimental research on the mechanism underlying the formation of FOD becomes an important issue for metacognition research.

1.2 Feeling of Difficulty and Its Underlying Mechanism

In the early 1970s, Borg et al. (1971) stated that performance and task demands are the objective basis of perceptions of difficulty. Perceptions of task difficulty are low for correctly solved problems and high for incorrectly solved problems (Borg et al., 1971). However, as already mentioned, the subjective experience of the FOD cannot be reduced to task demands only, because it is the interaction of task with the person's characteristics (e.g., ability or prior knowledge) that is important for it. Moreover, the experience of FOD may precede the formation of response and performance. Therefore, performance is important for metacognitive knowledge of the task and the self and, through them, for perceptions of difficulty. This explains why experts' perceptions of task difficulty have been repeatedly found to predict novices' performance (Ayres, 2001, 2006; Borg et al., 1971).

The relationship of FOD with performance is not as simple, however. For example, FOD is lower for simple problems as compared to complex ones, but the inverse holds true as well (Efklides, 2002). In this respect, when confronted with a non-routine task, high performance is likely to be associated with high FOD, increased effort, and increased processing time. When the same task gets familiar, high performance is likely to be associated to low FOD and decreased processing time (Efklides, 2002). Thus, FOD can correlate with increased processing time, but this relationship is mediated by the person's characteristics and the context in which the problem is solved (Delignières, 1998).

The above findings indicate that neither performance nor task difficulty per se can fully account for the variability in the FOD. What is, then, the mechanism that triggers FOD?

1.2.1 The Lack of Fluency Hypothesis

According to Efklides (2002, 2006) FOD monitors the lack of processing fluency at various levels. At the perceptual level, lack of fluency – and subsequently FOD – can be experienced in tasks in which stimuli vary in their presentation mode, for example, duration, stimulus luminance, figure-to-ground contrast, size of stimulus, letter font etc. (Alter et al., 2007; Reber, Fazendeiro, & Winkielman, 2002). At the conceptual level, FOD can be experienced in complex or developmentally advanced tasks that make demands on prior knowledge (Efklides et al., 1997, 1998). At the conceptual level, it also occurs when extant knowledge or processing schemas do not match task requirements or compete with each other creating conflict (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Moreover, FOD can be experienced in tasks that pose demands on processing resources such as working memory (Paas, Renkl, & Sweller, 2003). In the latter case, FOD reflects increased cognitive load (Ayres, 2006; Paas, Renkl, et al., 2003a; Sweller, 2003, 2006).

Lack of Fluency Due to Working Memory Load

Complex tasks require the execution of multiple cognitive acts, thus increasing demands on working memory, because the incoming information cannot be chunked or integrated into higher order units and has to be processed individually (Sweller, 2003). In this sense, FOD is due to the multiple cognitive acts performed rather than to lack of relevant schemas. This may explain why complex tasks have been repeatedly found to increase FOD ratings (Ayres, 2001, 2006; Efklides et al., 1997, 1998). Working memory load may also account for increases in FOD in dual tasks in which people have to perform two tasks concurrently (Paas, Tuovinen, Tabbers, & Van Gerven, 2003b; Whelan, 2007).

However, it needs to be pointed out that working memory load seems to implicate different brain mechanisms than those involved in the processing of difficult tasks. Barch et al. (1997) found increased activity in dorsolateral prefrontal cortex, Broca's area, and parietal cortex as a response to greater working memory demands (i.e., longer maintenance of information) but not as a response to greater task difficulty (e.g., stimuli degradation). On the contrary, the anterior cingulate cortex (ACC), which is implicated in metacognition (Fernandez-Duque, Baird, & Posner, 2000), showed increased activity in response to task difficulty, but not to working memory load.

Lack of Fluency Due to Conflict in Processing

Studies on ACC suggest a link between monitoring of lack of fluency and conflict in processing. It has been shown that ACC is activated in a variety of tasks and, particularly, in tasks in which effortful control is required (Botvinick, Cohen, & Carter, 2004; Botvinick et al., 2001; Bush, Luu, & Posner, 2000; Carter et al., 1998; Critchley et al., 2003; Duncan & Owen, 2000; Posner & DiGirolamo, 1998; van Veen & Carter, 2002). According to the conflict monitoring theory of ACC (Carter et al., 1998), lack of fluency in processing is attributed to conflict between simultaneously activated responses. Thus, FOD is experienced in tasks in which a prepotent response has to be inhibited in order to perform the correct response, as for example in Stroop tasks or go/no-go tasks (MacLeod, 2004). In these types of tasks, FOD seems to emerge as a response to a conflict created between incompatible (dominant and non-dominant) responses.

Furthermore, ACC is also activated during task processing in which a number of alternative responses exist and decision making is required, as in stem completion tasks, or in tasks in which participants have to generate a number of potential examples that fall into a certain category, as in verb generation task (Botvinick et al., 2004; Botvinick et al., 2001).

Finally, ACC is also activated in dual tasks (D'Esposito et al., 1995) as well as in speeded tasks (Dehaene, Posner, & Tucker, 1994). In dual tasks, lack of fluency is explained in terms of a conflict between the two simultaneously activated responses. Since most commission errors in speeded tasks are made impulsively, a conflict is also assumed to exist between correct and incorrect immature responses (Botvinick et al., 2001).

Taken together, these findings suggest that FOD arises when there is lack of fluency in: (a) dual tasks and tasks with many steps; (b) complex tasks in which many, attention demanding, acts have to be executed, and (c) conflict resolution tasks. The common characteristic shared by all these types of tasks is that they tend to be familiar, that is, tasks in which participants have available processing schemas, but integration into a new organization is needed or selection between different alternatives.

1.2.2 The Cognitive Interruption Hypothesis

Cognitive Interruption and Its Antecedent Conditions

Unless one has cognitive schemas that are relevant to task requirements, cognitive interruption is bound to occur (Mandler, 1975). A cognitive interruption occurs whenever a current schema fails to handle task requirements (Mandler, 1984). This is usual when the task in which one is working on is unfamiliar and some of its aspects are not well understood, that is, they are discrepant from our schemas.

Mandler (1975) argued that any discrepancy, conflict, or blockage in processing represents a potential source of cognitive interruption. According to the Interruption Theory of emotion (Mandler, 1975), cognitive interruptions result from unexpected events in processing and give rise to affective responses. Affective consequences, however, arise only when both of the following conditions are met: (a) unexpected events are encountered and (b) no appropriate alternative schemas are available (Mandler, 1975). Thus, unexpectedness is a necessary but not sufficient condition for cognitive interruption. Moreover, cognitive interruption occurs only in cases in which discrepancies are not automatically assimilated into extant schemas (Mandler, 1975).

Cognitive Interruption and Its Cognitive Consequences

According to schema theories of learning (Holyoak, Koh, & Nisbett, 1989), neural network computational models of learning (Phaf, Mul, & Wolters, 1994; Rumelhart, Smolensky, McClelland, & Hinton, 1986) and associative theories of learning (Rescorla, 2004; Schultz & Dickinson, 2000), discrepant events set off revision processes. The idea that schema reorganization follows when expectations are disconfirmed traces back to the pioneering work of Bruner and Postman (1949). In their classical study on perception of incongruity, they argued that perceptual reorganization occurs only when well-established expectations fail to be confirmed. Similarly, Posner, Strike, Hewson, and Gertzog (1982) suggested that students will not acquire alternative theories unless they become dissatisfied with the current ones.

In face of cognitive interruption, individuals may discover that their previous schemas (beliefs, theories) are inadequate to explain the new data, and this may lead them to consider or invent alternative hypotheses that can account for the new data. Discrepant events reflect inefficiency of extant schemas (Shallice & Burgess, 1996), and as such, they lead to changes in strategies that tackle the problem situation. However, the presence of cognitive interruption does not guarantee the efficiency of control processes assumed to operate after the detection of discrepancy (Alberdi, Sleeman, & Korpi, 2000). In fact, when confronted with cognitive interruption, errors are very likely to occur, especially in cases in which no alternative response schema is available (Mandler, 1975).

Whether efficient or not, reactions to discrepant events are largely time consuming, and for this reason, cognitive interruption generally increases processing time. Increases in RT, due to discrepancies in processing, have been demonstrated in attention studies (Horstmann, 2002, 2005). This increase in RT is interpreted as representing control processes recruited to safeguard performance (Kahneman, 1973; Norman & Shallice, 1986), for example, assimilation of the unexpected event to a changed schema.

Cognitive Interruption and Its Affective Consequences

As for the experiential aspect of cognitive interruption, it may take various forms. First, the mind seems to have no content (Izard, 1977; see also Moraitou & Efklides (2009) for the experience of blank in the mind). Also, a sense of 'wrongness' emerges, which may lead the person to recognize incongruity (Bruner & Postman, 1949). Furthermore, in the face of interruption, uncertainty emerges as the person lacks the knowledge necessary to decide what course of action to employ next (or whether a course of action will be met with success) (de Hoyos, Gray, & Simpson, 2004).

Besides the above experiences which are metacognitive in nature, positive or negative affect is generated depending on whether the unexpected events enhance or block further processing (Carver, 2003). According to Efklides (2006) and Frijda (1986), FOD is the metacognitive affective response typically generated in such a situation. Specifically, because discrepant events in most cases block the progress toward solution or increase the probability of error, cognitive interruptions are usually associated with frustration and negative affect (Frijda, 1986; Mandler, 1975).

Furthermore, cognitive interruptions are, by definition, unexpected, and for this reason, they may also trigger *surprise*, which is a response to schema-discrepant events (Meyer, Reisenzein, & Schützwohl, 1997). Indeed, according to the

Psychoevolutionary Model of surprise (Meyer et al., 1997; Reisenzein, 2000; Schützwohl, 1998), surprise is an emotional response to unexpectedness. Also, according to the Counterfactual Alternatives Hypothesis of surprise (Kanheman & Miller, 1986), surprise is elicited by a contrast between an event and its counterfactual alternatives constructed after the occurrence of discrepancy. In a similar vein, the Contrast Hypothesis of surprise suggests that surprise depends on the degree of contrast between an expected event and its unexpected outcome (Teigen & Keren, 2003). The link between surprise and cognitive interruption is also suggested by the Representational-Fit model of surprise (Maguire & Keane, 2006), according to which surprise varies with the difficulty with which an unexpected event is integrated into the discourse representation.

1.3 Cognitive Interruption and Feeling of Difficulty

When extant cognitive schemas prove inefficient, overcoming cognitive interruption requires the person to encode task elements in a new manner (Eysenck & Keane, 2005). FOD, thus, is associated with the need to create or select a schema for problem resolution than for applying a solution. Moreover, according to Mandler (1975), the ease with which discrepant events are integrated into extant schemas will finally determine the intensity of FOD. For example, low FOD is assumed to result from slightly incongruous data; higher FOD will emerge from incongruous data that can be assimilated by an alternative schema; even higher FOD will result from severe incongruous data that can be assimilated by a changed schema; and finally, the highest FOD will result from incongruous data that cannot be assimilated by any extant schemas. However, this will occur to the extent that the data are incongruous or discrepant to previous schemas. If there is no incongruity to extant schemas, as in the case of novel tasks, then we are not necessarily experiencing high feeling difficulty.

It can be argued, therefore, that monitoring of cognitive interruption constitutes a critical process that triggers control processes. Since metacognitive or executive processes are particularly involved in novel situations where no schema is immediately available to guide action (Fernandez-Duque et al., 2000), cognitive interruption can be considered as an object of metacognitive monitoring in a similar manner as fluency is (Koriat, 2007). Cognitive interruption, thus, is a cognitive phenomenon (schema failure), but its presence is metacognitively monitored (subjective experience of schema failure that may take the form of FOD). However, the relationship between cognitive interruption and FOD has not been previously addressed.

1.4 The Present Study

In the present study, several task manipulations were made in order to test the two hypotheses regarding the underpinnings of FOD, namely the lack of fluency hypothesis as represented by increased working memory load and the cognitive interruption hypothesis. The difference between the two hypotheses is that in the former case FOD may arise in familiar tasks too, such as tasks in which many acts have to be executed in working memory (Ayres, 2001). In the latter case FOD may arise in response to cognitive interruption caused by discrepant events (Efklides, 2006; Mandler, 1984). Since discrepancies give also rise to surprise, one would expect FOD and surprise to be related in the case of interruption, but not in the case of working memory load.

In Experiment 1, the lack of fluency hypothesis was examined by manipulating the working memory load while participants solved inductive reasoning tasks. The tasks consisted of number sequences that conform to a simple arithmetical rule. In Experiment 2, the working memory load hypothesis was further tested with the same design as in Experiment 1. In Experiment 3 the cognitive interruption hypothesis was examined through inductive reasoning tasks, consisting of number sequences that induced processing schemas and discrepancies. In all experiments the interrelations between metacognitive (i.e., FOD), cognitive (i.e., RT and performance) and affective (i.e., surprise) outcomes of working memory load and cognitive interruption manipulations were examined.

2 Experiment 1: First Test of the Working Memory Load Hypothesis

2.1 Aims and Hypotheses

The aim of Experiment 1 was to investigate the postulation that working memory load results in lack of fluency and increases FOD. Another aim of this experiment was to replicate previous findings that associate FOD with the self-reported effort exerted on task. To address this issue, participants rated their estimates of effort along with FOD.

The tasks, which were used in Experiment 1, involved mental arithmetic. According to DeStefano and LeFevre (2004), a domain in which all components of working memory appear to interact is mental arithmetic. Working memory resources are thought to be especially involved in encoding, calculating, and response generating (Adams & Hitch, 1997; Campell & Xue, 2001; Imbo & Vandierendonck, 2008; Rasmussen & Bisanz, 2005). Specifically, the phonological loop is involved in maintaining the operands in multi-digit problems, the central executive is involved in carrying and borrowing procedures, and the visual-spatial sketchpad in visually representing solutions (see DeStefano & LeFevre, 2004 for a review).

Multi-digit problems with carrying or borrowing (e.g., 23 + 29 or 31 − 2) are assumed to be more difficult (as defined by increase in errors and response times) than problems with no carrying or borrowing (e.g., 23 + 22 or 34 − 2) (Ashcraft, 1992). Moreover, the former problems are assumed to place greater demands on working memory than the latter (Adams & Hitch, 1997). For example, solving multiplication multi-digit problems has been found to require operations that increase working memory load (Imbo, De Rammelaere, & Vandierendonck, 2005). To test the lack of fluency hypothesis due to working memory load in the present study, task demands were operationalized in terms of the number of elements to be manipulated in working memory for the production of response to simple arithmetic problems.

2.1.1 Hypotheses

Regarding *performance* it was predicted that there will be lower accuracy in performance and increased RT in problems that impose greater working memory load (as indexed by the number of elements to be manipulated, e.g., carrying and borrowing, or number of digits) (Hypothesis 1).

Regarding *FOD* it was predicted that greater working memory load will increase the self-reported feeling of difficulty (FOD ratings). Moreover, FOD ratings will be positively related to ratings of estimate of effort, but not with ratings of surprise (Hypothesis 2).

Regarding *estimate of effort* it was predicted that greater working memory load will increase estimate of effort ratings (Hypothesis 3).

Regarding *surprise* it was predicted that, because arithmetic problems are familiar tasks including no discrepant elements, surprise ratings will not be affected by working memory load (Hypothesis 4).

2.2 Method

2.2.1 Participants

Ten undergraduate psychology students (four men and six women, mean age = 20.18 years, *SD* = 0.16) of Aristotle University of Thessaloniki volunteered to participate in this experiment.

2.2.2 Apparatus

An Intel PC with standard keyboard and a 17-in SVGA monitor was used for task presentation and response registration. Programming of tasks was completed with E-prime (Schneider, Eschman, & Zuccolotto, 2002), which allowed recording of response accuracy (0 or 1) as well as of RT in milliseconds.

2.2.3 Design

A one-way within subjects design was used. Within subjects designs are generally preferred to between-subjects ones when measuring task difficulty effects, due to individual differences in motivation and prior knowledge (Whelan, 2007). The type of arithmetic operation was the independent variable (addition, subtraction, multiplication, and division). Dependent variables were (a) accuracy of response to the problem; (b) RT of problem processing, that is, before response production, and RT of response production; (c) FOD ratings (seven-point scale) before and after response production; (d) ratings of estimate of effort (seven-point scale) before and after response production; (e) ratings of surprise (seven-point scale) before and after response production. Subjective measures of FOD have been found to be reliable and sensitive to difficulty levels (Ayres, 2006; Efklides, 2002). Moreover, they are easy to implement and analyze.

2.2.4 Task

A computerized inductive reasoning task was used, consisting of six-number sequences. All sequences comprised numbers that conformed to a simple arithmetic rule. The basic assumption underlying this task is that mental computations with carrying and borrowing in multi-digit problems require working memory resources and increase operational difficulty. There were two blocks of number sequences involving the four arithmetic operations. The order of presentation of the sequences in each block was randomized. For all sequences, the common operand was the single-digit number 3. The first term of both addition and multiplication sequences comprised numbers 1 through 4. Similarly, subtraction and division sequences were constructed so that the last term of the sequence would be a number between 1 and 4. Overall, there were two addition sequences (2, 5, 8, 11, 14, 17, *20* and 1, 4, 7, 10, 13, 16, *19*), two subtraction sequences (21, 18, 15, 12, 9, 6, *3* and 22, 19, 16, 13, 10, 7, *4*), two multiplication sequences (2, 6, 18, 54, 162, 486, *1,458* and 4, 12, 36, 108, 324, 972, *2,916*), and two division sequences (729, 243, 81, 27, 9, 3, *1* and 1,458, 486, 162, 54, 18, 6, *2*). Among the number sequences, only the multiplication and division sequences required carrying and borrowing operations in working memory for the production of a response. Participants were asked to find out the rule underlying each sequence (e.g., always add three to previous number) and generate the next number that should conform to the rule underlying the sequence. That is, they had to mentally compute the answer and fill it in at the appropriate place. There was no time constraint for the production of a response.

Thus, according to our hypotheses, participants were expected to have lower performance in multiplication and division sequences as compared to addition and subtraction ones. No differences were expected to emerge between addition and subtraction sequences since the number of elements to be manipulated in working memory was the same in both addition and subtraction sequences.

Ratings of FOD, estimate of effort, and surprise were measured at two points in time; after initial processing of the number sequence, that is, before generating the required seventh number, and after solving the task, that is, after generating the seventh number (FOD ratings: "How difficult was it to discover the rule?" and

"How much difficulty did you feel when generating the seventh number?"; ratings of estimate of effort: "How much effort did you exert in order to discover the rule?" and "How much effort did you exert to generate the seventh number?"; and ratings of surprise: "How surprising was the sequence rule for you?"). A seven-point Likert-type scale was used for the ratings, anchored at one (not at all difficult/no effort at all/not at all surprising) and seven (extremely difficult/extremely much effort/extremely surprising).

2.2.5 Procedure

Participants carried out the experiment individually. The opening instructions on the PC monitor described the experiment as an examination of "one's ability to discover rules" and outlined the experimental procedure. Participants read the instructions at their own pace.

The experimental session began after two practice trials that comprised two sequences of numbers ascending by 1 (e.g., 1, 2, 3, 4, 5, 6, and 4, 5, 6, 7, 8, 9, respectively). Once it was clear that participants understood the instructions, they were asked to first figure out the rule quickly and accurately because later on, they would have to generate, also quickly and accurately, the seventh number.

The sequences were presented one at a time in the centre of the computer screen. They were presented horizontally in Courier New format as dark characters on a light background. The sequences remained on screen until the participants responded. Immediately after processing each number sequence participants were required to rate FOD, estimate of effort, and surprise. Next, the sequence appeared again on the screen and participants provided their response (i.e., generated the seventh number). After having provided their response, participants rated their FOD, surprise, and estimate of effort and proceeded to the next sequence.

2.3 Results

As already mentioned, because of the construction of the tasks, multiplication and division items were predicted to be the most demanding and highest in FOD ratings before and after solution, followed by addition and subtraction with similar levels of FOD. In Experiment 1, dependent variables were the accuracy of response to the number sequence, the RT of sequence processing, the RT of response production as well as the FOD, estimate of effort, and surprise ratings. Because the number sequences were presented in two blocks (first block and second block), there were two separate measures for each dependent variable.

First, paired samples *t*-tests were employed to compare the first block and second block responses across all variables for each number sequence item. There were no statistically significant differences (in all cases $p > 0.05$). For this reason, a composite variable was developed for each dependent variable, by computing the mean score of responses for each arithmetic operation in the two blocks. Next, the data were submitted to one-way within subjects ANOVA with Arithmetic Operation (addition, subtraction, multiplication, and division) as the within subjects factor.

Means (and standard errors) of performance, FOD ratings, surprise ratings, and estimate of effort ratings as a function of the type of arithmetic operation are presented in Table 9.1.

2.3.1 Performance

Accuracy of Response

The main effect of arithmetic operation was significant, $F(3, 51) = 9.769$, $p = 0.002$, partial $\eta^2 = 0.37$. Bonferroni post hoc tests, $p < 0.05$, showed that the accuracy of response was lower for multiplication as compared to addition, subtraction, and division. This finding is in partial agreement with Hypothesis 1, because division was expected to be as difficult as multiplication but it turned out not to be so.

RT of Sequence Processing

A main effect of arithmetic operation was found, $F(3, 51) = 10.601$, $p < 0.001$, partial $\eta^2 = 0.38$. Post hoc comparisons showed that RT was faster for addition and subtraction than for multiplication and division. This finding is in accordance with Hypothesis 1 that predicted longer RT for multiplication and division.

RT of Response Production

Again, the main effect of arithmetic operation was statistically significant, $F(3, 51) = 46.490$, $p < 0.001$, partial $\eta^2 = 0.73$. Post hoc comparisons showed that RT on multiplication was longer than that on addition, subtraction, and division. This finding is in partial agreement with Hypothesis 1, because RT for division was shorter than the one in multiplication.

Therefore, these findings partly confirmed Hypothesis 1, suggesting that the higher working memory load in multiplication resulted in lower performance. However, contrary to expectations, multiplication had lower accuracy than division, although working memory load during the processing of the number sequences was similar in both multiplication and division sequences. This finding is probably due to the fact that in the division sequences the load was placed on working memory during sequence processing and not during response production, as was the case for multiplication sequences.

Table 9.1 Means and standard errors of performance, FOD, surprise, and estimate of effort as a function of the type of arithmetic operation in Experiment 1

	Accuracy		RT of processing		RT of response		FOD (before)		FOD (after)		Surprise (before)		Surprise (after)		Effort (before)		Effort (after)	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Addition	0.83	0.07	3.93	0.05	3.72	0.06	1.19	0.11	1.22	0.10	1.14	0.08	1.22	0.09	1.31	0.12	1.22	0.10
Subtraction	0.69	0.08	3.90	0.04	3.74	0.05	1.19	0.10	1.19	0.08	1.17	0.08	1.28	0.10	1.31	0.09	1.19	0.08
Multiplication	0.31	0.09	4.19	0.06	4.40	0.06	1.86	0.18	2.58	0.33	1.36	0.13	1.61	0.17	1.97	0.20	2.58	0.33
Division	0.67	0.08	4.21	0.06	3.80	0.05	1.85	0.18	1.44	0.18	1.31	0.10	1.39	0.14	1.78	0.20	1.44	0.18

FOD, feeling of difficulty; *EOE*, estimate of effort; *RT,* reaction time; *Before*, before the response production; *After*, after the response production

2.3.2 Feeling of Difficulty

FOD Before Response Production

A significant main effect of arithmetic operation was observed, $F(3, 51)=8.907$, $p<0.001$, partial $\eta^2=0.34$. Post hoc comparisons showed that FOD was higher for multiplication and division than for addition and subtraction.

FOD After Response Production

The main effect of arithmetic operation was significant, $F(3, 51)=17.719$, $p<0.001$, partial $\eta^2=0.41$. Post hoc comparisons showed that multiplication had higher FOD than addition, subtraction, and division did.

Our findings supported Hypothesis 2, that is, the working memory load hypothesis of FOD since greater working memory load resulted in higher FOD ratings. Interestingly, FOD ratings before response production did not differ across multiplication and division sequences. However, they did differ after the response production, suggesting that FOD before solution and FOD after solution have access to different sources of information. Moreover, in both cases FOD ratings were consistent with the RT data and FOD after response production was consistent with accuracy data as well as RT.

2.3.3 Surprise

Surprise Before Response Production

No main effect of Arithmetic Operation was found, $F(3, 51)=1.898$, $p>0.05$, suggesting that surprise ratings were similar across the four arithmetic operations, as predicted by Hypothesis 3.

Surprise After Response Production

Again, there was no significant main effect of arithmetic operation, $F(3, 51)=2.408$, $p>0.05$. Since no discrepant events were detected in the simple arithmetic tasks, surprise was not affected by working memory load.

2.3.4 Estimate of Effort

Estimate of Effort Before Response Production

Effort ratings were significantly affected by arithmetic operation, $F(3, 51)=7.175$, $p<0.001$, partial $\eta^2=0.30$. Post hoc comparisons showed that estimate of effort was higher for multiplication than for addition and subtraction. However, they did not differ from division, providing further support for the working memory load hypothesis.

Estimate of Effort After Response Production

The main effect of arithmetic operation was significant, $F(3, 51) = 14.456$, $p < 0.001$, partial $\eta^2 = 0.46$. Post hoc comparison tests, $p < 0.05$, showed that multiplication had higher effort ratings than did the other types of arithmetic operations. These findings are in line with Hypothesis 4 and suggest that estimate of effort ratings follow the same pattern as FOD ratings.

2.3.5 Intercorrelations Between the Various Measures

Pearson correlations were employed to investigate the relationship between cognitive, metacognitive, and affective (surprise) variables.

A positive and marginally significant correlation was found between accuracy of response and RT during sequence processing ($r = 0.58$, $p < 0.05$), but not during response production and only in the multiplication sequence. There were no significant correlations between accuracy of response and FOD, estimate of effort, and surprise. The same was true regarding RT of sequence processing and FOD, estimate of effort, and surprise before and after solution. There was only a significant correlation between RT of response production and FOD after the solution for division ($r = 0.63$, $p < 0.01$) as well as with estimate of effort ($r = 0.68$, $p < 0.01$) but not with surprise.

Before the response production, FOD and estimate of effort were intercorrelated positively in all sequences (all r values ranging from 0.60 to 0.81, $p < 0.01$). Surprise also correlated positively with FOD in subtraction only ($r = 0.67$, $p < 0.01$). After the response production, FOD and estimate of effort were also intercorrelated positively and highly in all sequences (all r values ranging from 0.67 to 0.94). There was no significant correlation between FOD and surprise and only a significant correlation of estimate of effort with surprise in the subtraction sequence ($r = 0.72$, $p < 0.01$).

2.4 Discussion

The aim of Experiment 1 was to test the hypothesis that working memory load increases FOD (see also Ayres, 2006). Indeed, the greater demands on working memory placed by the multiplication sequences with carrying operations reduced performance (as indexed by accuracy of response and RT) and increased FOD.

Our findings corroborate those of Imbo and Vandierendonck (2008) showing that multiplication is more demanding than addition. Moreover, multiplication resulted in higher FOD and estimate of effort ratings than did the other three operations. Given that simple arithmetic operations are familiar tasks, differing only in processing demands, the working memory load hypothesis can account for the findings regarding multiplication as compared to addition and subtraction; however, they cannot fully explain the findings regarding division.

Moreover, FOD ratings were differentiated before and after the production of response. Because working memory load was similar during the initial processing of multiplication and division sequences, FOD ratings before solution were similar across the two types of operations. At the production of response, however, the load decreased in the division sequences because the seventh number to be produced was a single digit number. This was associated with lower FOD ratings after the response production. On the contrary, in the multiplication sequences, response production required a three- or four-digit number to be calculated with carrying, which posed high working memory load. This can explain the differences in the production of response phase. During the processing of the sequence, however, since both multiplication and division had both single and multi-digit numbers to be calculated as well as carrying and borrowing, no differences were found.

However, a word a caution is needed here because one could argue that the differences between addition and subtraction, on the one hand, and multiplication and division on the other could be due to the fact that the former operations involved only single or two-digit numbers whereas the latter involved from single to four-digit numbers. Therefore, it was not only the carrying and borrowing that made the difference but also the number of digits in the numbers of the sequence.

Of particular interest is also the close relationship between FOD and estimate of effort in every sequence encountered; what is difficult seems to be also effort demanding. This relationship has been repeatedly found in previous studies (Efklides, 2002; Efklides et al., 2006). Finally, in accordance with our hypothesis, FOD ratings did not correlate with surprise ratings, indicating that no unexpected element was detected within the sequences. However, a close relationship between FOD and surprise is expected to appear when the task one is working on is unfamiliar and some of its aspects are not well understood, that is, they are discrepant from our schemas.

3 Experiment 2: Second Test of the Working Memory Load Hypothesis

3.1 Aim and Hypotheses

In Experiment 1, multiplication and division sequences comprised numbers with more than two-digits, as opposed to addition and subtraction that consisted of only two-digit numbers. Arithmetic operations with multi-digit numbers, however, make more demands on working memory resources and increase the operational difficulty as compared to operations with single-digit numbers (DeStefano & LeFevre, 2004). Therefore one could assume that the results of Experiment 1 are due to the differences in the number of digits in the sequences and not only to the carrying and borrowing operations manipulated in working memory. To exclude this alternative explanation, Experiment 2 was carried out with addition and subtraction sequences comprising three-digit numbers.

The hypotheses were the same as in Experiment 1.

3.2 Method

3.2.1 Participants

Ten undergraduate psychology students (three men and seven women, mean age = 18.1 years, $SD = 0.32$) of Aristotle University of Thessaloniki participated in this experiment. None of them had participated in Experiment 1.

3.2.2 Apparatus

The same as in Experiment 1.

3.2.3 Design

The same as in Experiment 1.

3.2.4 Task

The same as in Experiment 1 except for the numerical sequences. For all sequences, the common operand was the single-digit number 3. There were two addition sequences (302, 305, 308, 311, 314, 317 and 901, 904, 907, 910, 913, 916), two subtraction sequences (421, 418, 415, 412, 409, 406 and 822, 819, 816, 813, 810, 807), two multiplication sequences (2, 6, 18, 54, 162, 486 and 4, 12, 36, 108, 324, 972), and two division sequences (729, 243, 81, 27, 9, 3 and 1,458, 486, 162, 54, 18, 6).

3.2.5 Procedure

The same as in Experiment 1.

3.3 Results

Means (and standard errors) of performance, FOD, surprise, and estimate of effort ratings as a function of the type of arithmetic operation are presented in Table 9.2.

3.3.1 Performance

Accuracy of Response

As in Experiment 1, a main effect of arithmetic operation was found, $F(3, 27)=9.243$, $p<0.001$, partial $\eta^2=0.51$. Accuracy of response was significantly lower for multiplication than for addition (Bonferroni post hoc tests, $p<0.05$).

RT of Sequence Processing

Again, the main effect of arithmetic operation was significant, $F(3, 27)=7.062$, $p=0.001$, partial $\eta^2=0.42$. Bonferroni post hoc tests, $p<0.05$, showed that RT was faster for addition and subtraction than for multiplication and division.

RT of Response Production

The main effect of arithmetic operation was also statistically significant, $F(3, 27)=21.297$, $p<0.001$, partial $\eta^2=0.70$, with RT being longer on multiplication than on addition, subtraction, and division (Bonferroni post hoc tests, $p<0.05$).

These findings replicate the findings of Experiment 1 and further confirm Hypothesis 1.

3.3.2 Feeling of Difficulty

FOD Before Response Production

A significant main effect of arithmetic operation was observed, $F(3, 27)=5.385$, $p=0.005$, partial $\eta^2=0.37$. FOD was higher for multiplication and division than for addition and subtraction, although post hoc comparisons were not significant.

FOD After Response Production

The main effect of arithmetic operation was significant, $F(3, 27)=18.464$, $p<0.001$, partial $\eta^2=0.67$. As in Experiment 1, multiplication had higher FOD ratings than the other arithmetic operations (Bonferroni post hoc texts, $p<0.05$).

Table 9.2 Means and standard errors of performance, FOD, surprise, and estimate of effort as a function of the type of arithmetic operation in Experiment 2

	Accuracy		RT of processing		RT of response		FOD (before)		FOD (after)		Surprise (before)		Surprise (after)		Effort (before)		Effort (after)	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Addition	0.90	0.21	3.92	0.13	3.91	0.18	1.25	0.35	1.35	0.34	1.20	0.35	1.25	0.35	1.20	0.26	1.35	0.42
Subtraction	0.85	0.33	4.01	0.17	4.02	0.22	1.35	0.52	1.30	0.67	1.25	0.49	1.30	0.67	1.50	0.53	1.35	0.67
Multiplication	0.50	0.33	4.17	0.27	4.49	0.19	2.30	1.15	3.55	1.30	1.35	0.53	1.50	0.47	2.45	1.26	3.55	1.71
Division	0.75	0.35	4.24	0.19	3.85	0.35	2.20	1.00	2.15	1.18	1.50	0.47	1.55	0.64	2.25	0.98	2.00	1.18

FOD, feeling of difficulty; *EOE*, estimate of effort; *RT*, reaction time; *Before*, before the response production; *After*, after the response production

The findings further support the working memory load hypothesis of FOD, suggesting that higher FOD results from greater working memory load. Thus, Hypothesis 2 was confirmed.

3.3.3 Surprise

No main effect of arithmetic operation was found on either surprise ratings before response production, $F(3, 27) = 1.312$, $p > 0.05$, or surprise ratings after response production, $F(3, 27) = 2.364$, $p > 0.05$. These findings show that surprise (before and after response production) was not affected by the type of arithmetic operation, and consequently, by working memory load. Thus, Hypothesis 4 was confirmed.

3.3.4 Estimate of Effort

Estimate of Effort Before Response Production

A main effect of arithmetic operation was observed, $F(3, 27) = 6.194$, $p = 0.002$, partial $\eta^2 = 0.41$. Specifically, estimate of effort was higher for multiplication and division than for addition and subtraction, although post hoc comparisons were not significant.

Estimate of Effort After Response Production

Once again, the main effect of arithmetic operation was significant, $F(3, 27) = 12.582$, $p < 0.001$, partial $\eta^2 = 0.58$. As in Experiment 1, multiplication had higher ratings of estimate of effort than did the other types of arithmetic operations (Bonferroni pairwise comparison tests, $p < 0.05$).

Thus, Hypothesis 3 was confirmed.

3.4 Discussion

Experiment 2 replicated the findings of Experiment 1, suggesting that FOD increases with working memory load. Since all sequences comprised three-digit numbers, the results of Experiment 1 may be better explained by the lack of fluency hypothesis due to working memory load imposed from arithmetic operations with carrying and borrowing.

4 Experiment 3: The Cognitive Interruption Hypothesis

4.1 Aims and Hypotheses

Experiment 3 was carried out aiming to examine whether events that are discrepant to a prevalent processing schema, and which presumably create cognitive interruption, will trigger FOD. Given that surprise is a response to schema-discrepant events (Meyer et al., 1997), this experiment also aimed to investigate the interrelations between FOD and surprise. That is, contrary to Experiments 1 and 2, in this case it was predicted that there should be a correlation between FOD and surprise.

Another aim of Experiment 3 was to examine the effect of the moment in which interruption is produced. Empirical findings related to fluency of processing have shown that repetition of a stimulus, which results in fewer interruptions, facilitates cognitive processing (Reber et al., 2002). Schema theories of revision processes also hold that RT is longer in the first occurrence of a schema-discrepant event as compared to subsequent occurrences of the same event (Mandler, 1975). This is mainly due to automatic revision processes following the detection of discrepancy. In the same vein, FOD has been found to correlate negatively with feeling of familiarity (Efklides, 2002). This means that events that have been encountered in the past (i.e., familiar) and therefore have been repeated, are processed more fluently and create lower FOD. Surprise has also been found to decrease after the first occurrence of unexpected events (Meyer et al., 1997). All these findings suggest that the first time a discrepant event is encountered will create higher FOD and surprise as compared to the subsequent times of occurrence. Therefore, the cognitive interruption effects should vary depending on the order of presentation (first, second, third) of the discrepant event.

The cognitive interruption hypothesis was tested with an inductive reasoning task consisting of number sequences which were used in order to induce processing schemas and discrepancies. To test the presentation order effects on our measures, the number sequences were presented three times, that is, in three blocks of presentation.

4.1.1 Hypotheses

Regarding *accuracy of response* it was expected that cognitive interruption will reduce the accuracy of response in discrepant number sequences as compared to non-discrepant. This is so because discrepancies represent obstacles in processing (Hypothesis 5).

Regarding the *RT of sequence processing and of response production* it was expected that discrepant number sequences as compared to non-discrepant, will require longer RT for their processing (Hypothesis 6a). This is so because discrepant number sequences require resources for schema restructuring to resolve the discrepancy and overcome the cognitive interruption. Discrepant events that are more difficult to integrate into a new schema will result in longer RT as compared to more easily integrated discrepant events (Hypothesis 6b).

Furthermore, according to schema theories (Mandler, 1975), RT will be longer in the first block of presentation of the number sequences as compared to the second or third blocks for both the discrepant and the non-discrepant events (Hypothesis 7a). However, the neural network models of learning hold that the revision processes that follow the detection of novel events are always slow (Rumelhart et al., 1986). Thus, according to this view, RT will decrease in the second and third block of presentation only for non-discrepant events for which no revision processes are needed (Hypothesis 7b).

Regarding *FOD* it was expected that cognitive interruption will trigger FOD (Hypothesis 8a). Ratings of FOD will be higher the more difficult it is to integrate events into a new schema (Hypothesis 8b). Furthermore, FOD and surprise will be related to each other due to their common source, namely cognitive interruption (Hypothesis 8c). Also, FOD ratings will be higher in the first block of presentation than in the following blocks of presentation (Hypothesis 8d).

Regarding *surprise* it was expected that cognitive interruption will trigger surprise (Hypothesis 9a), because it is by definition unexpected. No hypothesis was formulated as regards the effect of the difficulty of schema restructuring processes on surprise. If surprise depends on unexpectedness only (Meyer et al., 1997), no differences are expected to emerge between easy and difficult to integrate events (Hypothesis 9b). However, according to the Representational-Fit model of surprise (Maguire & Keane, 2006), surprise varies with the ease with which an unexpected event is integrated into the discourse representation. In such a case, surprise will be higher in the more difficult to integrate discrepant events (Hypothesis 9c). Finally, the intensity of surprise will be affected by the order of presentation, with surprise ratings being higher in the first block of presentation as compared to the second or third block of presentation (Hypothesis 9d).

4.2 Method

4.2.1 Participants

Ten undergraduate psychology students (four men and six women, mean age = 19 years) of Aristotle University of Thessaloniki volunteered to participate in this experiment.

4.2.2 Apparatus

The same apparatus was used as in Experiment 1.

4.2.3 Design

There was a 3×3 within subjects design. Independent variables were the type of cognitive interruption (No Interruption, Interruption due to repetition, and Interruption due to an intervening number) and the presentation order (first, second, and third block of number sequences). Dependent variables were (a) accuracy of response to the numerical sequence; (b) RT after the sequence processing, that is, before response production, and RT of response production; (c) FOD ratings as in Experiment 1, and (d) surprise ratings before and after response production as in Experiment 1.

4.2.4 Task

A computerized inductive reasoning task was used, consisting of six-number sequences that induced processing schemas and discrepancies. Specifically, numbers were used to induce expectations about the rule underlying the sequence. For example, if the first two numbers are ascending by 2, the expectation is that the rest of the sequence numbers will be ascending by 2 as well. Discrepant to this expectation, is a number ascending by 4; this discrepancy will result in cognitive interruption.

In all number sequences, we manipulated the content of the induced schema via the first two numbers of the sequence, for example, numbers ascending by 2. There were number sequences that conformed to the schema and sequences with unexpected numbers that deviated from this expectation. Thus, there were three types of cognitive interruption: (a) *no interruption*: number sequences comprising expected numbers only (e.g., 22, 24, 26, 28, 30, 32, *34*); (b) *interruption due to repetition*: there was an unexpected repetition of a number of the sequence in between the primary number sequence that conformed to the initially induced schema (e.g., 22, 24, *24*, 26, *26*, 28, *28*). That is, the discrepant number repeated the one that occurred after the application of the schema for the production of the ascending number. The absence of the expected ascending number due to the repetition of the previous number temporarily interrupts the completion of the sequence. However, since the number following the discrepant one conforms to the initial schema, the initially induced schema is easily resumed. Repetition of the number was thus temporarily interrupting the processing, but the interruption could be easily integrated to the initial processing schema. (c) *Interruption due to an intervening number*: A number of the sequence is followed by an unexpected number that is not related in any clear way to the numbers preceding and following the discrepant numbers (e.g., 22, *13*, 24, *13*, 26, *13*, *28*). The intervening number not only interrupts the course of the primary sequence (i.e., 22, 24, 26) as in repetition, but also challenges the formation of a schema, at least at the beginning. In order to be able to find the seventh number, one needs to form an overarching schema that involves two independent subschemas, that is, numbers ascending by 2, and place 13 in between the successive numbers of the sequence.

Thus, although the two last number sequences involve cognitive interruption and the formation of an overarching schema that integrates two subschemas, there is a difference between them. Specifically, the "repetition" sequence facilitates the formation of the primary schema (e.g., 22 24) before coming across the interruption, whereas the "intervening sequence" does not. Furthermore, whereas in "repetition" the discrepant number is easily integrated to the primary schema, thus facilitating the formation of the overarching schema, in the "intervening sequence", there is no initial schema valid that can form an expectation for what follows. This leads to extensive search for a schema that can account for the numbers of the sequence. Consequently, this sequence will create higher FOD than the repetition sequence and the no discrepancy sequence.

All number sequences across the three cognitive interruption types were matched with respect to operational difficulty required for generating the seventh number (see Appendix for a full list of the sequences used).

Prior to the main experiment a pilot study was conducted to determine the computational difficulty of the sequences. Ten undergraduate psychology students (five men and five women) of Aristotle University of Thessaloniki were required to solve a set of arithmetic problems as accurately and quickly as possible. There were 37 addition problems that examined the problem size effect (sum smaller than 10, e.g., 3+4 vs. sum larger than 10, e.g., 3+9), the carry effect (addition with carry, e.g., 39+4 vs. addition with no carry, e.g., 33+4), the addend size effect (single-digit addend, e.g., 1+4, two-digit addend, e.g., 11+4, and three-digit addend, e.g., 111+4) and the tens digit effect (the tens digit ranged from 1 to 9) on accuracy and RT. The common addend for all additions was number 4.

Results showed that large number problems, that is, sum larger than 10, took longer RT than did problems with sum smaller than 10, $t(9)=5.97$, $p<0.001$ ($M=3.34$, $SD=0.10$ and $M=3.13$, $SD=0.05$, respectively). Furthermore, problems with carrying required longer RT than did problems with no carrying, $t(9)=30.06$, $p<0.001$ ($M=3.82$, $SD=0.09$ and $M=3.36$, $SD=0.06$, respectively). Also, RT on single-digit addends ($M=3.24$, $SD=0.06$) was faster than that on two-digit addends ($M=3.43$, $SD=0.07$), which were faster than RT on three-digit addends ($M=4.30$, $SD=0.14$), $F(2, 18)=672.316$, $p<0.001$, partial $\eta^2=0.98$. Finally, the tens digit effect was not significant, $F(8, 72)=1.607$, $p=0.138$, indicating that the value of the tens digit did not affect RT. Comparable results were obtained by other studies on mental arithmetic (Ashcraft, 1992).

Therefore, the following factors were taken into account when constructing the sequences: (a) addition was selected as the arithmetic rule defining all sequences to avoid increases in difficulty due to computations; (b) calculations were kept small to eliminate the number-size effect; (c) carrying appeared once in each sequence. Moreover, because the tens digit had no effect on RT, various numerals were used to provide some variety to the sequences.

Participants were asked to find out the rule underlying each sequence. Specifically, the rule for *no interruption* was "Always add 2 to the previous number"; for *interruption due to repetition* was "Add 2 and always repeat the sum"; for *interruption* *due to an intervening number* was “Add 2 and then interpose the number 13 between the numbers”. Then, they had to generate the next number that should conform to the rule underlying the sequence. Ratings of FOD and surprise were taken at two points in time, as in Experiments 1 and 2, that is, before and after providing the response to the sequence.

4.2.5 Procedure

The same procedure was used as in Experiment 1, only that we controlled for order effects by employing randomization without replacement. The number sequences were presented randomly with the constraint that no two consecutive sequences had the same interruption type.

4.3 Results

Means (and standard errors) of performance, FOD and surprise ratings as a function of interruption type and presentation order are given in Tables 9.3 and 9.4.

4.3.1 Performance

Accuracy of Response

A repeated measures ANOVA showed that there was no main effect of cognitive interruption, $F(2, 18) = 2.186$, $p > 0.05$, or presentation order, $F(2, 18) = 1.00$, $p > 0.05$. Also, there was no statistically significant interaction between cognitive interruption type and presentation order, $F(4, 36) = 1.323$, $p > 0.05$. These findings suggest that both the interruption due to repetition and interruption due to an intervening number did not affect the accuracy of response, contrary to Hypothesis 5.

RT of Sequence Processing

A main effect of cognitive interruption type, $F(2, 18) = 8.746$, $p = 0.002$, partial $\eta^2 = 0.49$, was found. Bonferroni post hoc tests, $p < 0.05$, showed that interruption due to an intervening number and interruption due to repetition resulted in longer RT than no-interruption, suggesting that cognitive interruption increases processing time, as Hypothesis 6a predicted. However, no difference was found between the two discrepant sequences, contrary to Hypothesis 6b.

A main effect of presentation order was also observed, $F(2, 18) = 5.980$, $p = 0.01$, partial $\eta^2 = 0.40$. Results showed that responses to the first block of number

Table 9.3 Means and standard errors of performance, FOD, surprise, and estimate of effort as a function of the type of interruption in Experiment 3

	Accuracy		RT of processing		RT of response		FOD (before)		FOD (after)		Surprise (before)		Surprise (after)	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
No Interruption	0.96	0.03	3.66	0.05	3.47	0.03	1.23	0.11	1.16	0.11	1.30	0.11	1.26	0.08
Interruption due to repetition	0.73	0.13	3.89	0.06	3.54	0.05	1.73	0.28	1.53	0.21	1.76	0.31	1.60	0.26
Interruption due to an intervening number	0.93	0.04	3.90	0.04	3.59	0.03	1.93	0.28	1.63	0.20	1.96	0.31	1.70	0.22

FOD, feeling of difficulty; *EOE*, estimate of effort; *RT*, reaction time; *Before*, before the response production; *After*, after the response production

Table 9.4 Means and standard errors of performance, FOD, and surprise as a function of the presentation order in Experiment 3

	Accuracy		RT of processing		RT of response		FOD (before)		FOD (after)		Surprise (before)		Surprise (after)	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
First block	0.83	0.05	3.89	0.04	3.56	0.05	1.87	0.26	1.57	0.21	2.10	0.29	1.87	0.26
Second block	0.90	0.05	3.81	0.05	3.56	0.03	1.67	0.19	1.47	0.16	1.60	0.25	1.43	0.14
Third block	0.90	0.05	3.76	0.04	3.48	0.04	1.35	0.15	1.30	0.14	1.33	0.18	1.27	0.16

FOD, feeling of difficulty; *EOE*, estimate of effort; *RT*, reaction time; *Before*, before the response production; *After*, after the response production

sequences required longer RT compared to the third block (Bonferroni post hoc tests, $p < 0.05$). There was no significant interaction between cognitive interruption type and presentation order, $F(2, 18) < 1$, *ns*. These findings support Hypothesis 7a but not Hypothesis 7b.

RT of Response Production

There was a statistically significant main effect of Cognitive Interruption Type, $F(2, 18) = 3.982$, $p = 0.037$, partial $\eta^2 = 0.31$. Bonferroni post hoc tests, $p < 0.05$, indicated that interruption due to an intervening number resulted in longer RT than no-interruption. There was no main effect of presentation order, $F(2, 18) = 2.230$, $p > 0.05$. Also, no significant interaction of cognitive interruption type with presentation order was observed, $F(2, 18) = 1.636$, $p > 0.05$. These findings also support Hypothesis 7a but not Hypothesis 7b.

4.3.2 Feeling of Difficulty

FOD Before Response Production

A significant main effect of cognitive interruption type was observed, $F(2, 18) = 4.162$, $p = 0.033$, partial $\eta^2 = 0.32$. Interruption due to an intervening number and interruption due to repetition tended to produce higher FOD ratings than no-interruption did, although Bonferroni pairwise comparison tests were not significant, $p > 0.05$. The main effect of presentation order was also significant, $F(2, 18) = 4.981$, $p = 0.019$, partial $\eta^2 = 0.36$. Ratings of FOD ratings tended to be lower in the third block than in second or first block, although Bonferroni pairwise comparison tests were not significant, $p > 0.05$. Moreover, no significant interaction between cognitive interruption type and presentation order was observed, $F(2, 18) < 1$, *ns*. These findings are in partial agreement with Hypotheses 8a and 8b.

FOD After Response Production

The main effect of cognitive interruption type was statistically significant, $F(2, 18) = 5.111$, $p = 0.017$, partial $\eta^2 = 0.36$. Bonferroni pairwise comparison tests, $p < 0.05$, showed that interruption due to an intervening number produced higher FOD ratings than no-interruption did. There was no significant main effect of presentation order, $F(2, 18) = 2.739$, $p > 0.05$, indicating that FOD ratings were similar across the three blocks. Also, the interaction between cognitive interruption type and presentation order was not statistically significant, $F(2, 18) < 1$, *ns*. These findings support Hypothesis 8a but not Hypothesis 8b.

4.3.3 Surprise

Surprise Before Response Production

The reported surprise was significantly affected by Cognitive Interruption Type, $F(2, 18)=4.537$, $p=0.041$, partial $\eta^2=0.28$. Surprise ratings tended to be higher for interruption due to an intervening number or interruption due to repetition compared to no-interruption, although Bonferroni pairwise comparison tests were not significant, $p>0.05$. This finding is in line with Hypothesis 9a.

A main effect of presentation order was also observed, $F(2, 18)=8.956$, $p=0.002$, partial $\eta^2=0.50$. Surprise was found to be lower in the third block than in first block as predicted by Hypothesis 9d. There was no significant interaction between cognitive interruption type and presentation order, $F(2, 18)<1$, *ns*, contrary to Hypotheses 9b and 9c.

Surprise After Response Production

There was no significant main effect of cognitive interruption type, $F(2, 18)=2.191$, $p>0.05$, contrary to Hypothesis 9a. However, a main effect of presentation order $F(2, 18)=5.537$, $p=0.013$, partial $\eta^2=0.38$, was observed. Surprise ratings were lower in the third block than in first block, although Bonferroni pairwise comparisons were not statistically significant, $p>0.05$ as predicted by Hypothesis 9d. Finally, no cognitive interruption type by presentation order interaction was observed, $F(2, 18)<1$, *ns*, contrary to Hypotheses 9b and 9c.

4.3.4 Intercorrelations Between the Measures

Pearson correlations were employed to examine (a) the interrelations between cognitive, metacognitive and affective implications of cognitive interruption, (b) the degree to which FOD and surprise are related to each other, as it was assumed due to their common source, namely cognitive interruption, and (c) the extent to which these two variables are differentiated from each other.

Results showed that accuracy of response did not correlate with RT of sequence processing or response production. Furthermore, an examination of intercorrelations between RT and surprise ratings indicated that RT correlated positively with surprise ratings (r values ranging from 0.78 to 0.91, $p<0.01$). Marginally significant positive correlations between RT and FOD ratings (before and after) were also found ($r=0.64$ and $r=0.71$, $p<0.05$, respectively) in the first block of interruption due to an intervening number, only. Most importantly, FOD correlated positively and highly with surprise (r values ranging from 0.78 to 0.93, $p<0.01$) in all types of cognitive interruption, suggesting that surprise and FOD shared common variance.

4.4 Discussion

The results of Experiment 3 showed that cognitive interruption did not increase errors. Because of control processes assumed to come into play in unexpected or novel situations (Norman & Shallice, 1986), participants allocated more time to the processing of repeating and intervening numbers, and successfully integrated those numbers into their overall schema. Indeed, as compared to no-interruption, cognitive interruption due to repetition and due to an intervening number resulted in longer RT during both sequence processing and response production.

However, in the case of interruption due to an intervening number an alternative explanation would suggest that increases in RT may not have been due to cognitive interruption, but instead due to carrying operations performed during sequence processing. The unexpected and expected numbers in this type of interruption (i.e., 40 *12* 42 *12* 44 *12*) had different tens digits and perhaps participants computed the difference between successive numbers. However, interruption due to intervening number and interruption due to repetition (i.e., 40 42 *42* 44 *44* 46) did not differ as regards RT, although they differed in the tens digits of unexpected numbers. Further experiments are needed, however, to replicate these findings with intervening numbers being in the same tens digit as the rest of the sequence numbers.

Most importantly, cognitive interruption gave rise to both the metacognitive experience of FOD and the emotion of surprise. Also, positive and high correlations were observed between surprise and FOD ratings, providing further evidence that the two variables have a common source, namely cognitive interruption. However, FOD and surprise are not the same affective experiences since there was a different response pattern in the presentation order effect. Therefore, the role of cognitive interruption in the differentiation between the two affective states should be investigated in the future.

5 General Discussion

The main aim of the present study was to test two hypotheses about the nature of events that lead to the metacognitive experience of FOD. The findings of Experiments 1 and 2 supported the working memory load hypothesis, whereas Experiment 3 confirmed the cognitive interruption hypothesis. Moreover, Experiments 1 and 3 showed that depending on the source of difficulty, FOD may correlate with surprise or not.

5.1 RT and Cognitive Interruption

What is worth noting is that in all experiments the accuracy of response was not affected by working memory load or interruption but RT was, particularly during the processing of the arithmetical sequence as compared to response production. The more demanding the processing the higher the RT is. In the case of cognitive interruption the discrepant numbers took longer time for their processing than the non-discrepant numbers. RT delays due to unexpected events have been reported in other studies (Horstmann, 2005), suggesting that overcoming cognitive interruption is a time consuming process. It has been argued that the detection of discrepancy sets off revision processes so that the discrepant events can be recognized when they are encountered again in the future (Meyer et al., 1997).

This assumption is further supported by the findings of Experiment 3 which showed longer RT in the first block of presentation than in the second or third block of presentation. The findings were similar to those reported by Schützwohl (1998). This study showed that subsequent discrepant trials led to no RT increase, implying that discrepant events were no longer discrepant. Furthermore, the findings are in agreement with the schema theories (Mandler, 1975) but contradict the neural network models of learning (Rumelhart et al., 1986) that predict no faster RT in subsequent discrepant events. Mandler (1975) proposed that the more difficult to integrate the discrepant events the higher will be the intensity of the triggered affect. This could explain why the discrepant numbers resulted in higher FOD and surprise ratings than the non-discrepant ones. However, it cannot explain why there were no differences in FOD and surprise ratings between the two types of cognitive interruption. This finding was contrary to our hypotheses.

One possible explanation could be that the two types of interruption did not differ in terms of their demands on the revision processes triggered by the discrepant events. Although the "intervening" numbers were assumed to be more difficult to integrate, as compared to "repeating" numbers, participants may have assimilated both types of discrepant numbers in a relatively easy way. Specifically, as opposed to "repetition" (i.e., 40, 42, *42*, 44, *44*, 46), the "intervening" sequence (i.e., 40, *12*, 42, *12*, 44, *12*) comprised discrepant and non-discrepant numbers that differed in their tens digits. This may have facilitated the schema revision via the uniform connectedness which according to Gestalt and recent theories of visual organization (Han, Humphrey, & Chen, 1999; Quinlan & Wilton, 1998) is an important principle determining grouping processes. Thus, it could be that participants may have segregated early in visual perception the "intervening" numbers into numbers that belonged in one and the other tens digits. This could not occur in "repetition" sequence however, since the numbers had the same tens digits and participants could not group them by similarity. Therefore, despite their differences, both types of discrepant numbers were presumably easily integrated into processing schemas. This assumption is additionally supported by the high accuracy of response to both types of cognitive interruption. Nevertheless, further research is needed to test this assumption.

A second explanation could be that the intensity of affect depends on the degree of unexpectedness and not on the difficulty of schema revision. In this case, FOD and surprise ratings would not differ between the interruption types simply because the two types of discrepant numbers were equally unexpected. This fits with the psychoevolutionary model of surprise proposed by Meyer et al. (1997), according to which the degree of unexpectedness is the critical factor for the consequences of interruption of cognitive processing. However, Schützwohl (1998) found that RT to discrepant events tended to increase with initial schema strength and he attributed this effect to the increasing difficulty of schema revision. Moreover, it has been shown that when a plausible explanation, which facilitates the revision processes, is provided along with the unexpected event, surprise ratings tend to decrease (Maguire & Keane, 2006). This implies that the difficulty of the revision process is another factor of importance in determining the intensity of affective responses generated by unexpected events (Ortony, 1991). However, further experiments are needed for investigating why feeling of difficult and surprise ratings did not differ between the two types of cognitive interruption.

5.2 Relations of FOD with Surprise

Another interesting finding that emerged from Experiment 3 was the positive and high correlation between FOD and surprise ratings. This was in agreement with our hypotheses, which indicated that FOD shares common variance with surprise. A problem is difficult because it cannot be assimilated by our schemas, and at the same time, is surprising because it deviates from extant schemas. We are not implying, however, that surprise underlies FOD. Rather we are suggesting that these two affective reactions may coexist when two conditions are met: a task is new and a task cannot be automatically integrated into our extant schemas. It is also important not to over-interpret the relation between FOD and surprise demonstrated in this study. It is highly likely that FOD is systematically related to additional affective states, such as anxiety (Efklides, 2006).

The differentiation between FOD and surprise is further supported by the presentation-order effect that showed a different pattern of results for surprise and FOD ratings. Surprise ratings after solution decreased as the discrepancies were repeated, whereas FOD ratings did not. Surprise decrease after the first presentation of unexpected events has been reported elsewhere (Horstmann, 2002; Schützwohl, 1998), suggesting that revision processes follow the detection of discrepancy.

Furthermore, in agreement with our hypotheses, no correlation between FOD and surprise was observed in Experiments 1 and 2. A possible explanation for this could be that the arithmetic tasks were familiar and, thus, no discrepancies were detected. Rather than being related to surprise, FOD ratings were closely related to the self-reported effort exerted on the task. This finding has been reported previously (Efklides, 2002), suggesting that working memory load may be, among others, a condition for triggering both FOD and effort exertion.

5.3 FOD and Working Memory Load

The basic assumption of the working memory load hypothesis is that difficulty can be experienced in tasks that place a load on working memory (Borg et al., 1971; Efklides et al., 2006). For example, perceptions of difficulty have long been found to vary linearly as a function of the number of items to be attended in a visual attention task and as a function of the items to be recalled in a working memory task (Borg et al., 1971). However, our findings suggest that FOD, unlike the perceptions of difficulty, is connected to lack of processing fluency rather than to task features per se, or to performance considerations. This assumption is based on the findings of our experiments, namely that accuracy of response was high in all cases. If FOD ratings depended only on performance, no differences should have been found between the sequence processing phase and the response production phase. Moreover, there would be no differences between the multiplication and division sequences, because they both comprised 3- and 4-digit numbers. Lack of processing fluency can explain why FOD intensity has been found to change during problem solving. The dynamic nature of FOD has been stressed by Efklides (2002) who also reported variations in its intensity depending upon the progress towards the solution.

In conclusion, the findings of Experiments 1 and 2 suggest that if working memory load increases, FOD will increase as well. However, Experiment 3 showed that FOD is likely to increase even over low working memory load, as predicted by the cognitive interruption hypothesis, and more specifically, in less familiar tasks where discrepancies are detected. Given these findings, the cognitive interruption hypothesis of FOD deserves further scrutiny. In any case, the working memory load and the cognitive interruption hypothesis should not be considered contradictory; rather, they should be seen as complementary. When the task is familiar, FOD tends to increase with greater working memory load; but when the task is less familiar or involves discrepant events, FOD tends to increase with cognitive interruptions.

5.4 Limitations of the Study

The extent to which these data generalize to real-life problems in general should be considered in light of the limitations within these experiments. The first point to note relates to the nature of the tasks used. Participants were given a highly structured reasoning task with artificial discrepant and non-discrepant events. A second note of caution should be made in relation to the expertise of participants, since the study has used basic arithmetic skills that are well established in university students. The influence of background knowledge in participants' experiences in other kinds of problems is likely to be influential. Thus, further studies should be undertaken to test if the results are replicable in more complex and realistic problems.

6 Appendix. The Number Sequences Used in Experiment 3

No Interruption

22, 24, 26, 28, 30, 32, **34**; 31, 33, 35, 37, 39, 41, **43**; 55, 57, 59, 61, 63, 65, **67**

Interruption due to repetition

22, 24, *24*, 26, *26*, 28, **28**; 31, 33, 33, 35, 35, 37, **37**; 58, 60, 60, 62, 62, 64, **64**

Interruption due to an intervening number

22, *13*, 24, *13*, 26, *13*, **28**; 35, *12*, 37, *12*, 39, *12*, **41**; 57, *14*, 59, *14*, 61, *14*, **63**

References

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology, 67*(1), 21–38.

Alberdi, E., Sleeman, D. H., & Korpi, M. (2000). Accommodating surprise in taxonomic tasks: The role of expertise. *Cognitive Science, 24*(1), 53–91.

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136*(4), 569–576.

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*(1–2), 75–106.

Ayres, P. L. (2001). Systematic mathematical errors and cognitive load. *Contemporary Educational Psychology, 26*, 227–248.

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction, 16*(5), 389–400.

Barch, D. M., Braver, T. S., Nystrom, L. E., Forman, S. D., Noll, D. C., & Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia, 35*(10), 1373–1380.

Borg, G. (1978). Subjective aspects of physical and mental load. *Ergonomics, 21*(3), 215–220.

Borg, G., Bratfisch, O., & Dornič, S. (1971). On the problems of perceived difficulty. *Scandinavian Journal of Psychology,12*, 249–260.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624–652.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8*(12), 539–546.

Brinkmann, K., & Gendolla, G. H. E. (2007). Dysphoria and mobilization of mental effort: Effects on cardiovascular reactivity. *Motivation and Emotion, 31*, 71–82.

Bruner, J., & Postman, L. (1949). On the perception of incongruity: A paradigm. *Journal of Personality, 18*, 206–223.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4*(6), 215–222.

Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General, 130*(2), 299–315.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, New Series, 280*(5354), 748–749.

Carver, C. S. (2003). Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect. *Cognition and Emotion, 17*(2), 241–261.

Critchley, H. D., Mathias, C. J., Josephs, O., O'Doherty, J., Zanini, S., Dewar, B.-K., et al. (2003). Human cingulate cortex and autonomic control: Converging neuroimaging and clinical evidence. *Brain, 126*, 2139–2152.

Davidson, J. E., Deuser, R., & Sternberg, R. J. (1994). The role of metacognition in problem solving. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 207–226). Cambridge, MA: Bradford.

de Hoyos, M., Gray, E., & Simpson, A. (2004). Uncertainty during the early stages of problem solving. *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education, 2*, 255–262.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science, 5*(5), 303–305.

Delignières, D. (1998). Perceived difficulty and resources investment in motor tasks. *European Yearbook of Sport Psychology, 2*, 33–54.

D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature, 378*, 279–281.

DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*(3), 353–386.

Dewey, J. (1910). *How we think.* London: Health.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Review, 23*(10), 475–483.

Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation, and self-regulation. In A. Efklides, J. Kuhl, & R. M. Sorrentino (Eds.), *Trends and prospects in motivation research* (pp. 297–323). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2002). The systemic nature of metacognitive experiences: Feelings, judgments, and their interrelations. In M. Izaute, P. Chambres, & P.-J. Marescaux (Eds.), *Metacognition: Process, function, and use* (pp. 19–34). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2005). Emotional experiences during learning: Multiple, situated and dynamic. *Learning and Instruction, 15*, 377–380.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.

Efklides, A., & Tsiora, A. (2002). Metacognitive experiences, self-concept, and self-regulation. *Psychologia: An International Journal of Psychology in the Orient, 45*, 222–236.

Efklides, A., Kourkoulou, A., Mitsiou, F., & Ziliaskopoulou, D. (2006). Metacognitive knowledge of effort, personality factors, and mood state: Their relationships with effort-related metacognitive experiences. *Metacognition and Learning, 1*(1), 33–49.

Efklides, A., Papadaki, M., Papantoniou, G., & Kiosseoglou, G. (1997). Effects of cognitive ability and affect on school mathematics performance and feelings of difficulty. *American Journal of Psychology, 110*(2), 225–258.

Efklides, A., Papadaki, M., Papantoniou, G., & Kiosseoglou, G. (1998). Individual differences in feelings of difficulty: The case of school mathematics. *European Journal of Psychology of Education, 13*(2), 207–226.

Efklides, A., & Petkaki, C. (2005). Effects of mood on students' metacognitive experiences. *Learning and Instruction, 15*(5), 415–431.

Efklides, A., Samara, A., & Petropoulou, M. (1999). Feeling of difficulty: An aspect of monitoring that influences control. *European Journal of Psychology of Education, 14*(4), 461–476.

Eysenck, M. W., & Keane, M. (2005). *Cognitive psychology: A student's handbook* (5th ed.). New York: Taylor & Francis.

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Awareness and metacognition. *Consciousness and Cognition, 9*(2), 324–326.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new era of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911.

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Gendolla, G. H. E. (1999). Self-relevance of performance, task difficulty, and task engagement assessed as cardiovascular response. *Motivation and Emotion, 23*(1), 45–66.

Gendolla, G. H. E., & Krüsken, J. (2001). The joint impact of mood state and task difficulty on cardiovascular and electrodermal reactivity in active coping. *Psychophysiology, 38*, 548–556.

Han, S., Humphreys, G. W., & Chen, L. (1999). Uniform connectedness and classical gestalt principles of perceptual grouping. *Perception and Psychophysics, 61*(4), 661–674.

Herbert, A. (1978). Perceived causes of difficulty in work situations. *Ergonomics, 21*(7), 539–549.

Holyoak, K. J., Koh, K. K., & Nisbett, R. E. (1989). A theory of conditioning: Inductive learning within rule-based default hierarchies. *Psychological Review, 96*, 315–340.

Horstmann, G. (2002). Evidence for attentional capture by a surprising color singleton in visual search. *Psychological Science, 13*(6), 499–505.

Horstmann, G. (2005). Attentional capture by an unannounced color singleton depends on expectation discrepancy. *Journal of Experimental Psychology: Human Perception and Performance, 31*(5), 1039–1060.

Hrubes, D., & Feldman, R. R. (2001). Nonverbal displays as indicants of task difficulty. *Contemporary Educational Psychology, 26*, 267–276.

Imbo, I., De Rammelaere, S., & Vandierendonck, A. (2005). New insights in the role of working memory in carry and borrow operations. *Psychologica Belgica, 45*, 101–121.

Imbo, I., & Vandierendonck, A. (2008). Effects of problem size, operation, and working-memory span on simple-arithmetic strategies: Differences between children and adults? *Psychological Research, 72*, 331–346.

Izard, C. E. (1977). *Human emotions*. New York: Plenum.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*(2), 136–153.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289–325). Cambridge, UK: Cambridge University Press.

Koriat, A. (2008). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory and Cognition, 36*(2), 416–428.

Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 464–470.

MacLeod, C. M. (2004). The stroop task in cognitive research. In A. Wemzel & D. C. Rubin (Eds.), *Cognitive methods and their application to clinical research* (pp. 17–40). Washington, DC: American Psychological Association.

Maguire, R., & Keane, M. T. (2006). Surprise: Disconfirmed expectations or representation-fit? *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 1765–1770). Hillsdale, NJ: Erlbaum.

Mandler, G. (1975). *Mind and emotion*. New York: Wiley.

Mandler, G. (1984). *Mind and body: Psychology of emotion and stress*. New York: Norton.

Meyer, W., Reisenzein, R., & Schützwohl, A. (1997). Toward a process analysis of emotions: The case of surprise. *Motivation and Emotion, 21*(3), 251–274.

Moraitou, D., & Efklides, A. (2009). The Blank in the Mind Questionnaire (BIMQ). *European Journal of Psychological Assessment, 25*, 115–122.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102–116.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: Bradford.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (pp. 1–18). New York: Plenum.

Ortony, A. (1991). Value and emotion. In W. Kessen, A. Ortony, & F. Craik (Eds.), *Memories, thoughts, and emotions: Essays in honor of George Mandler* (pp. 337–353). Hillsdale, NJ: Erlbaum.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1–4.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63–71.
Pecchinenda, A., & Smith, C. A. (1996). The affective significance of skin conductance activity during a difficult problem-solving task. *Cognition and Emotion, 10*(5), 481–503.
Phaf, R. H., Mul, N. M., & Wolters, G. (1994). A connectionist view on dissociations. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance: XV. Conscious and nonconscious information processing* (pp. 725–751). Cambridge, MA: MIT Press.
Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self-regulated learning. In J. C. Impara (Series Ed.) & G. Schraw & J. C. Impara (Vol. Eds.), *Issues in the measurement of metacognition* (pp. 43–97). Lincoln, NE: University of Nebraska-Lincoln.
Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 211–227.
Posner, M. I., & DiGirolamo, G. J. (1998). Executive attention: Conflict, target detection, and cognitive control. In R. Parasuraman (Ed.), *The attentive brain* (pp. 401–423). Cambridge, England: Cambridge University Press.
Quinlan, P. T., & Wilton, R. N. (1998). Grouping by proximity or similarity? Competition between the gestalt principles in vision. *Perception, 27*(4), 417–430.
Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology, 91*(2), 137–157.
Reber, R., Fazendeiro, T. A., & Winkielman, P. (2002). Processing fluency as the source of experiences at the fringe of consciousness**.** *Psyche, 8*. Retrieved October 26, 2007, from http://psyche.cs.monash.edu.au/v8/psyche-8-10-reber.html
Reisenzein, R. (2000). Exploring the strength of association between the components of emotion syndromes: The case of surprise. *Cognition and Emotion, 14*(1), 1–38.
Rescorla, R. A. (2004). Spontaneous recovery. *Learning and Memory, 11*, 501–509.
Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2: Psychological and biological models* (pp. 7–57). Cambridge, MA: MIT Press.
Schneider, W., Eschman, A., & & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh, PA: Psychology Software Tools Inc.
Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience, 23*, 473–500.
Schützwohl, A. (1998). Surprise and schema strength. *Journal of Experimental Psychology: Learning Memory and Cognition, 24*(5), 1182–1199.
Shallice, T., & Burgess, P. W. (1996). The domain of supervisory processes and the temporal organisation of behaviour. *Philosophical Transactions of the Royal Society of London B, 351*, 1405–1412.
Stepper, S., & Strack, F. (1993). Proprioceptive determinants of emotional and nonemotional feelings. *Journal of Personality and Social Psychology, 64*(2), 211–220.
Sweller, J. (2003). Evolution of human cognitive architecture. *The Psychology of Learning and Motivation, 43*, 215–265.
Sweller, J. (2006). Discussion of 'Emerging topics in cognitive load research: Using learner and information characteristics in the design of powerful learning environments'. *Applied Cognitive Psychology, 20*, 353–357.
Teigen, K. H., & Keren, G. (2003). Surprises: Low probabilities or high contrasts? *Cognition, 87*(2), 55–71.
van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience, 14*(4), 593–602.
Whelan, R. R. (2007). Neuroimaging of cognitive load in instructional multimedia. *Educational Research Review, 2*, 1–12.
Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology, 81*(6), 989–1000.

Chapter 10
Tracking On-Line Metacognition: Monitoring and Regulating Comprehension in Reading

Riitta Kinnunen and Marja Vauras

1 Introduction

Readers often encounter difficulties with different levels of texts, which challenge their understanding. Think about the following simple sentence:

> Early in the morning, a man slowly walked towards his house post under his arm.

Most probably, this sentence activates an image of a man who has fetched his morning mail from the mailbox. However, consider that, at some later point, the text continues:

> He had to dig a deep hole to make sure that the post was steady enough to hold the heavy sign telling that the house was on sale.

On reading this last sentence, a skilled reader notices that the initial mental image was incorrect, and the reader has to adjust his interpretation. The reader may also notice that both a typical (in this case false) expectation, and incorrect use of the language, largely led him to make this error, as the word 'post' was written without an article in the first sentence. The reader became aware of the lexical ambiguity and the need to adjust the mental image of the situation to proceed smoothly with reading. Written texts often contain these kinds of difficulties, such as lexical ambiguities, grammatical errors, unknown words, ill-structured sentences, semantic confusion, violation of common knowledge, or ambiguous information. A reader, who habitually monitors and regulates his or her level of comprehension during the ongoing reading process, is expected to take corrective regulative actions whenever these comprehension failures are detected.

As part of metacognition, comprehension monitoring refers to metacognitive acts, which enable good comprehension (see, e.g., Baker, 2002; Brown, 1980; Ruffman, 1996; Wagoner, 1983). Although it may merely be a passing feeling, at

R. Kinnunen (✉)
Centre of Learning Research, University of Turku, FIN-20014 Turku, Finland
e-mail: riitta.kinnunen@utu.fi

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_10,

its best it involves conscious metacognitive awareness, that is, gaining on-line information regarding the level of comprehension, detecting comprehension obstacles, and deciding to maintain or increase the level of comprehension and to overcome identified comprehension failures (e.g., in reading: decisions to slow down reading, to re-read, to stop to think or to search from memory) (Kinnunen, Vauras, & Niemi, 1998). It is metacognitive in the sense that it takes attention away from the actual meaning construction and shifts it to the ongoing comprehension process itself. Readers who monitor their comprehension, thus set understanding as their goal, and constantly keep track of what they are reading to decide if it makes sense. When decisions are made at a conscious level, this also acts to compete for valuable working memory resources that are involved in comprehension per se (Baker, 2002; Kinnunen et al., 1998).

In the 1980s, researchers interested in cognitive development started to investigate the age at which the awareness of comprehension/incomprehension reaches such a level that a child is able to show recognition of comprehension difficulties, and is able to find strategies to overcome them. Generally, the results showed that the older and the more mature a person is, then the more probable it is that he shows signs of comprehension monitoring. Concerning the age/maturity effects, however, empirical evidence from the early studies, often applying the so-called error-detection method (see, e.g., Markman, 1979), was contradictory, showing that sometimes adolescents and adults failed in this area, and sometimes the very young succeeded in monitoring their comprehension. The reasons for this incongruity were often attributed by critics as to methodological factors (including the nature and demands of the tasks, the level of processing required, the prompted or spontaneous attention to comprehension obstacles), which, of course, partially explained the contradictory results. On the other hand, in regard to individual paths in language and, generally speaking, cognitive development, large individual variability is expected, which surpasses the age/maturity effects. Thus, the multiplicity of comprehension monitoring processes (Baker, 1985, 2002; Baker & Brown, 1984) and their bond to other processes involved in comprehension makes the age issue, as such, less obvious.

When a child starts to attend to his/her own mental processes – usually after acquisition of theory of mind at the age of about 4 years (Bartsch & Wellman, 1995) – a child also gradually becomes aware of his/her feelings of comprehension or incomprehension in different communication contexts. Language learning offers a mass of opportunities for the development of attention and awareness of one's own feelings, and, also, an increasing number of concepts to support and to enrich this awareness. Various oral communication situations (instructions, explanations, story-reading, etc.) offer opportunities both to make use of, and to develop more sophisticated conceptual tools. Although the findings generally indicate that the older and more skilled a person is the more likely he or she is to monitor comprehension, empirical evidence shows that even very young (4- to 5-year-old) children monitor their (listening) comprehension in communication contexts, showing sensitivity, for example, regarding story inconsistencies or ambiguous messages (Baker, 1984; Markman, 1979; Nilsen, Graham, Shannon, & Chambers, 2008; Peterson & Marrie, 1988; Revelle, Wellman, & Karabenick, 1985). Besides personal aptitudes and qualities, opportunities offered by the child's social environment (e.g., degree and nature of guidance in play, chores and other communication settings) are important in determining the growth of monitoring skills.

In kindergarten and in school, comprehension and learning become intentional goals and explicit demands for children, offering further opportunities and motivation to enhance awareness surrounding their own feelings. Monitoring skills earlier acquired in listening comprehension situations must now be put into practice in reading comprehension. First, however, a child has to acquire decoding skills, which delay this process. Reading from texts also changes the nature of comprehension processes, for the reason that texts are used to intentionally learn from them. To give a simple example beyond the necessity to decode written symbols into language, comprehension-serving cues often available in oral communication contexts (e.g., tone of the voice, stress on certain messages) are missing in the reading situation.

Comprehension monitoring is strongly related to skilled and effective reading and comprehension. Numerous studies show that the more skilled and mature a reader/comprehender is, the more probable it is that signs of comprehension monitoring can be found in his or her reading, whereas deficient comprehension monitoring seems to characterize poor readers/comprehenders (for a review, see Wagoner, 1983). Poor readers, irrespective of age, may monitor their comprehension less effectively than good comprehenders because of several reasons: their decoding skills may be immature and hampered, for example by limited working memory capacity, poor allocation of attention or language-related deficiencies (Cain, Oakhill, & Bryant, 2004; Walczyk et al., 2007); they may lack relevant conceptual knowledge (Vosniadou, Pearson, & Rogers, 1988); they may be impaired in their ability to construct coherent text representations that integrate provided information (Rubman & Waters, 2000; Zwaan & Radvansky, 1998); or they may lack the knowledge and skills of appropriate comprehension-debugging strategies (Wagoner, 1983).

Moreover, comprehension monitoring always reflects the reader's goals, his or her intended level of comprehension, and, in regard to these goals, the utilization of different standards as criteria for comprehension. Baker (for reviews, see Baker 1985, 2002) distinguished three basic types of standards, which operate at different levels of text processing: (a) *lexical* standards for monitoring the comprehension of words, (b) *syntactic* standards for monitoring syntactical acceptability, and (c) *semantic* standards for monitoring the construction of semantic representations (e.g., external consistency criteria for checking that acquired world knowledge is not violated, or internal consistency criteria for checking inconsistent or contradictory information). As an example, consider a young novice reader. As the goal of becoming a fluent reader has a priority over comprehension at the early stage of schooling, it is plausible to assume that the beginning reader's intended level of comprehension concerns more individual words and propositions than, for example, seeking logical consistency or construction of gist and meanings of the text. Therefore, if comprehension monitoring can be traced in their reading, it most probably is visible at the word or propositional level, with overreliance on lexical standards (cf., Baker, 1985;

Zabrucky & Ratner, 1989). Repertoire and use of multiple standards become gradually more visible in skilled and mature reading (see Baker, 2002).

Comprehension monitoring can be argued to be intimately associated with metacognitive experiences, despite the lack of clear conceptual analyses and empirical evidence generated up to now. Metacognitive experiences (e.g., feelings of familiarity or difficulty) are what the person experiences *on-line* during a cognitive endeavor, and they are affectively charged (Efklides, 2001, 2008). It must be kept in mind, though, that comprehension failure may pass unnoticed, and does not necessarily give rise to affect-laden metacognitive experiences. However, detection of comprehension failure may associate with an unexpected feeling of unfamiliarity, difficulty or uncertainty, which may instigate either negative affects like anxiety, nervousness, embarrassment or confusion, or positive stimulating affects like surprise, discovery, amusement or interest, and, consequently, inhibit or strengthen active monitoring and regulating of the ongoing reading process.

Let us briefly consider the two sentences of the example presented in the beginning of this chapter. Most probably, the second sentence triggers amused reactions like “Oops”, “Oh yes, I see now, I was led to think…” in a confident reader, and comprehension is easily repaired. However, if comprehension difficulty is noticed, very different emotional reactions like “I can’t understand a word”, “I’m really dumb, I completely lost it” may arise and paralyse further attempts to understand, or even halt reading altogether, when a young urban reader encounters text like the following, including too much unfamiliarity:

> Soon foxgloves would be nodding on roadside verges, and lamb’s foot would expose fiery heads from the hedgerows and the drystone walls that defined individual fields in this part of the world. (Elisabeth George, 2008, p. 1)

The present chapter focuses on on-line comprehension monitoring in reading, and is structured in three sections. In the two main sections, we first describe and discuss the on-line methods to study comprehension monitoring, and present how we have applied two of these methods, namely, traced silent reading and eye-tracking, in our studies to examine elementary (Grade 1–6) school students’ monitoring and regulating comprehension. Second, on the basis of these studies, we present evidence on young students’ comprehension monitoring and discuss developmental trends as a function of grade, decoding skills, listening and reading comprehension skills and intervention. As an example, we present, in more detail, results from a recent study linking students’ comprehension monitoring, mood and metacognitive experiences. Finally, we conclude by briefly discussing the promise and prospects of technology-supported on-line comprehension monitoring methods for metacognition research. Further, the future promise of assessing affects associated with comprehension monitoring processes are briefly discussed, particularly from the point of view of the possibilities of modern technology allowing synchronized data collection of affective reactions and reading comprehension behavior.

2 Traced Silent Reading and Eyetracking Methods

On-line methods are still rarely used to study monitoring and regulating of reading comprehension. The very nature of the phenomenon, though, calls for sensitive methods, allowing tracing processes that are not necessarily fully conscious acts but are either automatised, or not yet adequately developed. We have earlier (Kinnunen et al., 1998; Kinnunen & Vauras, 1995) pointed out that the methods to track comprehension monitoring should preferably meet, at least, most of the following requirements: (a) On-line reading should be registered in a situation, where the task is to read for comprehension, rather than to search for comprehension obstacles. (b) Evaluative and regulative acts of comprehension should be measured to sensitively reveal forms, intensity and developmental levels of monitoring. (c) Tasks should provoke monitoring acts at different levels of text processing (e.g., in respect of the degree of text integration required). (d) The method should offer the possibility of tracking the developmental changes or contextual effects both in the level of monitoring and in the level of text processing.

If the method fulfils these requirements, then we have a possibility to trace a reader's typical reading processes, even in cases where they are unconscious or semi-conscious, and to draw ecologically valid conclusions on a person's reading patterns without distractions from secondary tasks (e.g., thinking aloud) or other-imposed attention channeling processes (e.g., error detection). However, this does not exclude the value of a prompted search for comprehension obstacles, which can be effectively used for certain purposes. For example, error detection may serve pedagogical purposes and reveal developmental proximal regions in young or poor readers' comprehension skills, before active or nuanced comprehension monitoring can be traced in their reading.

One of the traditional methods to come close to the requirements listed above is tracing pauses, hesitations, repetitions and self-corrections during *oral* reading (Garner & Reis, 1981; Oakhill, Hartt, & Samols, 2005; Paris & Myers, 1981). However, reading aloud is not a typical activity in reading, and may interfere with comprehension monitoring. Although reading aloud may be more characteristic to very novice readers, even they tend to mumble instead of producing clear articulation, and instructions to spell out words could induce pronunciation problems, for instance, to young readers that may interfere with the comprehension processes more than during silent reading, and, thus, impede valid conclusions. Some studies have relied on observations during listening or silent reading, and all verbal remarks or reactions (like, "No!", "Mum, it is not…") and nonverbal signals (like frowns or puzzled looks) have been noted down and interpreted in relation to target messages. For example, Skarakis-Doyle (2002) developed a controlled on-line monitoring method, the expectancy violation detection task (EVDT), on the basis of child book-reading observations by the parent. The methods based on observation have proved useful with very young children, where comprehension monitoring is studied first of all in oral communication contexts (Nilsen et al., 2008; Skarakis-Doyle & Dempsey, 2008). In regard to reading comprehension, though, both oral reading and observations of silent reading, however controlled, give us fairly rough information on comprehension monitoring, and they are also rather uneconomical in terms of data registration and analysis.

Therefore, in our studies we sought methods that, as far as possible, fulfilled the criteria listed earlier and allowed computer-assisted registration and analysis of the data. Two non-intrusive on-line methods were applied in our studies, *traced silent reading* and *eyetracking*. Even though two decades ago applications of both of these methods were already in use to study comprehension monitoring (Baker & Anderson, 1982; Grabe, Antes, Thorson, & Kahn, 1987; Grabe, Antes, Kahn, & Kristjanson, 1991; Zabrucky & Ratner, 1986), it is surprising how rare these studies still are. Yet today, when technology is significantly advanced and on-line methods, eyetracking in particular, have been efficiently used in reading research in general, only a few studies (van der Schoot, Vasbinder, Horsley, Reijntjes, & van Lieshout, 2009) employ the possibilities that computer-assisted on-line methods offer for metacognition research. Next, we briefly describe these two methods as they have been applied in our comprehension monitoring research with young (from 7- to 13-year olds) elementary school students (Kinnunen & Vauras, 1995, 2009; Kinnunen et al., 1998; Kinnunen, Vauras, & Kajamies, 2009; Vauras, Kinnunen, & Rauhanummi, 1999; Vauras, Rauhanummi, Kinnunen, & Lepola, 1999). Two critical methodological features of all our studies have been (a) the goal for reading was to comprehend in order to answer subsequent questions and not, for example, to search for peculiarities or to evaluate the comprehensibility of the text; and (b) no secondary tasks (e.g., thinking aloud, or less disturbing reading aloud) were used (in an attempt to avoid possible extra shifts of attention or memory load increase in young novice readers).

2.1 *Traced Silent Reading*

In traced silent reading (Kinnunen et al., 1998, 2009), elementary school students' reading was tracked to examine how they spontaneously detected comprehension obstacles in sentences and short texts. The text was displayed on the computer screen, and the program registered data on time spent reading, look-backs and re-reading. The exact tracking of on-line reading was made possible during the self-paced reading by displaying sentences one word at a time and displaying passages sentence by sentence. The students moved forward or backward in the text by pressing the arrow keys on the keyboard (see in detail, Kinnunen et al., 1998). Prior to reading, the text was masked by Os, appearing as sequences of zeros on the screen; each zero represented one letter in the text. When reading the single sentences, pressing the forward or backward key made the next or the previous word visible (e.g., "0000 000 000 met 00 000 000000..."). In short texts, pressing the key made a whole sentence visible. When the next or the previous word/sentence became visible, the formerly visible word/sentence was again replaced by zeros. This method is also called a self-paced *moving window* method (Just, Carpenter, & Wooley, 1982).

Since we were interested in the sensitivity to comprehension obstacles and the development of meaning construction and comprehension monitoring skills of young students, the selected comprehension obstacles represented different text levels, such as lexical, propositional, and local (see van Dijk & Kintsch, 1983, pp. 10–19). Therefore, at the sentence level, lexical difficulty (a very rare word), syntactical error (an incorrect inflection) or a violation of common knowledge (external inconsistency) was included. In the short texts, a contradiction between two sentences was created. The reading of these sentences or texts was compared to those not including any intended comprehension obstacles. In the study by Kinnunen et al. (1998) four sentence paragraphs were given, with no resolution, whereas in the study by Kinnunen and Vauras (2009), six sentence passages with resolution were used (see the passage example given below).

Below is, first, an example of a sentence (target words in italics) in its four versions both in English and in Finnish (note that even though the word order is not correct in English it is correct in Finnish):

Once the man met *on the island* an angry dog.	[Sentence OK]
[Kerran mies tapasi *saaressa* kiukkuisen koiran.]	
Once the man met *in the briefcase* an angry dog.	[Common knowledge violation]
[Kerran mies tapasi *salkussa* kiukkuisen koiran.]	
Once the man met *into the island* an angry dog.	[Syntactical error]
[Kerran mies tapasi *saareksi* kiukkuisen koiran.]	[Incorrect (saare*ksi*) vs. correct (saare*ssa*) inflection]
Once the man met *in the silo* an angry dog.	[Rare word in Finnish]
[Kerran mies tapasi *siilossa* kiukkuisen koiran.]	

In the short texts, there was a contradiction between two sentences in the passage (e.g., "All frogs have to breathe to live" was later followed by "Some frogs do not have a nose or lungs", or "In the ancient plays, only the men were allowed to perform" was later followed by "In the play, the women got often into awful situations"). Some of the contradictions were not merely internal, but violated the common knowledge of the children, since most of them have no representations of, for example, breathing without nose and lungs, or sleeping in the air (see the example below). No artificial contradictions were used, and since we were also interested in poor and good comprehenders' corrective acts, the initially apparent contradiction was later resolved (Kinnunen et al., 1998). Below, an example of the whole passage in its two versions is given in English (in Finnish, the sentences were equally long).

Swifts often fly far up in the sky.	
Young swifts *sleep several times* a day.	
After the rest, the swifts seek insects for food.	
They *stay in the air* their first years of life.	[With contradiction]
They sleep supported by the currents of air.	[With resolution]
As adults they come down for nesting.	
[Whereas in no contradiction passages the last three sentences were:]	
They make their nests into all kinds of holes.	[With no contradiction]
They make the nest for nestlings with care.	
Adult swifts always use the same nest.	

During the on-line reading, the computer saved the key presses with their forward/backward attributes, as well as the times the words/sentences were visible. From these data, it was possible to reconstruct the readers' acts, examine their reactions to the comprehension obstacles and make inferences about their comprehension monitoring. Several variables were computed in order to trace the possible reactions: mean times spent reading a target word/sentence, mean number of re-readings of a whole sentence/passage and mean number of target-related or other look-backs. These variables were computed separately for each type of comprehension obstacle. A score for consistency of the monitoring, separately for the time spent reading the target words/sentences and the number of target-related look-backs, was also computed to clarify individual differences. This score was simply the number of detected obstacles in relation to all possible obstacles (see, in detail, Kinnunen et al., 1998). Thus, the data highly resembled those obtained by registering eye movements with eyetrackers.

2.2 *Eyetracking*

Eyetracking, registering eye fixations, offers a precise and powerful method to track on-line reading with any text material. Reading and regulatory strategies can be effectively revealed by eye-movement patterns during on-line reading (e.g., changes in speed, fixation paths, gaze durations, re-reading in relation to main ideas, key concepts and difficult words or other comprehension obstacles) (Hyönä, Lorch, & Rinck, 2003; Kinnunen & Vauras, 1995). Compared to the traced reading method described above, eyetracking has an advantage, by allowing the display of the texts in their natural form on the computer screen. This gives the reader a possibility to inspect and to read the text as they choose. Also, the use of pictorial material (like, graphs, graphics, tables, photos, etc.) in addition to text is a further advantage. Overall, it can be claimed that eyetracking provides presently the method which best fulfils the methodological requirements stated earlier for reading comprehension monitoring research.

The eyetracking method has been rather extensively used in reading and reading comprehension research (see, for a comprehensive review, Rayner, 1998), but not from the metacognitive comprehension monitoring perspective. The focus of eye-movement research has strongly been "expert reading", and thus, it has been conducted with adults as participants. Despite the promise of the method, the lack of eye-movement studies with children is understandable from the point of view of eye-movement registration techniques. Besides, experiments had to be conducted in laboratory settings. Previously, all systems demanded that head movements were restricted (e.g., the head fixed to a chinrest), and calibration processes took a long time. There were also limitations concerning the size of the material on the screen, which meant that only rather short texts could be used. This kind of system was used in our earlier eye-movement study (Kinnunen & Vauras, 1995), which resulted in considerable loss of valid data. In particular, the experimental situation was stressful for anxious and restless students with learning difficulties. Rapid technological development has provided more flexible devices to track eye-movements. For example, EYELINK consists of miniature cameras mounted on a headband, allowing precise binocular eyetracking (see, Kaakinen, Hyönä, & Keenan, 2003; van der Schoot et al., 2009). In our recent studies, we have used Tobii Eye Tracker, in which sensors to register eye movements are attached to the screen's panels, permitting free head movements. The system is highly suitable and adequate for young children, and allows rapid calibration and precise enough analyses for the comprehension monitoring research, where the interest is more on meaning construction at the propositional or higher text levels than, for example, within-word level.

In the study by Kinnunen and Vauras (1995) with 10-year-old elementary (Grade 4) school students, short texts (three to four sentences) were used. At that time, the eyetracker device (Remote Eye View Monitor) restricted the material that could be used and, therefore, passages were kept short, because the screen area on which eye-movement recordings succeeded best was limited. The texts contained lexical difficulties, or external or internal inconsistencies, which were manipulated in different versions of the same passage. Inconsistencies were not resolved in these texts. Below, an example of a version with internal inconsistency (here written in italics) is given in English:

> In the winter the squirrel sometimes wakes up and hurries to its hoard of food. There it has stored cones, nuts and acorns. Then it returns *hungry* to its nest and falls asleep again.

Two more versions were prepared by replacing the target word with another, in this case by "*nilfyg*" producing lexical problems, "*flying*" inducing an external inconsistency, or "*satisfied*" making the text consistent. In this study, the strategy training effects were examined, and the text series were prepared to allow balanced prior- and after-training measurement (see for details, Kinnunen & Vauras, 1995).

Modern eye trackers have opened up new possibilities to use longer and more authentic text materials. Therefore, in our recent studies (Kinnunen & Vauras, 2009; Vauras, Kinnunen, Salonen, & Lehtinen, 2008), we have been able to shift our interest to monitoring processes while students are reading complex expository texts, which resemble those read in school, which pose serious comprehension demands for all young students (see the study example given above). Although experimental control is weaker in these conditions, the texts and reading conditions are ecologically more valid, and pedagogically very interesting.

3 Comprehension Monitoring in Reading

In what follows, we present empirical evidence from our studies, where the above described on-line methods to study comprehension monitoring have been applied, to illuminate the promise and prospects of these methods in metacognition research with even the youngest school-aged students. To demonstrate the application of the traced silent reading method with young elementary school students, we start with the study relating to beginning readers' comprehension monitoring, its predictive power and interactions with decoding and listening comprehension. Then, we move to the developmental studies concerning comprehension monitoring in elementary school students (Grades 1–3, 4 and 6). To demonstrate the eyetracking method, we firstly discuss our early comprehension monitoring study with fourth graders, in which the impact of strategy training was also assessed with the aid of eye fixations in reading. Finally, we present as an example, in more detail, results from a recent study linking fourth graders' comprehension monitoring, mood, social anxiety and metacognitive experiences.

3.1 Comprehension Monitoring in Beginning Readers

By using traced silent reading, Kinnunen et al. (1998) found that children start to monitor their reading comprehension as soon as they learn to decode and read short meanings, only after the first few months of Grade 1. As such, this might not be surprising, since they already have a long history of comprehension attempts in oral communication settings – i.e., becoming aware of unfamiliar words or phrases and trying to figure out or ask about their meaning have been frequent events in their lives, as well as accepting the explanations given by adults. However, the level of their monitoring, and the fact that they were readily able to transfer monitoring skills into reading comprehension context was an interesting finding.[1] As regards sentence comprehension, their meaning construction and monitoring clearly exceeded the lexical level; that is, they did not only react to lexical difficulties (rare words), but also to syntactic errors, and more than 80% of the students reacted even to the sentences violating factual knowledge. Most often, they slowed down their reading, but more than 40% of them also reacted with target-related look-backs (looking back or coming back from the subsequent word). The proportion of detected obstacles (consistency) was 62% when measured as increased reading time, and 20% when measured as increased number of look-backs.

[1]It must be noted, though, that by the end of Grade 1 most of the Finnish children learn to read rather fluently; not only letters and words, but also longer text passages, and an increasing number of them acquire decoding skills prior to schooling at home or kindergarten. A concrete indication of the first graders' ability to rapidly and correctly decode words is that in a lexical decision task (syllable length of the words = 2; 36 words in series; $n = 226$), the mean decision time for words was 2.96 s ($SD = 1.44$ s) with the mean error rate 1.04 ($SD = 1.15$). In the word naming task, on average, it took less than 3 seconds for them to correctly name short (one- or two-syllable) words. Interestingly (and validating the differing methods), almost one-to-one correspondence in reading times of words in the naming task, and the target words of equal length in traced silent reading was observed. Further, most of the first graders were already able to read short stories at moderate speed and with remarkable accuracy (Kinnunen et al., 1998).

The level of decoding affected comprehension monitoring, creating problems in constructing meanings at a higher level than the lexical one, although the effects of stumbling decoding were not as severe as it could be expected (Kinnunen et al., 1998). As Stothard and Hulme (1996) had shown, readers may have advanced comprehension skills despite poor decoding. One-third of our first-grade students, those with the poorest decoding skills (that is, decoding one word took them several seconds, and they did not manage to read short texts), reacted to syntactic errors and common-knowledge violations by slowing down their reading. They seemed to have no resources left, though, for more active monitoring. This was evidenced by the lack of look-backs; that is, they made no active attempts to overcome comprehension obstacles. Interestingly, the detrimental effect of poor decoding seemed to be reduced by good listening comprehension skills, and increased the consistency of comprehension monitoring in poor decoders. The first signs of comprehension monitoring could also be detected at the more demanding (local) text level (four-sentence paragraphs), although only in a few first graders. Most of the beginning readers, despite good decoding and listening comprehension skills, did not yet monitor their comprehension at the local level, demanding integration and meaning construction of successive sentences (Kinnunen et al., 1998).

The traced silent reading method proved highly useful in uncovering even very young and novice readers' comprehension monitoring behavior. Beginning readers already had the ability and the goal to monitor their comprehension, and the better the decoding and listening comprehension skills were, the higher the level of meaning construction and monitoring (Kinnunen et al., 1998). Since this study was part of a longitudinal and intervention study, the importance of early emerging reading comprehension monitoring skills could later be demonstrated. In a follow-up, Vauras, Kinnunen, et al. (1999) showed that, along with metacognitive knowledge, comprehension monitoring skills at the first grade were strong predictors of later reading comprehension and mathematical word problem-solving proficiency at the end of the third grade. Typically, poor third grade learners' difficulties in word problems were accompanied by inferior early reading and metacognitive skills, as they lagged far behind the other students in both metacognitive knowledge and comprehension monitoring. In particular, the active regulative aspect of comprehension monitoring (i.e., in the form of look-backs indicating efforts to overcome comprehension obstacles) had not been present in their behaviour. Furthermore, results concerning the intervention effects confirmed the significant role of early developing comprehension monitoring skills (Vauras, Kinnunen, et al., 1999; Vauras, Rauhanummi, et al., 1999). Despite overall success of the reading comprehension and math problem-solving intervention (extending over 4 months), great variation in individual gains was observed, strongly associated with metacognitive factors. In contrast to responsive students, the students not benefiting from the training had shown lower metacognitive knowledge and comprehension monitoring earlier in the first grade. These results not only supported the general idea that self-regulation and the disposition to apply it are significant right from the beginning of schooling, but also lend support to the validity of the applied traced reading method to study comprehension monitoring.

3.2 *Development of Reading Comprehension Monitoring in Elementary Grades*

The traced silent reading method was also applied in two studies by Kinnunen et al. (2009), where we, first, followed the beginning readers from Grade 1 to Grade 2, and, second, compared comprehension monitoring skills in Grade 2, 4, and 6 students in a cross-sectional study. Comparisons between the first and second graders showed significant developmental changes, reflecting an increase in reading comprehension competence. The results indicate distinct qualitative changes in young children's comprehension monitoring during the second grade. The children had not only acquired higher reading speed, but their reading was marked with an increased number of obstacle-related look-backs, an increased obstacle-detection rate (both by reading times and look-backs) and spontaneous comments concerning the obstacles. This was particularly true for the individual sentences, although the general increase in look-backs and, further, obstacle-detection rates were also observed at the text level. Such a result indicates higher effort for comprehension on the part of the second graders, which probably reflects more active, regulatory, and consistent comprehension monitoring.

Only the short texts were given to students when comparing comprehension monitoring of second, fourth and sixth graders (Kinnunen et al., 2009). Although the evidence seen in the cross-sectional data is not as powerful as in follow-ups, trends on the development of comprehension monitoring were, at least, indicative. Reading speed did not significantly increase from Grade 2 onwards, showing that second graders were already rather fluent readers. The most significant results concerning the progress in comprehension monitoring were observed in obstacle-detection rates (by look-backs), time spent on sentences resolving the contradiction and spontaneous comments concerning the obstacles. It should be noted that the detection rate of comprehension obstacles assessed by reading times was equal in the three age groups. These results seem to point out that although comprehension monitoring strengthens during the early school years, the less active monitoring (i.e., reflected in time spent reading target sentences and detection rates by reading times) is already equally present in second graders and the development is not as rapid after the first 2 years in school as could be expected. The sixth graders' monitoring was more active in nature, also reflecting regulative aspects. Still, the sixth graders were far from what can be characterized as expert readers. For example, although spontaneous comments concerning comprehension obstacles increased from less than 2% (second graders) to 20% (sixth graders) and, respectively, detection rates by look-backs from less than 6 to 27%, this means that less than 30% of all internal inconsistencies in short texts were clearly detected by the sixth graders.

By using eyetracking, Kinnunen and Vauras (1995) studied the low- and high-achieving students' comprehension monitoring in relation to their text-processing skills at Grade 4. The study, being part of a longitudinal and intervention study (see, Vauras, Lehtinen, Kinnunen, & Salonen, 1992), allowed us also to examine low-achievers' comprehension monitoring from Grade 4 to Grade 5, that is, prior to, and about 4 months after a 16-week intervention on reading comprehension and/or socio-emotional coping. The results on group comparisons were in accordance with those found in our studies with younger students (Kinnunen et al., 1998, 2009), and those found in other on-line studies with older students (sixth graders; e.g., Zabrucky & Ratner, 1986). The high-achieving fourth graders were more advanced than the low-achieving ones in their text comprehension; the former being faster readers, indicating highly automatised decoding and easier access to meanings. They also showed active comprehension monitoring; that is, by slowing down reading and re-reading (look-backs) when confronted not only with lexical difficulties or common-knowledge violations, but also with internal inconsistencies. The low-achievers' slower reading was accompanied by fewer indications of comprehension monitoring, present primarily at the lexical and propositional (common knowledge violations) levels. In line with Kinnunen et al. (2009), rather minor developmental effects were observed from Grade 4 to Grade 5.

In the study by Kinnunen and Vauras (1995), we were interested in the relationship between comprehension monitoring and text integration, and, therefore, analysed the association between the eye movement data (three to four sentence texts) and text comprehension data concerning longer expository texts (approx. 180 words; e.g., on the life of Athenians in ancient Greece). Text integration and comprehension monitoring were closely associated, and particularly, the number of look-backs was significantly related to both the coherence and selection of main ideas in the summaries written after studying the longer texts, but only in the group of high achievers. The low-achievers' look-backs were rare and their summaries short and often incoherent, but this was not unexpected.

Within the low achievers, the transfer effects of comprehension strategy training were further examined. The trained students' longer reading times for internal inconsistencies indicated a slight increase in their tendency towards, and in proficiency of comprehension monitoring. However, active monitoring was still missing in their monitoring acts, since no reliable increase in look-backs was observed. Interestingly, though, the number of students reacting to internal inconsistencies increased as a function of training, and these students wrote more coherent summaries than the others after the training. The training helped the low-achieving students to raise their level of comprehension monitoring even without explicit training in this skill, and if this progress was achieved, it coincided with the rise in their ability to construct more coherent text representations integrating textual information.

3.3 Affect and Comprehension Monitoring in Reading Complex Texts

As a part of our larger longitudinal and intervention study (Vauras, 2006), we analyzed elementary (Grade 4) school students' eye movements during on-line reading, in an attempt to investigate students' comprehension monitoring in reading familiar and unfamiliar expository texts. In this context, we further explored the relationships of an affectively charged state of mind (mood) and experiences (metacognitive experiences) with comprehension monitoring and recall of main ideas. For the eye-movement study, the participants ($n=46$), including good ($n=17$) and poor ($n=19$) comprehenders, were selected on the basis of a reading comprehension task (a one-page, 220-word expository text on air pollution, including three types of comprehension tasks) given to all participants ($n=313$). The participants in the eye-movement study read two expository texts (familiar and unfamiliar, both approx. 160 words). The core content of the texts described the defense or the preying mechanism of a familiar animal – either a viper (biting its prey) or a lizard (losing its tail) – and of an unfamiliar animal – either a bombardier beetle (beetle defending itself by firing a boiling hot toxic fluid at predators) or a pistol shrimp (prawn with a sonic claw gun). The parts of these texts that were of specific interest were those detailing the process of the mechanism. Below, a part of the process description is given as an example (from the bombardier beetle):

> The cells of the beetle's back body produce two substances, which are carried through thin channels into spherical stores. When the beetle is frightened, these substances move through other channels towards the chambers at the tail end. These are like explosion compartments of rockets. The glands of the chambers produce a third substance, which brings about the explosion when it blends with those two other substances…

After reading, the participants assessed their own metacognitive experiences (feeling of familiarity, feeling of difficulty, certainty of comprehension and answering) and mood (e.g., happy, calm, worried or nervous) in relation to their feelings at the moment (Vauras, Efklides, Kinnunen, Salonen, & Junttila, 2007). We also used other emotion-related data on social anxiety (fear of negative estimations, and social anxiety and avoidance; see, La Greca & Lopez, 1998) and motivational vulnerability (ego-defensiveness; see Vauras, Salonen, Lehtinen, & Kinnunen, 2009) gathered prior to the eye-movement study. Our examples of the results are based on the following eye-movement variables depicting reading and comprehension monitoring:

(a) *The mean time* (per letter) *spent reading target lines*. The sum of first-pass fixation times (see, e.g., Kaakinen et al., 2003), including both the progressive and the regressive (re-reading) fixations (look-backs), was divided by the number of letters and marks in the line in order to correct the small differences in the line lengths.
(b) *The mean proportion of time spent re-reading phrases/propositions* (at least two successive words; e.g., "When the beetle is frightened" and in Finnish [Kun kiitäjäinen pelästyy]) *in the target lines*. The time spent on re-reading (the sum of first-pass regressive fixation times in the reread area) was divided by the sum of all the first-pass fixation times in the target lines.

The results indicated that both poor and good comprehenders monitored their reading by slowing down reading, re-reading difficult words, and looking back while constructing understanding of a complex process (Vauras et al., 2008). However, the monitoring patterns of the poor and good readers were significantly different, as indicated by fixation, re-reading and outcome patterns. First, the mean time spent reading target lines (process descriptions) differed significantly as a function of reading comprehension skill. Second, in the mean proportion of time spent re-reading the target lines, significant effects were observed as a function of text type and reading and comprehension skill. Opposite to expectations, fewer look-backs were made while reading the unfamiliar text. As expected, these results indicate that the poor readers were not such fluent decoders and read longer than the good readers but, instead, the good readers had more look-backs, that is, they re-read parts of the text more frequently. The good readers also clearly outperformed the poor readers in the recall of main ideas. It is interesting that no text effect (familiar vs. unfamiliar) on recall was found. Further, better recall was related to faster reading and more look-backs within both groups.

The lack of expected text effects may be a result of the fact that both texts included rather specific process descriptions, which in themselves were unfamiliar to the students, although they had general knowledge on vipers and lizards, but not on the two strange animals. This general familiarity may further have been instrumental for deeper understanding, that is, assimilation of new specific information with their existing knowledge. Thus, it seems that the young readers invested more effort and monitoring (as indicated by the number of look-backs) into the text, where assimilation of new information was more readily done than to the text, which demanded more (unassisted) bottom-up construction of knowledge.

Figure 10.1 (upper panel) illuminates the reading pattern of a skilled reader (the extract is the same as the one given earlier in English, i.e., from the bombardier beetle). The circled areas represent gaze durations (the larger the circle, the longer the fixation) and the lines and the numbers show gaze paths (regressive and progressive fixations). A rather systematic reading pattern can be seen in this extract. For example, look at the second line (note that only from this point, eye-movement patterns are visualized here). The student carefully reads the sentence “When the beetle is frightened, these substances move through other channels towards the chambers at the tail end.” We can see regressive and progressive fixations, particularly at the crucial points about frightening and about substances moving into the

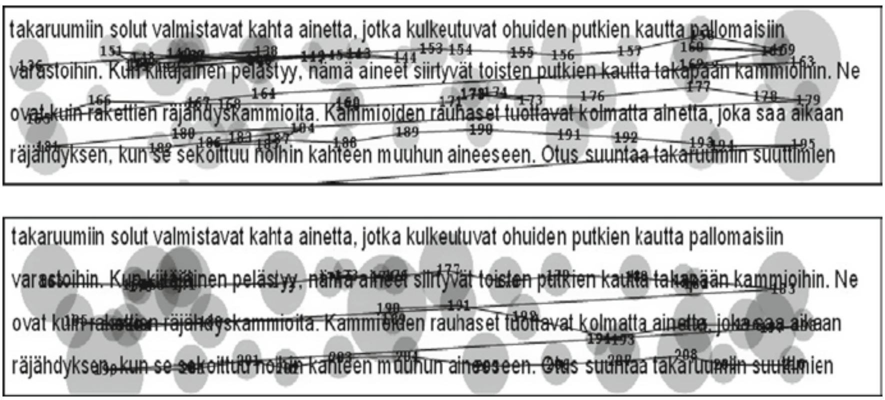


Fig. 10.1 Examples of good (*upper panel*) and poor (*lower panel*) readers’ eye-fixation patterns

chambers, with very long gaze duration on the last word (which in the Finnish version is “chambers”). He continues reading in a similar fashion throughout the whole critical region describing the complex process.

Rather different reading patterns can be seen in Fig. 10.1 (lower panel), which illuminates the eye-movements of a poor reader. The circles are larger, indicating long gaze durations, resulting in long total (first-pass) reading times per line. More interesting from the monitoring point of view is that a rather linear reading path, typical of poor readers, characterises their reading.

The study was carried out in the university laboratory, and sometimes it is claimed that this kind of the context creates negative affective effects on performance. In our study, we could compare the self-assessed mood in comparable reading situations in the classroom (pre-test task on air pollution) and the laboratory (eye-movement study). Therefore, it was highly interesting that the students reported being significantly in a more positive mood (e.g., happy, friendly, excited) and in a less negative mood (e.g., bad, worried, annoyed) in the laboratory than in the classroom context (Vauras et al., 2008).

Metacognitive experiences (ME) of feeling of familiarity, feeling of difficulty and certainty adequately varied as a function of text difficulty and unfamiliarity. After reading the unfamiliar text, both the poor and the skilled readers showed lower feeling of familiarity, higher feeling of difficulty, and lower levels of certainty (concerning both understanding and prospective answering of the questions) than after reading a familiar text. No significant correlations between ME and reading comprehension were observed, though. Social and emotional vulnerability (fear of negative estimations, social anxiety and avoidance and ego defensiveness) had a detrimental effect on comprehension monitoring processes and recall of main ideas, whereas a moderate degree of situational nervousness seemed to increase effort and positively affected comprehension monitoring and outcome. This situational arousal associated with favorable effects seemed to be more typical of the skilled readers, whereas generalised emotional vulnerability associated with unfavorable effects was more typical of the poor readers. However, it is plausible to assume that the affectively-charged experiences, states and interpretations have complex interdependent effects on reading behaviors, and bidirectional associations do not reveal this complexity. However, the number of participants in this study limited the possibilities to be able to analyse these complex associations.

4 Conclusion

4.1 The Promises of Technology-Supported On-Line Methods to Study Reading Comprehension Monitoring

Technologically-supported on-line methods, that is, traced silent reading and eye-tracking, have proved to be well applicable in studying comprehension monitoring in reading. Rather surprisingly, the advantages of these methods have not been fully exploited in the current research on comprehension monitoring, though. Along with a few other studies (e.g., van der Schoot et al., 2009), our studies show that these methods are subtle enough to give detailed information on students' reading processes that can be interpreted as mirroring both evaluative (evidenced, e.g., by prolonged reading times at information-demanding additional attention points) and regulative aspects (evidenced, e.g., by increased look-backs from, and to information creating and resolving problems) of comprehension monitoring. Naturally, there is always room for unfounded interpretations, since, e.g., long gaze durations may merely reflect the situation where a reader has halted reading and, for some reason or other, shifted attention to completely different thoughts. These effects are random, though, and can be avoided, for instance, if target information is clearly defined, intra- and inter-individual variation is recognized, and relative times are used (which also controls the reading speed effects). Most importantly, if consistent and confirming evidence is obtained from groups of readers and from several studies, more convincing conclusions become warranted. For example, our studies described here show rather systematic and converging support for young elementary school students' comprehension monitoring in reading.

Our studies have shown how beginning readers start monitoring their comprehension as soon as they sufficiently master decoding skills. In novice readers, though, the monitoring is more evaluative than regulative in nature, although in skilled young readers the signs of regulatory corrective acts of monitoring are already present. Slow decoding itself may not be a hindrance for monitoring to occur, and good listening comprehension skills seem to compensate for poor decoding (Kinnunen et al., 1998). After Grade 1, comprehension monitoring strengthens, along with increasing decoding proficiency, and more signs of regulation of comprehension become more visible (Kinnunen & Vauras, 2009; Vauras et al., 2008). However, the level of comprehension monitoring characterizing expert learners is not yet matured by the end of Grade 4 or even Grade 6, although great individual variability can be observed. These findings concerning fourth, fifth, and sixth graders' comprehension monitoring outcomes are in accordance with the other studies in which on-line methods are used (van der Schoot et al., 2009; Zabrucky & Ratner, 1986). In addition to consistent results on developmental trends, strong support for the validity of the on-line methods was offered by the results concerning the predictive power of early developed comprehension monitoring in reading (Vauras, Kinnunen, et al., 1999). Next, we briefly turn to the methodological issue of eyetracking and silent traced reading.

4.2 Methodological Issues

Eye movements have, from the beginning of psychological research, been considered to reflect cognitive processes, and as early as 1928, Tinker published an article on eye movements in reading with different materials and for different purposes (Tinker, 1928). As technology developed, eyetracking became ever more popular, particularly in basic psychological research. Rapid technological development during the past 10 years has increased the potential to use eyetracking in many areas and interests of psychological and educational research, as well as with participants of different ages. Whereas older eye trackers were in many respects (e.g., due to restrictions of materials and devices to hinder bodily and head movements) hard and inconvenient to use with young children or authentic materials, modern eye trackers have widened the potential use of this method and have also opened new opportunities for comprehension monitoring research. In reading, regressive fixations are particularly seen to reflect problems in comprehension, that is, regressions from, and to critical text information triggering or resolving the problems (cf., Hyönä et al., 2003). Since slowing down reading and regressive fixations may reflect various cognitive processes, like laborious decoding, the emphasis must be placed on reading materials used in the studies. With careful preparation of materials to be read, distinctions between comprehension monitoring and other reading processes, like decoding, can be reliably deduced.

Similarly with eyetracking, the traced silent reading method has proved to be powerful in allowing exact tracking of reading processes on-line without asking the student to read or think aloud, thus without attention shifting or other intrusive tasks being present. The inferences are also made from a very similar kind of data, that is, reading times and look-backs. Displaying words or sentences one-by-one may interfere with the results and conclusions, though. Rather recently, the two methods could be compared by means of two experiments by Rinck, Gamez, Diaz, and De Vega (2003) on processing of temporal text information. The same experiments were run, one using eyetracking and the other self-paced reading of sentences, showing similar results (cf., Just et al., 1982). However, to what extent Rinck et al.'s (2003) conclusions are valid for young students' comprehension monitoring in reading, remains open. Although such comparisons have not been made, our results from traced silent reading (Kinnunen et al., 1998, 2009) and eyetracking studies (Kinnunen & Vauras, 1995) yield similar developmental results and conclusions on young students' comprehension monitoring. Further, measures offered by traced silent reading have proved to be very strong and consistent predictors of later reading comprehension and mathematical word problem solving (Vauras, Kinnunen, et al., 1999). Our studies have further proved that traced reading is a convenient and adequate method to study reading comprehension even with beginning 7-year-old (Grade 1) readers (Kinnunen et al., 1998; Kinnunen & Vauras, 2009; Vauras, Kinnunen, et al., 1999; Vauras, Rauhanummi, et al., 1999). From an educational point of view, the computer-assisted traced silent reading method also provides a method that is applicable not only in research laboratories, but in schools, allowing a practical way to test comprehension monitoring.

In general, however, to make valid conclusions from eye movements or data from traced silent reading, the comprehensive nature of reading must be taken into consideration when preparing materials, tasks and designs for comprehension monitoring studies. Monitoring serves the level of comprehension that is going on, and the nature and level of comprehension monitoring is interweaved with many other acts and processes of the reader – i.e., the intentions, goals and/or demands of reading set different criteria for monitoring, and the strategies the reader masters determine their intent to make regulatory decisions. Our study (Kinnunen & Vauras, 2009) also lends support to the important role of the level of conceptual knowledge in comprehension monitoring. Adequate conceptual knowledge is needed to construct coherent text representations, which is required in order to monitor and regulate comprehension at textual local and macro-levels (cf., Rubman & Waters, 2000). This further means that with a good method, one gets information not only on the reader's monitoring, but also, e.g., about the semantic structures he is trying to make, or is capable of making meaningful and coherent.

We have argued that current technology-supported on-line methods offer interesting prospects for the research on comprehension monitoring. Not only is the use of more divergent and authentic tasks and materials made possible, but also the studies with very young children, like beginning readers without stressful, process inflicting side effects. These methods further allow well controlled and replicable studies, which enhances their use in longitudinal developmental research as well as devices to scrutinize effectiveness of interventions. Besides cognitive factors, other attention-captivating processes, emotional and motivational, affect comprehension monitoring. Regrettably, apart from our own unpublished data (Vauras et al., 2008), studies combining synchronized observational data about emotional reactions with eye movements in comprehension monitoring are non-existent. Yet today, the synchronized on-line data collection and analysis with the aid of modern technology (like eye-movement registration in synchrony with audio- and video-analysis of speech and facial expressions) is made possible in some modern eye trackers (such as Tobii T60 XL with Tobii Studio Software). The two different sources enable on-line information of the students' affective reactions, and make it possible to relate them exactly and dynamically to the content being read and the comprehension strategies being used.

Acknowledgements The research was supported by Grant 114048 from the Council of Cultural and Social Science Research, The Academy of Finland, to the first author.

References

Baker, L. (1984). Spontaneous versus instructed use of multiple standards for evaluating comprehension: Effects of age, reading proficiency, and type of standard. *Journal of Experimental Child Psychology*, *38*, 289–311.

Baker, L. (1985). How do we know when we don't understand? Standards for evaluating comprehension. In D. Forrest-Pressley, G. McKinnon, & T. Waller (Eds.), *Metacognition, cognition and human performance* (pp. 155–205). New York, NY: Academic.

Baker, L. (2002). Metacognition in comprehension instruction. In C. Collins Block & M. Pressley (Eds.), *Comprehension instruction. Research-based best practices* (pp. 77–95). New York, NY: Guilford.

Baker, L., & Anderson, R. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. *Reading Research Quarterly*, *17*, 281–293.

Baker, L., & Brown, A. (1984). Metacognitive skills and reading. In P. J. Pearson, M. Kamil, M. R. Barr, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353–394). New York, NY: Longman.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. New York, NY: Oxford University Press.

Brown, A. L. (1980). Metacognitive development and reading. In R. J. Spiro, B. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 453–481). Hillsdale, NJ: Erlbaum.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31–42.

Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation and self-regulation. In A. Efklides, J. Kuhl, & R. Sorrentino (Eds.), *Trends and prospects in motivation research* (pp. 297–323). Dordrecht, the Netherlands: Kluwer.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, *13*, 277–287.

Garner, R., & Reis, R. (1981). Monitoring and resolving comprehension obstacles: An investigation of spontaneous text lookbacks among upper-grade good and poor comprehenders. *Reading Research Quarterly*, *16*, 569–582.

George, E. (2008). *Careless in red.* London: Hodder & Stoughton.

Grabe, M., Antes, J., Thorson, I., & Kahn, H. (1987). Eye fixation patterns during informed and uninformed comprehension monitoring. *Journal of Reading Behavior*, *19*, 123–140.

Grabe, M., Antes, J., Kahn, H., & Kristjanson, A. (1991). Adult and adolescent readers' comprehension monitoring performance: An investigation of monitoring accuracy and related eye movements. *Contemporary Educational Psychology*, *16*, 45–60.

Hyönä, J., Lorch, R., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Amsterdam: Elsevier.

Just, M., Carpenter, P., & Wooley, J. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*, 228–238.

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2003). How prior knowledge, WMC, and relevance of information affect eye fixations in expository text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 447–457.

Kinnunen, R., & Vauras, M. (1995). Comprehension monitoring and the level of comprehension in high- and low-achieving primary school children's reading. *Learning & Instruction*, *5*, 143–165.

Kinnunen, R., & Vauras, M. (2009). *Comprehension monitoring in good and poor comprehenders: An eye fixation study with elementary school students*. Unpublished manuscript.

Kinnunen, R., Vauras, M., & Niemi, P. (1998). Comprehension monitoring in beginning readers. *Scientific Studies in Reading*, *2*, 353–375.

Kinnunen, R., Vauras, M., & Kajamies, A. (2009). *Development of comprehension monitoring in elementary school students: Evidence from traced on-line reading*. Unpublished manuscript.

La Greca, A. M., & Lopez, N. (1998). Social anxiety among adolescents: Linkages with peer relations and friendships. *Journal of Abnormal Child Psychology*, *26*, 83–94.

Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development*, *50*, 643–655.

Nilsen, E. S., Graham, S. A., Smith, S., & Chambers, C. (2008). Preschoolers' sensitivity to referential ambiguity: Evidence for a dissociation between implicit understanding and explicit behavior. *Developmental Science*, *11*, 556–562.

Oakhill, J., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing: An Interdisciplinary Journal*, *18*, 657–686.

Paris, S. G., & Myers, M. (1981). Comprehension monitoring, memory, and study strategies of good and poor readers. *Journal of Reading Behavior*, *13*, 5–22.

Peterson, C., & Marrie, C. (1988). Even 4-year-olds can detect inconsistency. *The Journal of Genetic Psychology*, *149*, 119–126.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Revelle, G., Wellman, H., & Karabenick, J. (1985). Comprehension monitoring in preschool children. *Child Development*, *56*, 654–663.

Rinck, M., Gámez, E., Díaz, J., & De Vega, M. (2003). Processing of temporal information: Evidence from eye movements. *Memory & Cognition*, *31*, 77–86.

Rubman, C., & Waters, H. (2000). A, B seeing: The role of constructive processes in children's comprehension monitoring. *Journal of Educational Psychology*, *92*, 503–514.

Ruffman, T. (1996). Reassessing children's comprehension monitoring skills. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 33–67). Mahwah, NJ: Erlbaum.

Skarakis-Doyle, E. (2002). Young children's detection of violations in familiar stories and emerging comprehension monitoring. *Discourse Processes*, *33*, 175–197.

Skarakis-Doyle, E., & Dempsey, L. (2008). The detection and monitoring of comprehension errors by preschool children with and without impairment. *Journal of Speech, Language and Hearing Research*, *51*, 1227–1243.

Stothard, S., & Hulme, C. (1996). A comparison of reading comprehension and decoding difficulties in children. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 93–112). Mahwah, NJ: Erlbaum.

Tinker, M. (1928). Eye movement duration, pause duration, and reading time. *Psychological Review*, *35*, 385–397.

van der Schoot, M., Vasbinder, A., Horsley, T, Reijntjes, A., & van Lieshout, E. (2009). Lexical ambiguity resolution good and poor comprehenders: An eye fixation and self-paced reading study in primary school children. *Journal of Educational Psychology*, *101*, 21–36.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies in discourse comprehension.* London: Academic.

Vauras, M., Lehtinen, E., Kinnunen, R., & Salonen, P. (1992). Socioemotional and cognitive processes in training learning disabled children. In B. Wong (Ed.), *Contemporary intervention research in learning disabilities: An international perspective* (pp. 163–189). New York, NY: Springer.

Vauras, M., Kinnunen, R., & Rauhanummi, T. (1999). The role of metacognition in the context of integrated strategy intervention. *European Journal of Psychology of Education*, *14*, 555–569.

Vauras, M., Rauhanummi, T., Kinnunen, R., & Lepola, J. (1999). Motivational vulnerability as a challenge for educational interventions. *International Journal of Educational Research*, *31*, 515–531.

Vauras, M. (2006). SOLE. The social in learning: Upbringing socially and academically competent, motivated, and self- and co-regulated young learners in school and family contexts. Unpublished research plan. Academy of Finland.

Vauras, M., Efklides, A., Kinnunen, R., Salonen, P., & Junttila, N. (2007). *Mood and metacognitive experiences scale for children*. Unpublished assessment scale.

Vauras, M., Kinnunen, R., Salonen, P., & Lehtinen, E. (2008). Emotions in comprehension monitoring and regulation. *International Journal of Psychology*, *43*, 17.

Vauras, M., Salonen, P., Lehtinen, E., & Kinnunen, R. (2009). Motivation in school from contextual and longitudinal perspective. In M. Wosnitza, S. A. Karabenick, A. Efklides, & P. Nenniger (Eds.), *Contemporary motivation research: From global to local perspectives* (pp. 1–23). Cambridge, MA: Hogrefe-Huber.

Vosniadou, S., Pearson, P., & Rogers, T. (1988). What causes children's failures to detect inconsistencies in text? Representation versus comparison difficulties. *Journal of Educational Psychology*, *80*, 27–39.

Wagoner, S. (1983). Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly*, *18*, 329–346.

Walczyk, J. J., Wei, M., Zha, P., Griffith-Ross, D. A., Goubert, S. E., & Cooper, A. L. (2007). Development of the interplay between automatic processes and cognitive resources in reading. *Journal of Educational Psychology*, *99*, 867–887.

Zabrucky, K., & Ratner, H. (1986). Children's comprehension monitoring and recall of inconsistent stories. *Child Development*, *57*, 1401–1418.

Zabrucky, K., & Ratner, H. (1989). Effects of reading ability on children's comprehension evaluation and regulation. *Journal of Reading Behavior*, *21*, 69–83.

Zwaan, R., & Radvansky, G. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185.

Part II
Developmental and Educational Implications of Metacognition

Chapter 11
Metacognition in Young Children: Current Methodological and Theoretical Developments

David Whitebread, Qais Almeqdad, Donna Bryce, Demetra Demetriou, Valeska Grau, and Claire Sangster

1 Introduction

The established orthodoxy within metacognition research has been that metacognitive skills emerge around the age of 8–10 years (Veenman, Van Hout-Wolters, & Afflerbach, 2006) and are necessarily preceeded by other cognitive abilities such as the development of theory of mind (Wellman, 1985). However, this position has been challenged by recent research on both methodological and theoretical grounds. As regards methodology, it is increasingly recognised that research relying on self-report or verbally-based experimental methodologies may significantly underestimate the metacognitive and self-regulated performance of young children (Van Hout Wolters, 2000; Whitebread et al., 2005; Winne & Perry, 2000). Recent studies, adopting a range of more age-appropriate methodologies, have identified and begun to analyse metacognitive and self-regulatory behaviours in much younger children. The development of these new methodological approaches, involving various kinds of systematic observation, is reviewed in the following section of this chapter.

These methodological advances have facilitated the development of new understandings concerning the emergence of metacognition in young children and the general structure and inter-relations between metacognitive processes. These various theoretical strands, concerned with conscious and nonconscious processes, social aspects of regulation, and interactions between affective and metacognitive aspects of self-regulation have been recently and persuasively drawn together by Efklides (2008) within a new multifaceted and multilevel model of metacognition. Further discussion of these theoretical advances is presented within the third section of the present chapter.

D. Whitebread (✉)
University of Cambridge, 184, Hills Rd, Cambridge, UK
e-mail: dgw1004@cam.ac.uk

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_11,

The final section of the chapter reviews current research projects, carried out by members of the University of Cambridge Faculty of Education Self-Regulated Learning Research Group, involving children aged 3–11 years, in the UK, Cyprus, Chile, Jordan and Canada. The methodological and theoretical developments referred to above are exemplified by these projects. They have predominantly used observational methods, interventions and tasks which reduce the demands on young children's verbal abilities, and demonstrated that clear indicators of early metacognition can be discerned by these methods within typically developing young children, and, indeed, young children with learning and motor difficulties. Within the typically developing group, studies are reported which have explored the relation between early metacognition and inhibition, theory of mind and conceptual development. These studies support a model of metacognition which is multifaceted, within which different facets exhibit differential developmental trajectories, and within which particular facets, namely those defined by Efklides (2008) as metacognitive experiences and metacognitive skills, are fundamental to the functioning of the human brain and are, in emergent forms, discernable within the behaviour of very young children. The present study concerned with conceptual development and the intervention studies with children with learning and motor difficulties have also recognised the impact of motivational and contextual factors (for example, "mastery-orientation", task characteristics, nature and quality of adult mediation) on young children's metacognitive learning and performance. The intervention studies have also demonstrated that young children, even those with learning difficulties, may have weak metacognitive skills, but can benefit from certain kinds of metacognitive training in ways which enhance their performance.

The chapter is organised into three sections, dealing in more detail with, first, methodological developments in studies with young children; second, the associated theoretical developments; and, third, the findings and issues raised by the Cambridge studies. What emerges is a new field of investigation within the metacognition literature which has the potential to make important contributions to understanding the complexities of early development, and the nature of metacognitive processes themselves.

2 Methodologies for Identification, Assessment, and Measurement of Metacognition in Young Children

Within much of the metacognition literature, there has been an emphasis on using self-report techniques as a way of understanding individuals' metacognitive processes (Winne & Perry, 2000). Self-report interview or questionnaire methodologies are intimately linked to a theoretical model of metacognition as an essentially explicit and declarative set of processes, and depend upon the respondents' ability to give reliable reports of their own mental experiences. As we discuss in the following section, however, there is now evidence to support the view that some aspects of monitoring and control processes are not available to conscious awareness (Efklides, 2008; Fitzsimmons & Bargh, 2004; Reder, 1996; Siegler, 1996). Indeed, even in classic early work exploring metacognitive processes in young children, there was clear evidence showing them to be often capable of performing tasks (such as Piaget's oddity problem) and adapting their behaviour effectively (e.g., rehearsing when asked to remember some items), but being unable to report verbally on what they had done (Flavell, Beach, & Chinsky, 1966; Piaget, 1977).

The recognition of the role of conscious and implicit processes in metacognition has major implications for the identification, assessment and measurement of metacognitive processes, particularly in young children with limited verbal abilities. Veenman (2005), in an exhaustive review of methodologies used in metacognition research, for example, has argued that multi-method designs should be developed to investigate metacognitive phenomena, and these should importantly include what he terms "on-line" methods of data collection, including the systematic observation and recording of behaviour. As we noted above, a range of such observational methodologies has been developed in recent work with young children.

Winne and Perry (2000) have argued that observational data regarding metacognition in young children has at least three advantages. It records what learners actually do, rather than what they recall or believe they do. It allows links to be established between learners' behaviours and the context of the task. And, particularly crucial for young children, it does not depend on the verbal abilities of the participants. We would also wish to argue for two further benefits. First, systematic observation, particularly where it involves video-recording, affords the opportunity to record non-verbal as well as purely verbal behaviour. Intriguingly, increasing evidence is emerging of the role of non-verbal behaviour in the development of young children's conceptual understandings and self-regulatory processes. In the closely related area of theory of mind, for example, Ruffman, Garnham, Import, and Connolly (2001) have demonstrated that 3-year-old children sometimes look to the correct location but give an incorrect verbal answer in a place-change false belief task. Analysis of their eye-gaze behaviour thus indicated a stage of implicit knowledge before fully conscious awareness which they were able to articulate. Further, recent work concerned with the role of gesture in conceptual learning and strategy development (Goldin-Meadow, 2002; Pine, Lufkin, & Messer, 2004) suggests that conscious articulation is only a part of the process of development in these areas. In other words, it seems highly probable that non-verbal behaviour is not only indicative in young children of metacognitive processes, but might also be an important part of the processes by which they are acquired.

A second further advantage afforded by observational methods in naturalistic, educational settings relates to the opportunity to record social processes involved in the development of metacognitive and self-regulatory abilities. There is, of course, as we have indicated above, a significant body of theoretical and empirical work, within the Vygotskian, socio-cultural tradition, which suggests that social processes have a crucial role to play in this area (Salonen, Vauras, & Efklides, 2005; Zimmerman & Schunk, 2001). Much of this work has emphasized the significance of mediation by an adult, and the impact of sensitive and contingent "scaffolding" in supporting children's learning. A range of studies, however, have also explored the significance of children's collaborative or peer-assisted learning of various kinds in the process of internalization of learning, and particularly in relation to the development of metacognitive and self-regulatory abilities (Elias & Berk, 2002; Karpov, 2005; Salonen et al., 2005; Whitebread et al., 2007).

Alongside observations of children's behaviours in naturalistic, educational settings (Perry, 1998; Whitebread et al., 2005), other significant developments have included the development of observational instruments and coding frameworks (Dermitzaki, Leondari, & Goudas, 2009; Mayr & Ulich, 2009; Ponitz et al., 2008; Rothbart, Ahadi, Hershey, & Fisher, 2001; Whitebread et al., 2009), observations of children's behaviours on age-appropriate problem-solving tasks (Annevirta & Vauras, 2006; Dermitzaki et al., 2009) and the development of interventions incorporating dynamic assessment through graded mediation (Lidz & Gindis, 2003; Polatajko & Mandich, 2004; Saldaña, 2004). Each of these methodological approaches are demonstrated within the Cambridge studies reported later in the chapter.

As regards observational instruments and coding frameworks, those which most directly and centrally address young children's metacognitive and self-regulatory development are the Strategic Behaviour Observation Scale (SBOS) developed by Dermitzaki et al. (2009) and the Children's Independent Learning Development (CHILD 3–5) instrument, and the Cambridgeshire Independent Learning (C.Ind. Le) coding framework developed by Whitebread et al. (2009). Dermitzaki et al.'s (2009) SBOS was developed for use with first- and second-grade children aged 6–8 years, and enables trained observers to rate children on 12 different cognitive, metacognitive and motivational strategic behaviours during their engagement with a cube assembly task. The three-factor structure of the scale was confirmed by factor analysis (with the cognitive and metacognitive strategic factors combining into a second-order cognitive self-regulation factor), using data from 168 children's performance on the task, and a high level of inter-rater agreement (intra-class coefficient 0.77) was achieved between observers. Cognitive self-regulation was shown to be related to task performance, and motivational strategic behaviour with domain-specific self-concept. Whitebread et al.'s (2009) CHILD observational instrument was designed for use with children aged 3–5 years (UK Foundation Stage) by their classroom teachers, based on their observations of the children's classroom performance. It is divided into four sections dealing with cognitive, emotional, motivational and social self-regulation, derived from a model developed by Bronson (2000). The 22-item scale was developed based on ratings of 192 children by 32 teachers, and has shown high levels of internal consistency (Cronbach's alpha = 0.97) and inter-rater agreement (95.5% of ratings within one judgement category). Early indications of good external validity of this instrument as a measure of metacognition and self-regulation in young children are also reported, but further research to refine and validate it is on-going in four European countries.

Within the C.Ind.Le project (Whitebread et al., 2005, 2007, 2009) a detailed coding framework was developed which identifies verbal and non-verbal behaviours indicative of metacognitive knowledge (of persons, tasks, and strategies), metacognitive regulation (planning, monitoring, control, and evaluation) and emotional and motivational regulation (monitoring and control). This framework was developed from an analysis of 582 sequences of behaviour video-recorded in UK Foundation Stage classrooms, which showed evidence of these facets of metacognition in children aged 3–5 years. Analysis of contextual variables indicated that the nature and prevalence of metacognitive and self-regulatory behaviours among these young children, as identified using the C.Ind.Le coding framework, were significantly influenced by the extent of adult intervention or direction, by the size of group in which the children were working, and by the nature of the task. Opportunities for collaborative groupwork and various kinds of peer-tutoring, which encouraged the children to articulate their ideas and explain their reasoning, were found to be significantly effective in stimulating metacognitive and self-regulatory behaviours (Whitebread et al., 2007). This framework has formed the basis for the coding of metacognitive and self-regulatory behaviours in several of the Cambridge studies reported later in this chapter.

While these observation-based methodological developments have enabled the clear identification of early metacognitive skills in young children, there remain, however, a range of methodological challenges in this kind of approach as Bakeman and Gottman (1997) and Veenman (2005) have usefully reviewed. The children's goals and intentions have to be inferred and their internal representations are not available (as they might be, to some extent through the use of think-aloud procedures with older children and adults). Only directly observable behaviours can be coded – so, for example, theory would predict that all control behaviours must be preceded by internal monitoring, but this cannot be coded unless it is directly observable (by, for example, eye gaze movements in checking behaviours). The high level of inference involved in identifying the "socially-based" behaviours involved in this kind of analysis is clearly a challenge to reliability. Levels of agreement concerning which behaviours constitute a unit of analysis in this kind of study are commonly around 60–70%, which clearly urges caution. It would clearly be ideal, for example, wherever resources allow, for all behaviour to be dual-coded and only those behaviours on which there is absolute agreement to be included in any analysis. However, the observation of behaviour in natural contexts, the involvement of class teachers who know the children and the classroom context well, and the use of video-recorded data, which can be viewed repeatedly by a team of researchers, all help in facing some of these challenges.

3 Theoretical Developments and Issues

These methodological advances have facilitated two important contributions to the study of metacognition. First, they have facilitated the development of new understandings concerning the emergence of metacognition in young children and its relationships with other early aspects of development. Second, work using observational and more age-appropriate methodologies has contributed significantly to our understandings concerning the general structure and inter-relations between metacognitive processes (Efklides, 2008).

As regards the emergence of metacognition in young children, research using these new methodologies has begun to present a much more positive picture of young children's metacognitive capabilities. Many of the limitations of metacognitive knowledge and "production deficiencies" attributed to young children by earlier research (Flavell, 1979; Kreutzer, Leonard, & Flavell, 1975) may well have been, at least to some extent, methodological artefacts. Schneider (1985), for example, demonstrated that many early studies examining the relationship between metacognition and performance in children were not particularly careful as regards the aspects of metacognition measured, often measuring just metacognitive knowledge using self-report methods, which under-estimated children's regulatory abilities. A range of other studies have demonstrated that when young children are presented with age-appropriate tasks related to memory (Istomina, 1975) or problem solving (Blöte, Resing, Mazer, & Van Noort, 1999; Deloache, Sugarman, & Brown, 1985), which are set in contexts that are playful and meaningful to them, they can demonstrate emergent metacognitive abilities. Bronson (2000) has provided a useful and comprehensive review of the large and blossoming research related to the development of cognitive, emotional, social, and motivational self-regulation in children from birth to the end of primary school (i.e., age 11 years).

Recent research has also begun to explore the relations and interactions of early metacognitive processes with other early cognitive developments. For example, observational methodologies in both naturalistic and problem-solving situations have uniquely facilitated the identification of non-verbal indicators of early metacognitive processes, such as eye gaze, gesture, pauses and changes in behaviour (Pino Pasternak, Whitebread, & Tolmie, 2010; Whitebread et al., 2007, 2009) which has contributed to theoretical developments recognising the role of implicit, or nonconscious, as well as conscious processes in metacognition (Efklides, 2008; Fitzsimmons & Bargh, 2004; Reder, 1996; Siegler, 1996). Extensive study of the relations between explicit metacognitive knowledge and performance, particularly in the area of metamemory, for example, has shown that, while they become stronger with age, they are never particularly high (Lockl & Schneider, 2007; Schneider & Bjorklund, 1998). Siegler (1996), in his theory relating to children's development of cognitive strategies, has also concluded that the metacognitive processes involved in strategy selection, certainly in children, are predominantly of an implicit nature and unavailable to conscious awareness. As part of this increasing recognition of the role of implicit processes, there is a developing empirical and theoretical literature beginning to explore links between the early emergence of metacognition, early executive functions such as inhibition, working memory and attention (Blair, Zelazo, & Greenberg, 2005; Fernandez-Duque, Baird, & Posner; 2000; Whitebread, 1999) and theory of mind (Flavell, 2004; Kuhn, 2000).

At the same time, observational studies in naturalistic contexts have contributed to understandings concerning metacognition more generally. For example, studies of this kind with young children have enabled and supported research adopting a more socio-cultural perspective which has emphasised the role of social context and social processes in early metacognition (Boekaerts & Corno, 2005; Iiskala, Vauras, & Lehtinen, 2004; Meyer & Turner, 2002; Perry, 1998; Salonen et al., 2005;

Whitebread et al., 2007) and the complex interactions between metcognition and motivation, or affect, within the broader conceptual notion of self-regulation (Boekaerts & Niemivirta, 2000; Efklides, 2006; Pintrich, 2000).

Central to the increasing recognition of the importance of social processes in the development of metacognitive and self-regulatory abilities is the Vygotskian notion of learning as a process of acculturation or internalization whereby the child moves from being other-regulated to being self-regulated (McCaslin & Hickey, 2001). Research using this framework in analysing regulatory processes in classroom learning, by such as Iiskala et al. (2004), has shown that regulation can be directed to the self (self-regulation), or to others (co-regulation), or can be shared (shared-regulation) (see also Efklides, 2008). Quite a body of work has now shown that, even with quite young children, social, collaborative forms of learning, including group-work and peer-tutoring, can enhance metacognitive processes and learning in classroom situations (Brown, 1997; Whitebread et al., 2007).

The relation of metacognition to the broader notion of "self-regulation" has been a further theoretical issue to which research with young children has made a significant contribution (Efklides, 2008; Pino Pasternak et al., 2010). Despite some early confusion over definitions and terminology – with "regulation" being used initially to refer to purely cognitive monitoring and control processes (Brown, 1987) – a broad consensus is now emerging that metacognition refers specifically to the monitoring and control of cognition, while "self-regulation" refers more generally to the monitoring and control of all aspects of human functioning, including emotional, social and motivational aspects (Boekaerts & Niemivirta, 2000; Pintrich, 2000). This shift in meanings has paralleled and been necessitated by important theoretical developments which have recognized that metacognitive abilities have an impact on behaviour and performance, but that this also depends upon the degree of effort that the individual decides to exert in relation to any particular task. The individual's beliefs about the value of the task, their affective response to it – e.g., feeling of difficulty (Efklides, 2006) – and the attributions they make based on previous success and failure on similar tasks all impact upon their "goal-orientation" and, thus, their metacognitive performance (Boekaerts & Niemivirta, 2000; Pintrich, 2000). This recognition has led Paris and Paris (2001) to refer to self-regulated learning as the "fusion of skill and will" (p. 98). An extensive range of empirical studies have confirmed these theoretical inter-relationships between metacognitive and emotionally and motivationally self-regulatory processes (Dermitzaki et al., 2009; Efklides, 2006; Pintrich, 2000).

4 Current Studies

In the remainder of this chapter, in order to illustrate the current theoretical and methodological trends in research related to metacognition in young children, we review a number of our own research studies carried out using observational methods, interventions and tasks which reduce the dependence on young children's verbal abilities. These studies have explored metacognitive development in typically developing young children, and in young children with learning and motor difficulties. Within the typically developing group, studies are reported which have explored the relation between early metacognition and inhibition, theory of mind and conceptual development.

4.1 Early Metacognition and Executive Functioning

This renewed interest in the early stages of metacognitive development has been supported by, and in turn has contributed towards, theoretical developments bringing together previously separate literatures concerned with metacognitive processes, executive functioning and theory of mind. These literatures have largely run in parallel until the very recent past, but in the last few years there have been attempts to explore relationships and conceptual equivalences between them which might facilitate a more integrated and improved understanding of the nature and course of early cognitive development (Efklides, 2008; Fernandez-Duque et al., 2000; Flavell, 2004; Kuhn, 2000).

A study conducted by Donna Bryce in the UK (Bryce, 2007; Bryce & Whitebread, 2008) has examined the relations between early metacognitive development and executive functioning in children aged 5 and 7 years. Metacognitive skilfulness was defined in this study as a child's ability to monitor and control their own learning (Nelson & Narens, 1990). The particular executive function considered, namely "inhibitory control", has been defined as the ability to suppress information, actions or emotions when they are inappropriate or no longer relevant (Deak & Narasimham, 2003).

The theoretical framework adopted in this study was based on an early model of metamemory proposed by Nelson and Narens (1990). The basis of the model is that all metacognitive processes can be considered as either monitoring or control processes. Monitoring processes involve updating one's mental representation of the current situation while control processes assert some action, such as changing strategy. Therefore, "error detection" could be considered a monitoring process, and "error correction" a control process (Nelson & Narens, 1994). The contribution of Bryce's (2007; Bryce & Whitebread, 2008) framework is to propose that all executive functions may similarly be involved in either monitoring or control processes. Specifically, it was hypothesised that inhibitory control would be related to metacognitive control processes.

In this study (Bryce, 2007; Bryce & Whitebread, 2008) the children's natural use of metacognitive skills, including monitoring skills such as checking, error detection and self-commentary, and control skills such as changing strategy, using gesture to support a cognitive activity and organising or grouping materials, was observed during a simple, familiar problem-solving task. The task involved a child making a train track from a range of wooden train track pieces to match a blank shape in a plan (adapted from Karmiloff-Smith, 1979). Minimal assistance was given with the task, with only gentle encouragement if the child directly asked for help. Both direct verbalisations and non-verbal behaviour were coded as reflecting either monitoring or control processes, adhering to Nelson and Narens' (1990, 1994) model.

Inhibitory control was assessed in the classical manner for work on executive functioning – a computer-based task with easy to follow instructions. In this case a novel Animal Stroop task was devised and used, where two animals of different physical sizes were presented on a screen and the child was required to select the animal that was larger *in real life*. Both this task and the train track task were designed to be minimally dependent on verbal skills and previous knowledge, so as to reflect the true abilities of young children.

The study provided support for a strong link between inhibitory control and the metacognitive control behaviours listed above. Having controlled for a range of potentially confounding variables, including age, non-verbal IQ, receptive vocabulary and verbal working memory, significant correlations were found between measures of inhibitory control and behavioural indicators of control processes. For example, reaction time on incongruent trials ($r = -0.41$, $p = 0.032$) and the difference in reaction times between congruent and incongruent trials ($r = -0.46$, $p = 0.016$) in the animal Stroop task correlated significantly with the rate per minute of control codes in the train track task. At the same time, intriguingly, no significant correlations were found with monitoring behaviours. Further, strong support was found for the validity of the metacognitive skills coding scheme, as control rates correlated positively with end-product quality, was supported by parent questionnaire, and a high inter-rater agreement (91%) was achieved.

However, although the theoretical framework adopted in this study appears to be supported by empirical evidence, the precise relationships between individual executive functions and early metacognition are not yet established. For instance, do executive functions act as precursors to metacognitive skilfulness, how interrelated are all control executive functions, and are they all equally important for metacognitively skilled behaviour? Can you, for example, be metacognitively skilled even if you have poor inhibitory control? It may be that executive functions are necessary but not sufficient precursors of metacognitive control behaviours. The study (Bryce, 2007; Bryce & Whitebread, 2008) also produced evidence that inhibitory control was more strongly related to metacognitive skilfulness at age 5 than at age 7. When the correlations reported above were examined within age groups, they increased markedly within the 5-year-old group ($r = -0.67$, $p = 0.017$, and $r = -0.73$, $p = 0.007$, respectively) and reduced below statistically significant levels for the 7-year-old group.

Executive functions have traditionally been considered as necessary, basic skills for normal functioning (e.g., you either can or cannot inhibit inappropriate responses) and the lack of them has been implicated in the aetiology of some developmental disorders (e.g., autism, ADHD; Happe, Booth, Charlton, & Hughes, 2006). However, metacognitive skilfulness is a rather fuzzy concept. It can be considered as a person's propensity to use these "basic skills" in everyday situations. As such, it is highly dependent on motivation for the task, emotional state, experience of the task, and so on. Indeed, as we have noted earlier, some influential models of self-regulation place great emphasis on goal orientation and motivation. In short, executive functions may be seen as a person's basic cognitive functions, while metacognitive skilfulness possibly reflects a person's tendency to use these skills successfully in problem-solving situations.

4.2 Early Metacognition and Theory of Mind

While work examining the relationships between executive functions and metacognition is just emerging, there is a fairly longstanding and established body of research which has investigated the relationship between executive functions, particularly inhibitory control, and theory of mind (Perner & Lang, 1999). At the same time, while there have been some interesting theoretical speculations (Bartsch & Estes, 1996; Kuhn, 2000; Wellman, 1985), there has been little empirical work examining relationships between theory of mind and metacognition. What work there has been, however, has supported the general view we noted at the outset of this chapter, that theory of mind is an earlier development and predicts later metacognitive abilities (Lockl & Schneider, 2007).

A study carried out by Demetra Demetriou (Demetriou, 2009; Demetriou & Whitebread, 2008) aimed to provide a comprehensive account of false belief understanding development and, to this end, used a longitudinal (repeated measures) design. The selected children were Cypriot and just under 4 years of age at the start of the study. They were assessed at three test points (T1, T2 and T3) with an interval of 6 months between them. As recently conducted research studies have shown theory of mind is not a clear-cut developmental achievement which emerges at one point in time (Carlson, Wong, Lemke, & Cosser, 2005; Flynn, O' Malley, & Wood, 2004), this design appeared to be appropriate to examine its early emergence and development.

At each test point performance was examined on two false-belief tasks alongside performance on a range of tasks designed to assess significant aspects of cognitive and metacognitive development, namely verbal and non-verbal IQ, working memory, inhibition control, metarepresentational abilities (i.e., understanding that signs can misrepresent reality), language skills, and source memory. At test point 3 (T3) the children's teachers were also asked to complete the CHILD 3–5 checklist for each child to measure the overall levels of self-regulation they had achieved by the end of the testing period (Whitebread et al., 2005, 2009).

Source memory (Gopnik & Graf, 1988) involves knowing the source of one's own knowledge, and is a type of metacognitive knowledge which seems highly likely to be related to understanding false beliefs. A review of the literature reveals, however, that this possible relationship has been surprisingly little explored (Taylor, Esbenson, & Bennett, 1994; Wellman & Liu, 2004), although some interest has emerged in a number of recent cross-sectional studies, using a variety of measures (Bright-Paul, Jarrold, & Wright, 2008; Lind & Bowler, 2009; Naito, 2003).

In our longitudinal study source memory was assessed using two tasks. The first was devised by Taylor et al. (1994) and required children to report, as part of a colouring game, when they had learnt a new colour word (when they were a baby, before the game or during the game) and how they had learnt it (a puppet told them). The second was devised by Gopnik and Graf (1988) and required children to report how they had learnt the contents of some drawers (by being shown, by being told, or by being given a clue). As we shall see, the longitudinal nature of Demetriou's study has provided new evidence concerning the nature of the relationship between source memory and theory of mind.

Simple correlational analysis demonstrated significant and in some cases strong relationships among false belief understanding and metarepresentational ability, inhibition control, working memory, language, source memory, verbal IQ and the composite score of the CHILD 3–5 checklist (ranging from $r=0.31$, $p<0.05$. for inhibition control and working memory at T1 to $r=0.60$, $p<0.01$ for language at T3). However, the strongest relationship at all three test points was that between false belief understanding and source memory, and this relationship became progressively stronger, achieving $r=0.83$, $p<0.01$ at T3, when the children were around 5 years of age. Further, this relationship between the constructs of theory of mind and source memory remained significant after controlling for the effect of executive functions (using partial correlations), such as working memory and inhibition control, and after controlling for the effect of language ability (T3: working memory, $r=0.75$, $p<0.01$; inhibition control, $r=0.79$, $p<0.01$; language, $r=0.72$, $p<0.01$). The CHILD 3–5 checklist, which was employed as a broader measurement of self-regulation and metacognition, also correlated significantly with the composite variable of false belief understanding at T2 and T3. The strongest correlation traced between the two constructs was at T2 and was moderate in size, $r=0.50$, $p<0.01$.

To further examine the nature of the relationship between theory of mind and source memory, regression analyses with variables within and across testing points were conducted, and a procedure was adopted which was proposed by Astington and Jenkins (1999), and further employed by Lockl and Schneider (2007) in order to explore the direction through time of the relationship between the two abilities. According to this procedure (known as "cross-lagged panel analysis"), the direction of influence between, for example, source memory and false belief understanding was assessed as follows. First, in order to predict source memory at a later test point (let's say T2), source memory at T1 is entered into the regression model at Step 1 and false belief understanding at T1 is entered at Step 2. Then, false belief understanding at T2 is predicted by entering false belief understanding at T1 followed by source memory at T1. The key point about this method is that by entering the earlier ability of the skill under investigation at Step 1 it becomes possible "to assess the contribution of each variable to changes in the other variable" (Lockl & Schneider, 2007, p. 159).

These analyses established important and interesting aspects of the relationship between theory of mind and source memory. First, the relationship found between these two developing aspects of cognition could not be accounted for by a third variable such as language, verbal IQ, inhibition control or working memory.

The predictors of longitudinal regressions between the three test points of each of the two constructs were very different. Those involved in the false belief understanding longitudinal regressions, apart from earlier false belief understanding, were working memory, verbal IQ and source memory, while the predictors involved in source memory longitudinal regressions were inhibition control, false belief understanding and language.

Second, the findings provide evidence that the relationship between false belief understanding and source memory may well be bidirectional. Between T1 and T2 earlier source memory significantly predicted later false belief understanding (14.5%) and earlier false belief understanding predicted later source memory (5.6%); between T2 and T3 earlier false belief predicted later source memory performance (4.8%) and between T1 and T3 earlier source memory predicted later false belief understanding (24.4%).

The findings of Demetra Demetriou's study taken together suggest that the development of theory of mind is strongly inter-related with that of source memory and broader measures of self-regulation (the CHILD 3–5 checklist score correlating significantly with false belief understanding at T3, $r=0.41$, $p<0.01$). The data demonstrating that earlier source memory predicts later theory of mind development opens a new chapter in this research literature and suggests that the notion that theory of mind is an earlier socio-cognitive achievement and metacognition a later achievement may have been a methodological artefact which needs to be re-examined. Interesting findings have also been reported recently concerning relations between source memory and later cognitive skills (e.g., for text comprehension see Strømsø, Bråten, & Britt, 2010) which would support the view that source memory is a significant element in metacognitive processing which has early and continuing significance within cognitive development.

4.3 Metacognition and Conceptual Development

A third area in which current research concerned with early metacognition is endeavouring to develop theoretical relationships with other aspects of development concerns the relationship between self-regulated learning and conceptual development. A study carried out by Valeska Grau (2008a, b) has attempted to explore these developing relationships with Chilean children in the domain of biological science. This domain was chosen because it involves declarative and procedural knowledge, constitutes a core domain in human cognition, involves certain commonly found difficulties for conceptual learning and, consequently, raises interesting challenges for teaching and learning (Inagaki & Hatano, 2002). In line with recent theoretical developments, this research has highlighted, in particular, the significance of motivational aspects of metacognition and self-regulation, such as a mastery-orientation towards learning (Dweck, 2000; Pintrich, 2000) and the importance of investigating young children's learning and development in natural school contexts and within a socio-cultural framework.

As with the studies outlined above, Valeska Grau's (2008a, b) study was located in an area in which there have been interesting theoretical developments, but little empirical work at present. Two promising theoretical frameworks currently attempting to link metacognition and conceptual development are those related to "adaptive expertise" (Hatano & Oura, 2003) and "intentional conceptual change" (Limón, 2003; Sinatra & Pintrich, 2003). There are links within both of these approaches to the mastery-oriented motivational pattern proposed by Dweck (2000) and recognition of the significance of emotional and motivational processes involved in self-regulation (Boekaerts & Niemivirta, 2000; Pintrich, 2000). The "adaptive expertise" approach emphasises intellectual adventurousness, flexibility and playfulness, while intentional conceptual change emphasises volitional aspects, implying that individuals must want to change, considering change as an intrinsic personal goal, and not one imposed by others.

The methodological approach of this study consisted of a multiple case study of eight third grade primary school children in Chile (aged 8 years). The design partly followed a microgenetic approach with repeated measurements of the children's performance on particular tasks over a 5-month period (a semester). The data collection process was as naturalistic as possible in order to understand the processes of conceptual development and self-regulated learning in the real context in which they occur. The methods of data collection included mainly videotapes of children working by themselves and in groups in activities related to biological concepts.

Different kinds of qualitative and quantitative procedures were used to analyse the different sets of data. In the case of self-regulated learning two coding schemes were developed (one for individual activities and another for group activities) which included verbal and non-verbal indicators of metacognitive and regulatory processes, including planning, monitoring, control, and evaluation. The group coding included coding for self-, other-, and shared-regulation as identified by Iiskala et al. (2004). Also, when analysing the behaviour of children solving a task on their own, a thematic analysis of post-task reflections was included. Finally, in the case of the observations of children while engaged in a group-task, a sociocultural discourse analysis was added as a way to grasp phenomena related to shared-regulation of learning, when the children were collectively regulating the activity (e.g., sequences of utterances that built constructively upon one another towards some shared understanding or strategic decision within the group). All the different kinds of data gathered were brought together within cases and in a cross-case analysis which compared groups of cases who showed high, medium, or low levels of self-regulation overall.

A key finding from this study was that the children who tended to show good general metacognitive and regulatory skills at the beginning of the semester, such as high abilities for reflecting on their own knowledge and skills and high levels of mastery orientation, were the children that were more successful in developing domain-specific strategies to solve the tasks specifically related to biological science, once they had acquired an adequate level of knowledge. This finding gives support to models of learning which bring together domain-general and domain-specific capacities of human cognition (Hirschfeld & Gelman, 1994) and models of expertise which relate the acquisition of knowledge with the development of metacognition and self-regulated learning skills, such as the Model of Domain Knowledge formulated by Alexander (Alexander, 2003; Alexander, Jetton, & Kulikowich, 1995) or the Model of Intentional Conceptual Change of Limón (2003).

The analysis also revealed that the children who were more playful, more mastery-oriented in their motivational style and more “intellectually adventurous” (e.g., attempting a wider range of strategies to solve tasks) were also the children who showed a greater extent of metacognitive and regulatory skills across different contexts, giving support to the influence of motivational beliefs, emotional processes, and broader contextual issues related to the educational opportunities provided within different school classrooms (Perry, 1998; Whitebread et al., 2005, 2007).

4.4 Metacognition and Learning Difficulties

Alongside studies of metacognitive and self-regulatory processes in young typically developing children which, as we have seen, have focused mainly on the relationships between early metacognitive and other aspects of development, a further important focus of current research concerns the efficacy of metacognitive training in children with various kinds of learning difficulties (LDs), including problems with attention, memory, problem solving, reasoning, transfer of learning, and language and literacy (Gersten, Baker, Pugach, with Scanlon, & Chard, 2001). It is well established, of course, that children with LDs demonstrate some metacognitive and self-regulation deficits, namely relatively limited abilities to use goal-oriented strategies effectively, efficiently and flexibly (Cameron & Reynolds, 1999). A growing body of research has produced promising results in relation to using metacognitive strategies with children who experience difficulties in learning (Davis & Florian, 2004; Florian, 2007). The intervention studies reported here, however, have shown that this approach is equally possible with younger children. Moreover, as we shall see, our studies have predominantly adopted a socio-cultural framework within which observational methods were used to analyse the dynamic interactions between teacher and pupil. The two studies have involved teaching mathematics to children with LDs and children with motor development difficulties, respectively.

Qais Almeqdad (2008a, b) carried out a study exploring the use of self-explanations as a metacognitive strategy. Typically, self-explanation is a technique whereby the learner is required to explain their reasoning, or the strategy they adopted, when carrying out a problem-solving task. This requires the learner to engage in metacognitive processes involving self-monitoring and self-questioning (Kavale, 2007). Previous research evidence, predominantly with older students and typically developing children, has suggested that the generation of self-explanations aids learners to build and transfer knowledge (Renkl, Stark, Gruber, & Mandl, 1998), become more metacognitively aware of their own strategies and level of understanding and develop as learners in a wide variety of knowledge domains (Chi, 2000; Siegler, 2002). This study therefore attempted to explore the extent to which children with LDs could be trained to benefit from using self-explanation strategies, and, if so, which types of self-explanation (explaining concepts and relationships, prediction, task solving procedures, cause and effect, etc.) could they use most productively?

The Qais Almeqdad (2008a, b) study was designed as an exploratory intervention, in which three resource room teachers in Jordanian primary schools were recruited and trained on the use of the self-explanation technique. They worked with 20 children aged 7.5–11.5 years who had LDs, and taught them basic mathematics for a period of 14 weeks. Systematic video observations were conducted during the teacher-pupil interactions, and a refined coding scheme was developed to code and analyze this data. So, for example, if a child explained their understanding of the functions and use of a mathematical symbol, that was coded as “concepts and relationships”; if a child explained how they planned to solve a particular problem, that was coded as “prediction”; if they explained the strategies they had used to solve a task, that was coded as “task solving procedures”; if they explained why they had used a particular strategy and why it worked, this was coded as “cause and effect”.

The findings showed that the participant LD children were capable of generating different types of self-explanations in response to a variety of questions posed by their teachers. In particular, they were capable of identifying concepts and relationships, engaging in predictions of mathematical operations, procedures and outcomes, articulating the procedures which they applied to solve a given task, explaining the cause and effect of their own reasoning, explaining the cause and effect of the teacher’s reasoning, and, finally, evaluating their own previous knowledge. The number of self-explanations generated by the children was found, not surprisingly, to be associated with the number of questions asked by the teachers. However, an interesting finding which emerged was that as the instruction proceeded, the number of both the teachers’ questions and the children’s self-explanations decreased in most of the self-explanations categories, with the exception of the categories of “predictions” and “task solving”, in which both questions and self-explanations increased in the later sessions, which is in line with previous research with older and typically developing learners (Renkl et al., 1998). This appears to suggest that the teachers recognised that the children demonstrated better self-explanations in these categories and, therefore, that they responded in their questioning strategies to this aspect of the children’s performance.

In the previous literature, there have been contrasting results in relation to the ability of children with LDs to generalise their use of learning strategies (Davis & Florian, 2004; Florian, 2007). However, in this study it was clear that the children were able to maintain their use of self-explanations in a follow-up session, 2 weeks after the end of the intervention. Informal reports by class teachers also indicated that all 20 of the children, to varying extents, also transferred the use of self-explanations to their work in the mainstream classroom. One pair of children, for example, were observed by their class teacher to have begun playing the “self-explanation game” in class, with one of the children taking on the role of the teacher, asking the questions, and the other child providing the self-explanations. The findings also suggested that there was a positive relationship between the children's achievement on the pre- and post-tests and their overall generation of self-explanations. A comparison of the frequency of self-explanations between groups of children with high, medium and low achievement scores on a standard pre-test revealed a highly significant difference, $\chi^2(2, N=20)=6.88$, $p=0.011$. At post-test a simple statistical comparison was not possible, as the children were each given different individual post-tests based on their learning objectives, which were identified based on their pre-test performance. However, of the 20 children in the sample, the seven children with good scores (above 75%) on their post-test were all in the top 12 as regards frequency of self-explanations, while the four children who scored poorly (below 50%) in their post-test were all in the bottom seven as regards frequency of self-explanations. There was also, once again, some evidence of the teachers responding to these differences, as high achievers on the individual achievement post-tests were found to generate more fully articulated self-explanations and to receive relatively less follow-up and evaluation questions from their teachers.

This study clearly indicated the benefits of using self-explanations in teaching children with LDs, since it enabled the teachers and students to engage in interactive learning processes which supported the teachers' practices, and enabled the children to become more metacognitively aware of their own learning and develop as strategic self-regulated learners.

4.5 Metacognition and Motor Learning Difficulties

The last study we wish to report in this chapter has been carried out by Claire Sangster (2009a, b) and concerns a cognitively-oriented intervention program for children with motor LDs. The use of observational methodology for investigating metacognition and self-regulation in young children becomes particularly pertinent when the skill under investigation is a motor task. Indeed, the process of acquiring a motor skill produces a wealth of evidence for self-regulation and metacognitive knowledge that is not readily accessible through the verbal reports of children, but rather through the observation of motor performance and practice. There has been much recent interest in the role of cognitive and metacognitive skills in motor skill acquisition, and their influence amongst children with motor LDs (Kitsantas & Zimmerman, 2002; Lloyd, Reid, & Bouffard, 2006).

Recently a cognitive learning paradigm has been applied to the study of children with developmental co-ordination disorder (DCD), a specific learning difficulty characterised by impaired performance of motor skills (Sugden & Chambers, 2005). This paradigm is based on the assumption that DCD children have fewer cognitive and metacognitive skills with which to acquire motor skills and solve motor performance problems. It has been applied to studying the nature of the problems experienced by children with DCD and developing appropriate intervention strategies (Sugden & Chambers, 2005). However, individual differences in the use of self-regulatory and metacognitive skills during motor learning have received little attention in the research literature. The Sangster (2009a, b) research aimed to develop a better understanding of the self-regulatory differences exhibited during motor learning between children with and without difficulties, whether these differences contribute to the performance problems experienced by children with DCD and how a cognitively oriented intervention program could facilitate change in self-regulatory skills during motor learning.

Fifteen children aged 7–9 years participated in ten intervention sessions, which aimed to improve performance on a motor task selected by the children themselves. The children were divided into three ability groups, namely a group of children with DCD (DCD), a group of children with DCD and co-occurring difficulties in learning and/or attention (DCD+), and a group of typically-developing children (TDC). The intervention employed was a program called Cognitive Orientation to daily Occupational Performance (CO-OP), which assists children in learning and applying a problem-solving strategy aimed at enabling them to discover and use cognitive and metacognitive skills in order to overcome motor performance difficulties (Polatajko & Mandich, 2004). The video-recorded sessions of all children were analysed in detail using a coding framework developed using existing models of self-regulation and motor learning (Ferrari, 1996; Kirschenbaum, 1984), pre-existing coding schemes (Sangster, Beninger, Polatajko, & Mandich, 2005) and the data itself. This framework included coding schemes aimed at observing children's self-regulatory behaviour (i.e., goal setting, planning, monitoring or evaluating performance, and strategy use) as well as evidence of task and metacognitive knowledge (including knowledge about oneself as a motor performer, about task, environmental and strategy variables that could influence performance, and the role of self-regulation in influencing improvement in performance).

Previous research has suggested that, while children with motor difficulties are often able to provide evidence of cognitive and metacognitive skills on task, these skills are often ineffective or inappropriate to the situation (Lloyd et al., 2006). For example, a child experiencing difficulties in motor performance might select a goal that is well beyond his or her current ability or select a strategy for task practice that is ineffective for improving performance. As such, it was the aim of the current research to determine specifically whether there existed group differences in the quality of self-regulation observed and whether the intervention program served to modify this quality. To do so, each coded behaviour was additionally given a designation of quality on one of four levels:

- *Independent*: Spontaneous self-regulation (e.g., "I'm going to see if I can hit the target five times in a row!")
- *Cued*: Self-regulation in response to a simple prompt, question, or cue (e.g., "So, what's the goal for this activity?", "Do you have any strategies for reaching your goal?")
- *Mediated*: Self-regulation occurring under direct and supportive external mediation (e.g., "Why don't you watch me and see if you can tell me what I'm doing to make sure the ball gets all the way to you?")

- *Ineffective*: Self-regulation that is inappropriate or ineffective in the context of the situation, despite cueing and mediation (e.g., after failing to dribble a ball around some skittles, "I think I did it too slow, I' m going to try going faster!" followed by running at it and losing the ball immediately on the next attempt)

This coding scheme is an excellent example of a form of dynamic assessment through graded mediation, which is an approach which has been successfully applied in a number of recent studies (Lidz & Gindis, 2003; Polatajko & Mandich, 2004; Saldaña, 2004). In this case it was applied to an in-depth review of video-recorded sessions, after which code data was used to conduct a comparison across groups. At the outset of the program, children in the TDC group had the highest levels of independent self-regulation and the lowest levels of mediated and ineffective regulation. Children in the DCD+ group appeared to have the most difficulty with the performance of self-regulation skills, exhibiting the lowest levels of independent self-regulation and most frequent instances of mediated and ineffective forms of self-regulation skills. While children in the DCD group managed to exhibit slightly more frequent indicators of independent regulation than their peers in the DCD+ group, ineffective self-regulation skills was similarly frequent in this group in early sessions. In all groups, children's demonstration of cued SR skill varied considerably across cases and was argued to be dependent on the complex interaction between each child's emerging skill, the demands of the task and the practice context. Together, these results suggest that, before intervention, children without any motor difficulties exhibited the most highly developed repertoire of self-regulation skills. In contrast, children with DCD were more likely to ineffectively self-regulate their own motor performance or require significant levels of external support to do so effectively. More promisingly, the SR skill of children in both DCD groups improved over the program, where growing levels of independent regulation and a reduction in ineffective regulation were observed. Arguably, this finding reflects a positive effect of the intervention program on the development of self-regulation skills during motor task practice.

A comparison of expressed task and metacognitive knowledge revealed that the children in the TDC group consistently expressed the highest levels of knowledge while those in the DCD+ group exhibited the least frequent expressions of knowledge. More interestingly, individual case analysis across DCD groups revealed that verbal indicators of knowledge were not always consistent with children's demonstration of self-regulation skills during performance. Namely, while some children were able to explicitly report a wealth of relevant knowledge concerning the selected motor task and variables influencing task performance, they were unable to effectively apply such knowledge to practice through self-regulated performance. Conversely, some children exhibited effective self-regulation skills during practice but subsequently failed to report an understanding or awareness of such ability during reflective discussion. This apparent dissociation between children's on-task performance of self-regulation skills and verbal expression of knowledge is consistent with previous research, which argues that children with DCD often fail to integrate the knowledge, monitoring and planning elements of the self-regulation of motor performance (Lloyd et al., 2006). It further supports current understanding regarding the nature of metacognition itself, where both the implicit performance of self-regulation skills and the conscious articulation of knowledge are equally critical reflections of metacognitive ability (Whitebread & Pino Pasternak, 2010). It is also consistent with the argument that the relationship between these two elements in the early stages of metacognitive development is not necessarily a strong one (Schneider & Bjorklund, 1998).

To summarize, the use of an observational approach to examine metacognitive and self-regulatory behaviour in young children with and without DCD allowed for the discovery of clear group differences in the acquisition and application of self-regulation skills during on-task motor practice. The present research further confirms that while children with DCD do engage in self-regulatory activity during motor learning, they often do so poorly. It also suggests that a cognitively-oriented intervention program, previously demonstrated to have a positive effect on motor performance itself (Polatajko & Mandich, 2004), additionally has a positive influence on improving the quality of self-regulatory activity during motor skill acquisition. Finally, it has illustrated the complex and dynamic relationship between self-regulated performance and articulated knowledge and, as a result, highlights the importance of examining both implicit and conscious forms of self-regulation and metacognition through the use of multi-method frameworks that combine both observational and self-report methods.

5 Conclusion

It is hoped that the present chapter has amply demonstrated that research in metacognition in young children is a fruitful and productive enterprise. We have reviewed studies of typically developing children as young as 3 years of age, and of young children with learning and motor difficulties. As we argued at the outset of the chapter, these studies have been made possible by the emergence and development of observational methodologies and the development of age-appropriate tasks which are meaningful to young children and reduce the dependence on the children's verbal abilities. It seems likely that further developments in these kinds of methodologies, including the development of agreed and refined observational coding frameworks in particular domains, of neuroscientific methods, and of forms of dynamic assessments facilitated by incorporating social contexts into our research frameworks, will continue to facilitate productive investigations in this area.

These new methodologies have enabled each of the present studies to be innovative, also, in the scope of their investigations. We have seen that the exploration of metacognitive processes in young children can illuminate their functional relationships with executive functions, with theory of mind, with conceptual development and with affective and motivational aspects of self-regulation.

Specifically, the study by Bryce (Bryce, 2007; Bryce & Whitebread, 2008) has shown that metacognitive monitoring processes, such as checking, error detection and self-commentary, are evident in children as young as 5 years of age and may be relatively independent of the executive functioning processes involving inhibitory control. However, the ability to act upon the information derived from this monitoring, through metacognitive control processes such as changing strategies, using gesture to support a cognitive activity or organising or grouping materials, appear to be initially, in 5-year-olds, strongly associated with inhibitory control capabilities. The study by Demetriou (Demetriou, 2009; Demetriou & Whitebread, 2008) has further shown that the relationship between aspects of early metacognitive processing and the initial achievement of a theory of mind by children aged 4–5 years are much more reciprocal than previously understood. Specifically, the ability to monitor the source of knowledge and to hold this information as declarative, metacognitive knowledge, was shown to be achieved by children in this age group, and to be strongly implicated in their performance on false belief tasks. As regards the relationships between early metacognitive abilities and conceptual development, the study by Grau (2008a, b) has shown that, by 8 years of age, children can use domain-general metacognitive abilities to develop domain-specific strategies as knowledge is acquired in the particular domain. This study also provided evidence that children's motivational style was instrumental in the development of their metacognitive and self-regulatory skills across different contexts.

We have also shown that observational research with children can reveal the impact and nature of metacognitive processes and interventions in a range of different domains of learning, including biological science, mathematics and motor development. The use of observational methodologies, enabling the collection and analysis of verbal and non-verbal behaviours, has allowed the investigation of the implicit as well as the conscious processes which comprise metacognitive performance. The studies of Almeqdad (2008a, b) and Sangster (2009a, b) used observational and dynamic assessment methodologies to investigate the impact of metacognitive interventions on teaching mathematics to children with learning difficulties and on teaching motor skills to children with DCD. In both cases clear evidence was produced that children as young as 7 years of age could respond productively to such innovations, and that improvements in self-regulatory abilities were clearly linked to improvements in performance. These studies raise the issue of the intriguing relation between metacognitive control and declarative metacognitive knowledge. In Almeqdad's study the use of the self-explanation technique explicitly involved children in articulating their mathematical understandings and there was evidence of children adopting and transferring the technique to their regular classroom learning. The evidence from Sangster's study of children with DCD, however, suggested a more complex and dynamic relationship between self-regulated performance and declarative metacognitive knowledge.

There is currently, as we have reviewed, considerable and growing concern to pull together the various literatures concerned with early metacogniton, self-regulation, executive processing, theory of mind, motivation, conceptual development and so on, and to establish a more coherent and holistic view of the processes by which children and adults learn and, in turn, develop their capabilities as learners (Efklides, 2008). It is quite clear, we would argue, that the study of metacognition in young children has the potential to make an important contribution in this endeavour.

References

Alexander, P. A. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher, 32*(8), 10–14.

Alexander, P. A., Jetton, T. L., & Kulikowich, J. M. (1995). Inter-relationship of knowledge, interest and recall: Assessing a model of domain learning. *Journal of Educational Psychology, 87*(4), 559–575.

Almeqdad, Q. (2008a). *Self-explanation and explanation in children with learning difficulties.* Unpublished PhD thesis, University of Cambridge, UK.

Almeqdad, Q. (2008b). *Self-explanation in children with learning difficulties: differences in age, ability and academic achievement.* Paper presented at 3rd Biennial Meeting of EARLI Special Interest Group 16: Metacognition, Ioannina, Greece.

Annevirta, T., & Vauras, M. (2006). Developmental changes of metacognitive skill in elementary school children. *Journal of Experimental Education, 74*(3) 197–225

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*(5), 1311–1320.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction* (2nd ed.). Cambridge, UK: Cambridge University Press.

Bartsch, K., & Estes, D. (1996). Individual differences in children's developing theory of mind and implications for metacognition. *Learning and Individual Differences, 8*(4), 281–304.

Blair, C., Zelazo, P. D., & Greenberg, M. T. (Eds.). (2005). Measurement of executive function in early childhood. *Developmental Neuropsychology, 28*(2), 561–729.

Blöte, A. W., Resing, W. C., Mazer, P., & Van Noort, D. A. (1999). Young children's organizational strategies on a same-different task: A microgenetic study and a training study. *Journal of Experimental Child Psychology, 74*, 21–43.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review, 54*(2), 199–231.

Boekaerts, M., & Niemivirta, M. (2000). Self-regulated learning: Finding a balance between learning goals and ego-protective goals. In P. Pintrich, M. Boekaerts, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 417–450). San Diego, CA: Academic.

Bright-Paul, A., Jarrold, C., & Wright, D. B. (2008). Theory-of-mind development influences suggestibility and source monitoring. *Developmental Psychology, 44*(4), 1055–1068.

Bronson, M. B. (2000). *Self-regulation in early childhood.* New York: Guilford.

Brown, A. L. (1987). Metacognition, executive control, self-regulation and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 65–116). London: Erlbaum.

Brown, A. L. (1997). Transforming schools into communities of thinking and learning about serious matters. *American Psychologist, 4*, 399–413.

Bryce, D. (2007). *Investigating the relationship between metacognitive skill and executive functions: A pilot study to assess developmental changes in four to eight-year-olds.* Unpublished MPhil thesis, University of Cambridge, UK.

Bryce, D., & Whitebread, D. (2008). *Investigating the relationship between metacognitive skills and executive functions.* Paper presented at 3rd Biennial Meeting of EARLI Special Interest Group 16: Metacognition, Ioannina, Greece.

Cameron, R. J., & Reynolds, A. R. (Eds.) (1999). *Psychology in education port-folio: Learning style and metacognition*. Berkshire, UK: NFER-NELSON.

Carlson, S. M., Wong, A., Lemke, M., & Cosser, C. (2005). Gesture as a window on children's beginning understanding of false belief. *Child Development, 76*, 73–86.

Chi, M. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Mahwah, NJ: Erlbaum.

Davis, P., & Florian, L. (2004). Teaching strategies and approaches for pupils with special educational needs: A scoping study (Research Report No 516). London, UK: Department for Education and Science.

Deak, G. O., & Narasimham, G. (2003). Is perseveration caused by inhibition failure? Evidence from preschool children's inferences about word meanings. *Journal of Experimental Child Psychology, 86*(3), 194–222.

Deloache, J. S., Sugarman, S., & Brown, A. L. (1985). The development of error correction strategies in young children's manipulative play. *Child Development, 56*, 125–137.

Demetriou, D. (2009). *Inter-relationships in children's socio-cognitive development: The case of false belief understanding.* Unpublished PhD thesis, University of Cambridge, UK.

Demetriou, D., & Whitebread, D. (2008). *Theory of mind and metacognitive knowing: Have we been investigating similar constructs without realising it?* Paper presented at 3rd Biennial Meeting of EARLI Special Interest Group 16: Metacognition, Ioannina, Greece.

Dermitzaki, I., Leondari, A., & Goudas, M. (2009). Relations between young children's strategic behaviours, domain-specific self-concept, and performance in a problem-solving situation. *Learning and Instruction, 19*, 144–157.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality and development.* Philadelphia: Psychology Press.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*(4), 277–287.

Elias, C. L., & Berk, L. E. (2002). Self-regulation in young children: Is there a role for sociodramatic play? *Early Childhood Research Quarterly, 17*(2), 216–238.

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition, 9*, 288–307.

Ferrari, M. (1996). Observing the observer: Self-regulation in the observational learning of motor skills. *Developmental Review, 16*, 203–240.

Fitzsimmons, G. M., & Bargh, J. A. (2004). Automatic self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory and applications* (pp. 151–170). New York: Guilford.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34*, 906–911.

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly, 50*(3), 274–290.

Flavell, J. H., Beach, D. R., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development, 37*, 283–299.

Florian, L. (Ed.). (2007). *The SAGE handbook of special education.* London: Sage.

Flynn, E., O' Malley, C., & Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Developmental Science, 7*(1), 103–115.

Gersten, R., Baker, S., Pugach, M., with Scanlon, D., & Chard, D. (2001). Contemporary research on special education teaching. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 695–748). Washington, DC: AERA.

Goldin-Meadow, S. (2002). Constructing communication by hand. *Cognitive Development, 17*, 1385–1405.

Gopnik, A., & Graf, P. (1988). Knowing how you know: Young children's ability to identify and remember the sources of their beliefs. *Child Development, 59*, 1366–1371.

Grau Cárdenas, V. (2008a). *In the cross-roads of self-regulated learning (SRL) and conceptual understanding: To what extent does it explain individual differences in learning and development?* Paper presented at 3rd Biennial Meeting of EARLI Special Interest Group 16: Metacognition, Ioannina, Greece.

Grau Cárdenas, V. (2008b). *Self-regulated learning and conceptual development in biology: A naturalistic study with primary school children.* Unpublished PhD thesis, University of Cambridge, UK.

Happe, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition, 61*, 25–39.

Hatano, G., & Oura, Y. (2003). Commentary: Reconceptualizing school learning using insight from expertise research. *Educational Researcher, 32*(8), 26–29.

Hirschfeld, L., & Gelman, S. (1994). Toward a topography of mind: An introduction to domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 3–35). Cambridge, UK: Cambridge University Press.

Iiskala, T., Vauras, M., & Lehtinen, E. (2004). Socially shared metacognition in peer-learning? *Hellenic Journal of Psychology, 1*, 147–178.

Inagaki, K., & Hatano, G. (2002). *Young children's naïve thinking about the biological world.* New York: Psychology Press.

Istomina, Z. M. (1975). The development of voluntary memory in preschool age children. *Soviet Psychology, 13*, 5–64.

Karmiloff-Smith, A. (1979). Problem-solving processes in children's construction and representations of closed railway circuits. *Archives de Psychologie, 17*, 33–59.

Karpov, Y. V. (2005). *The neo-Vygotskian approach to child development.* Cambridge, UK: Cambridge University Press.

Kavale, K. (2007). Quantitative research synthesis: Meta-analysis of research on meeting special educational needs. In L. Florian (Ed.), *The SAGE handbook of special education* (pp. 207–221). London: Sage.

Kirschenbaum, D. S. (1984). Self-regulation of sport psychology: Nurturing an emerging symbiosis. *Journal of Sport Psychology, 6*, 159–183.

Kitsantas, A., & Zimmerman, B. J. (2002). Comparing self-regulatory processes among novice, non-expert and expert volleyball players: A microanalytic study. *Journal of Applied Sport Psychology, 14*(2), 91–105.

Kreutzer, M. A., Leonard, S. C., & Flavell, J. H. (1975). An interview study of children's knowledge about memory. *Monographs of the Society for Research in Child Development, 40*(1).

Kuhn, D. (2000). Theory of mind, metacognition and reasoning: A life-span perspective. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 301–326). Hove, UK: Psychology Press.

Lidz, C.S., & Gindis, B. (2003). Dynamic assessment of the evolving cognitive functions in children. In A. Kozulin, B. Gindis, V. A. Ageyev, & S. A. Miller (Eds.), *Vygotsky's educational theory in cultural context* (pp. 99–116). Cambridge, UK: Cambridge University Press.

Limón, M. (2003). The role of domain-specific knowledge in intentional conceptual change. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change* (pp. 133–170). Mahwah, NJ: Erlbaum.

Lind, S. E., & Bowler, D. M. (2009). Recognition memory, self-other source memory, and theory-of-mind in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders* [On-line]. Available: http://www.springerlink.com/content/m5558v5x1n337617/

Lloyd, M., Ried, G., & Bouffard, M. (2006). Self-regulation of sport specific and educational problem-solving tasks by boys with and without DCD. *Adapted Physical Activity Quarterly, 23*(4), 370–389.

Lockl, K., & Schneider, W. (2007). Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development, 78*(1), 148–167.

McCaslin, M., & Hickey, D. T. (2001). Self-regulated learning and academic achievement: A Vygotskian view. In B. J. Zimmerman & D. H. Schunk *Self-regluated learning and academic achievement* (2nd ed., pp. 227–252). Mahwah, NJ: Erlbaum.

Mayr, T., & Ulich, M. (2009). Social-emotional well-being and resilience of children in early childhood settings – PERIK: An empirically based observation scale for practitioners. *Early Years, 29*(1), 45–57.

Meyer, D., & Turner, J. C. (2002). Using instructional discourse analysis to study scaffolding of student self-regulation. *Educational Psychologist, 37*, 17–25.

Naito, M. (2003). The relationship between theory of mind and episodic memory: Evidence for the development of autonoetic consciousness. *Journal of Experimental Child Psychology, 85*, 312–336.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–141). New York: Academic.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist, 36*, 89–101.

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences, 3*, 337–344.

Perry, N. (1998). Young children's self-regulated learning and contexts that support it. *Journal of Educational Psychology, 90*(4), 715–729.

Piaget, J. (1977). *The crasp of consciousness.* London: Routledge & Kegan Paul.

Pine, K. J., Lufkin, N., & Messer, D. (2004). More gestures than answers: Children learning about balance. *Developmental Psychology, 40*(6), 1059–1067.

Pino Pasternak, D., Whitebread, D., & Tolmie, A. (2010). A multi-dimensional analysis of parent-child interactions during academic tasks and their impact on children's self-regulated learning. *Cognition and Instruction, 28*(3)*,* 1–54.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego, CA: Academic.

Polatajko, H. J., & Mandich, A. (2004). *Enabling occupation in children: The Cognitive Orientation to daily Occupational Performance (CO-OP) approach*. Ottawa, ON: Canadian Association of Occupational Therapists Publications.

Ponitz, C. E. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioural regulation in early childhood. *Early Childhood Research Quarterly, 23*, 141–58.

Reder, L. M. (Ed.). (1996). *Implicit memory and metacognition.* Mahwah, NJ: Erlbaum.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23*(1), 90–108.

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at 3–7 years: The Children's Behaviour Questionnaire. *Child Development, 72*(5), 1394–1408.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology, 80*, 201–224.

Saldaña, D. (2004). Dynamic Master Mind: Interactive use of a game for testing metacognition. *School Psychology International*, 25, 422–438.

Salonen, P., Vauras, M., & Efklides, A. (2005). Social interaction – What can it tell us about metacognition and coregulation of learning? *European Psychologist, 10*(3), 199–208.

Sangster, C. A. (2009a). *The role of self-regulation in motor learning: Studying the differences between children with and without motor learning difficulties.* Unpublished PhD thesis, University of Cambridge, UK.
Sangster, C. A. (2009b, August). *The role of self-regulatory and metacognitive skill in motor learning: exploring responses to intervention amongst children with and without motor difficulties.* Paper presented to the 13th EARLI Biennial Conference, Amsterdam, Netherlands.
Sangster, C. A., Beninger, C., Polatajko, H. J., & Mandich, A. (2005). Cognitive strategy generation in children with developmental coordination disorder. *Canadian Journal of Occupational Therapy, 72*, 67–77.
Schneider, W. (1985). Developmental trends in the metamemory-memory behaviour relationship: An intergrative review. In D. L. Forrest-Pressley, G. E. McKinnon, & T. G. Waller (Eds.), *Metacognition, cognition and human performance: Vol. 1 Theoretical perspectives* (pp. 57–109). New York: Academic.
Schneider, W., & Bjorklund, D. F. (1998). Memory. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology: Vol. 2, Cognition, perception and language* (5th ed., pp. 467–521). New York: Wiley.
Siegler, R. S. (1996). *Emerging minds: The processes of change in children's thinking.* Oxford, UK: Oxford University Press.
Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition process in development and learning* (pp. 31–58). Cambridge, UK: Cambridge University Press.
Sinatra, G. M., & Pintrich, P. R. (2003). *Intentional conceptual change.* Mahwah, NJ: Erlbaum.
Strømsø, H. I., Bråten, I. & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension, *Learning and Instruction*, *20*(3), 192–204.
Sugden, D., & Chambers, M. (2005). Models of intervention: Towards an eco-developmental approach. In D. Sugden & M. Chambers (Eds.), *Children with developmental coordination disorder* (pp. 189–211). London: Whurr.
Taylor, M., Esbensen, B. M., & Bennett, R. T. (1994). Children's understanding of knowledge acquisition: The tendency for children to report that they have always known what they have just learnt. *Child Development, 65*, 1581–1604.
Van Hout-Wolters, B. H. A. M. (2000). Asessing active self-directed learning. In R. Simons, J. van der Linden, & T. Duffy (Eds.), *New learning* (pp. 83–101). Dordrecht, The Netherlands: Kluwer.
Veenman, M. V. J. (2005). The assessment of metacognitive skills: What can be learned from multi-method designs? In C. Artelt & B. Moschner (Eds.), *Lernstrategien und Metakognition: Implikationen für Forschung und Praxis* (pp. 75–97). Berlin: Waxmann.
Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition & Learning, 1*, 3–14.
Wellman, H. M. (1985). The origins of metacognition. In D. L. Forrest-Pressley, G. E. McKinnon, & T. G. Waller (Eds.), *Metacognition, cognition and human performance: Vol. 1. Theoretical perspectives* (pp. 1–31). London: Academic.
Wellman, H. M., & Liu, D. (2004). Scaling of theory of mind tasks. *Child Development, 75*, 523–541.
Whitebread, D. (1999). Interactions between children's metacognitive processes, working memory, choice of strategies and performance during problem-solving, *European Journal of Psychology of Education, 14*(4), 489–507.
Whitebread, D., Anderson, H., Coltman, P., Page, C., Pino Pasternak, D., & Mehta, S. (2005). Developing independent learning in the early years. *Education 3–13, 33*, 40–50.
Whitebread, D., Bingham, S., Grau, V., Pino Pasternak, D., & Sangster, C. (2007). Development of metacognition and self-regulated learning in young children: The role of collaborative and peer-assisted learning. *Journal of Cognitive Education and Psychology, 3,* 433–455.

Whitebread, D., Coltman, P., Pino Pasternak, D., Sangster, C., Grau, V., Bingham, S., Almeqdad, Q., & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning, 4,* 63–85.

Whitebread, D., & Pino Pasternak, D. (2010). Metacognition, Self-Regulation & Meta-Knowing. In Littleton, K., Wood, C. & Kleine Staarman, J. (eds) *International Handbook of Psychology in Education* (pp. 673–711). Bingley, UK: Emerald.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In P. Pintrich, M. Boekaerts, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). San Diego, CA: Academic.

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement* (2nd Ed). Mahwah, NJ: Erlbaum.

Chapter 12
Metacognitive Development in Early Childhood: New Questions about Old Assumptions

Kristen E. Lyons and Simona Ghetti

1 Introduction

An effective system must be able to monitor and regulate itself. That is, it must be able to evaluate the current state of the system (and how well the system is progressing towards its goals), and the system must be able to alter its behavior, based on these evaluations, in order to achieve the desired goal more accurately or efficiently. In humans (and perhaps nonhuman species; see Smith, Shields, & Washburn, 2003), monitoring and regulation of cognitive operations are achieved via metacognition (i.e., thinking about thinking; see Flavell, 1979).

When an individual evaluates the current state of his or her basic cognitive operations (e.g., memory, perception, problem solving), metacognitive monitoring is said to occur. When an individual uses the output of his or her monitoring evaluations (in conjunction with the current task demands, e.g., for speed or accuracy), to regulate their basic cognitive operations, metacognitive control is said to occur (Koriat & Goldsmith, 1996; Nelson & Narens, 1990, 1994).

Metacognitive monitoring and control may take many forms, depending on the type of task at hand and the stage of the task. This point may be illustrated by considering the example of a student preparing to take an exam. As the student is preparing for the exam, she will need to evaluate how well she has learned the material (i.e., a judgment of learning; see Schneider, Vise, Lockl, & Nelson, 2000), and direct her time and attention to studying materials that she is struggling with (i.e., allocation of study time; Dunlosky & Connor, 1997). During the exam, she may encounter test items for which she does not immediately know the answer; in this case she will have to assess the likelihood that she will be able to remember the answer (i.e., a feeling-of-knowing judgment; see Butterfield, Nelson, & Peck, 1988; Hart, 1965) to decide if she ought to allot extra time to

K.E. Lyons (✉)
Department of Psychology & Center for Mind and Brain, University of California, Davis, 1 Shields Avenue, Davis, CA 95616
e-mail: kelyons@umn.edu

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_12,

these test items or if her time would be better spent in working on other test items. Finally (if there is a penalty for providing an incorrect response, as is common on many standardized tests), when answering each question, the student must evaluate how certain she is about the likely accuracy of her response (i.e., a confidence judgment; see Roebers, 2002), and decide whether to provide that response (if the level of certainty is sufficiently high) or to refrain from answering the test question (if the level of certainty is low) (Koriat & Goldsmith, 1996). Thus, metacognitive monitoring and control can be thought of as a quality control system, whose function is to ensure that only the most accurate and appropriate output is produced.

For this reason, age-related improvements in children's ability to monitor and regulate their mental operations are widely considered to be a driving force in cognitive development, underlying age-related improvements in accuracy on a wide variety of tasks (Koriat, Goldsmith, Schneider, & Nakash-Dura, 2001; Kuhn & Pease, 2006; Plude, Nelson, & Scholnick, 1998). Accordingly, the development of these skills in childhood has been a central focus of metacognitive research. This work has revealed that critical developments in monitoring and control are observed during middle and late childhood (Ghetti, 2008; Koriat et al., 2001; Lockl & Schneider, 2002; Roebers & Howie, 2003; Son, 2005).

Nevertheless important questions about the emergence and early development of metacognition remain unanswered, as the majority of research has excluded younger children. In large part, this exclusion is due to prevailing views that young children have extremely limited metacognitive skills. This view may be based on findings indicating striking deficits in older children's ability to monitor and regulate their mental activity (Flavell, Green, & Flavell, 1995, 2000; Lockl & Schneider, 2004), which have lead to the logical inference that young children's metacognitive skills may be extremely limited.

However, there is some evidence from a small but compelling body of literature to suggest that young children may be more metacognitively skilled than previously assumed (see also Whitebread et al., this volume). In the following sections we review this evidence. Taken together, these findings raise several intriguing questions about the emergence and early development of metacognition, which are discussed in the latter half of this chapter.

2 Metacognition in the Preschool Years

Although young children are often assumed to have limited metacognitive skills, several studies have provided direct or indirect evidence of metacognitive monitoring and control in preschoolers. In the following sections we review this research. First, we consider the evidence indicative of metacognitive monitoring in early childhood; then we examine the evidence indicative of metacognitive control in early childhood.

2.1 Metacognitive Monitoring in Early Childhood

Findings from several lines of research suggest that young children may be more aware of their ongoing mental activity than previously assumed, including studies of children's ability to monitor their knowledge states (Marazita & Merriman, 2004; Patterson, Cosgrove, & Obrien, 1980), their mental imagery (Estes, 1998), and their memory processes (Cherney, 2003; Cultice, Somerville, & Wellman, 1983). Although some of this research was not originally conducted with the intention of assessing metacognitive monitoring, each offers evidence that young children are capable of reflecting on at least some aspects of their ongoing mental activity. However, the extant literature is sparse and rather disjointed. Thus, in the following paragraphs we discuss each of these lines of research individually, and then provide an integrative discussion.

2.1.1 Monitoring of Knowledge States

Children begin to produce mental verbs referring to knowledge states (e.g., *know*, *think*, *I don't know*) at the end of the second year of life. Although the first appearance of these words in conversation may simply reflect knowledge of the pragmatics surrounding their use, mental verbs are employed to refer specifically to one's own or others' mental states by the third year of life (with *know* and *think* being the most prevalent; Moore, Bryant, & Furrow, 1989; Moore, Furrow, Chiasson, & Patriquin, 1994; Shatz, Wellman, & Silber, 1983), suggesting that children as young as 3 years of age may be able to reflect upon their states of knowledge or lack of knowledge.

More direct evidence of this ability (to evaluate when one does or does not know something) comes from research on metalinguistic monitoring (Marazita & Merriman, 2004). In this research, children were asked to report whether or not they knew the meaning of real or fake words (read aloud by an experimenter) or whether or not they knew the word labels for real and fake objects (depicted in drawings). Results indicate that children as young as 2.5-years-old are capable of discriminating between instances when they know something (e.g., the meaning of a real word spoken aloud to them or the word label for a real object depicted in a line drawing) and instances when they do not know something (e.g., the meaning of a fake word spoken aloud to them or the word label of a fake object depicted in a line drawing; Marazita & Merriman, 2004), suggesting that well before the age of 3 years, children can reflect upon their current knowledge state, discriminating between states of knowledge and ignorance.

However, monitoring of one's knowledge state involves much more than being able to evaluate whether or not one has knowledge. Often, individuals have some knowledge, but this knowledge is insufficient for them to be able to respond accurately (either because their knowledge state is incomplete or unclear). In these instances, individuals must not only reflect upon whether or not they have knowledge, but how good their knowledge is, in a more graded fashion. One could imagine that this may prove to be more challenging than simply evaluating whether or not one does or does not have an answer in mind.

To date, this ability, to evaluate the graded nature of one's knowledge state has typically been assessed using comprehension monitoring tests, which assess children's ability to evaluate how well they understand another speaker's message. Typically, in these tasks children are given a task to complete (e.g., to follow tape-recorded instructions for building a tower out of blocks), and are given prompts which are problematic (i.e., ambiguous or incomplete messages, such as "put the red one on top", when there is more than one red object in the stimuli set), and prompts that are non-problematic (i.e., clear, comprehensive, and unambiguous messages). Comprehension monitoring is assessed by examining whether children's verbal or non-verbal behaviors differ in response to problematic versus non-problematic messages.

Using these procedures, clear evidence of comprehension monitoring has been observed even in very young preschoolers. For example, Patterson et al. (1980) found that preschoolers delayed their responding, spent more time looking at the experimenter, and physically wavered between response options more frequently when they were presented with uninformative messages as compared to when they were presented with informative messages. Children as young as 2.5-years of age have also been found to evince verbal indicators of comprehension monitoring (Revelle, Wellman, & Karabenick, 1985). In this research, children played a game with an experimenter, who requested that the children retrieve items in the playroom. Requests included non-problematic (i.e., clear, unambiguous requests) as well as problematic requests (i.e., inaudible, ambiguous, or impossible requests). Children's statements were coded for an indication that they detected the problematic nature of the request (e.g., asking the experimenter to repeat or clarify his/her message, stating that the request was not feasible). While younger preschoolers only appeared to monitor comprehension failures due to experimenters' requests being inaudible, or impossible (e.g., the item requested was absent or too heavy for the child to lift), older preschoolers additionally detected comprehension failures due to ambiguous messages and memory overloads (i.e., the experimenter asking the children to retrieve a long list of items, which the children could not keep in mind).

The results of this line of research suggest that the ability to monitor comprehension may develop in a step-like process, with children first exhibiting awareness of comprehension failures due to striking problems with the incoming messages (e.g., instructions which are impossible to follow). Later, children develop awareness of more subtle problems in comprehension (e.g., ambiguous messages). Furthermore, the somewhat contrasting results of Patterson et al. (1980) and Revelle et al. (1985) suggest that implicit awareness of these more subtle distinctions may be observed before children become able to note them explicitly. Consistent with this notion are findings from a more recent study indicating that preschoolers evince nonverbal indicators of comprehension monitoring (e.g., quizzical facial expressions) much more frequently than verbal indicators of comprehension monitoring (Skarakis-Doyle, 2002). It should be noted, however, that requests for clarification or repetition of a message involve not only metacognitive monitoring (of one's comprehension) but also metacognitive control (of one's knowledge state by seeking additional information). Thus, although comprehension monitoring is often assessed using these indicators, in such paradigms it is not possible to disentangle the contributions of monitoring and control to children's performance.

2.1.2 Monitoring of Mental Imagery

Another domain in which children's awareness of their ongoing mental activity has been explored is mental imagery. To do so, Estes (1998) asked children to report how they made their judgments on a mental rotation task (Shepard & Metzler, 1971). Children were presented with two images in different spatial orientations (rotated around a center point), and were asked to judge whether the images were the same or different (i.e., "Are these monkeys holding up the same arm or different arms?"). In order to make the judgment, children had to mentally represent one of the images in their minds, mentally rotate the image, and compare that (rotated) mental image to the target object.

When preschoolers were asked how they made their judgments, 40% of 4-year-olds and 56% of 5-year-olds referred to mental activity in their responses. Critically, the behavioral response pattern of these children corroborated their reports (i.e., children who provided "mental" explanations evinced a pattern of responding consistent with actual use of a mental rotation strategy: judgment reaction times increased with the angle of rotation). Children who did not provide mental explanations for how they performed the task did not exhibit this pattern, suggesting that children who referred to their mental activity seemed to be reliably reporting on their mental imagery – for additional indications of early awareness of mental imagery see, also, Estes, Wellman, and Woolley (1989) and Rieser, Garing, and Young (1994). These results provide a straightforward and compelling indication that children, as young as 4 years of age, are capable of metacognitively monitoring the contents of their mental imagery.

2.1.3 Memory Monitoring

Evidence of metacognitive monitoring in young children also comes from studies of preschoolers' ability to monitor their memory operations. Although relatively few studies have examined memory monitoring in young children, the studies that have done so have provided compelling evidence that young children are capable of reflecting on several dimensions of the retrieval process. For example, Cultice et al. (1983) investigated preschoolers' ability to provide feeling-of-knowing judgments (i.e., judging whether they would be able to recognize the names of schoolmates, pictured in photographs, whose names they could not currently recall). Results indicated that participants' feeling-of-knowing judgments corresponded to their subsequent memory performance. Given that, in adults, feeling-of-knowing judgments can arise as a function of the familiarity of the test cues (Koriat & Levy-Sadot, 2001), these results suggest that children as young as 4-years-old may be capable of introspecting on their sense of familiarity for test materials.

More recently, Cherney (2003) examined the extent to which 3- and 5-year-olds' spontaneous utterances of mental terms indicating subjective certainty (e.g., *know*, *remember*, *forget*) and subjective uncertainty (e.g., *think*, *guess*, *bet*) corresponded to the accuracy of their memory on a spatial location recall task. Although mental terms were uttered infrequently (uttered by about 20% of 3-year-olds and 30% of 5-year-olds), children's utterances were moderately likely to be consistent with their memory performance (e.g., saying "I know" and then providing a correct response or saying "I forget" and then providing an incorrect response). Correspondence was better for older compared to younger children and was better for terms expressing certainty compared to uncertainty. This latter finding is consistent with research in older children suggesting that children may gain mastery of certainty monitoring earlier than they master uncertainty monitoring (Roebers, von der Linden, & Howie, 2007), suggesting relative continuity in development of metacognitive functioning over the course of childhood.

Typically, monitoring of subjective certainty is indexed using confidence judgments (i.e., a scale-based rating of how confident one is about the likely accuracy of one's response). These judgments are among the most widely used indices of metacognitive monitoring in research with older children and adults (e.g., Roebers, 2002) and nonhuman animals (Shields, Smith, Guttmannova, & Washburn, 2005; see also, Beran, Couchman, Coutinho, Boomer, & Smith, this volume). This approach to the study of metacognitive monitoring has at least two major advantages. First, confidence judgments can be collected for a wide variety of decisions (e.g., perceptual decisions, memory decisions), and second, collecting confidence ratings in conjunction with decisions affords the opportunity to examine the degree to which individuals' *reports* of their subjective mental experiences (i.e., their confidence judgments) correspond to their *actual* mental experiences (i.e., their memory decisions), providing an objective indicator of monitoring ability (see also, Allwood, this volume).

If confidence judgments associated with accurate responses are higher than confidence judgments associated with inaccurate responses, metacognitive monitoring is inferred. The logic of this approach is that, if an individual is reflecting on their subjective sense of certainty about the likely accuracy of their decisions, they should, on average, report higher confidence for correct compared to incorrect responses. This pattern of results has been observed on memory tasks in children as young as 5 years of age (Ghetti, Qin, & Goodman, 2002). Five-year-olds also report higher confidence ratings on memory tests when recognizing words that have been studied with a picture compared to words that have been studied without a picture (Ghetti et al., 2002) and report that they remember more details about highly memorable life events compared to low memorability life events (Ghetti & Castelli, 2006), offering evidence that young children are capable of evaluating the strength of their memory representations at retrieval.

In sum, the results of several investigations of memory monitoring in young children strongly suggest that children as young as 4 years of age can reflect on critical aspects of their memory operations at retrieval. These dimensions include the feeling of knowing that one will be able to remember something that cannot currently be retrieved, and the strength and likely accuracy of one's memory representations.

2.1.4 Conclusion

Taken together, these findings suggest that young children may be much more adept at monitoring their mental activity than is often assumed. The extant literature indicates that children as young as 30 months of age can evaluate whether they do or do not know something (Marazita & Merriman, 2004; Moore et al., 1989; Moore et al., 1994). While this ability may be rudimentary, it arguably provides a foundation for subsequent developments in metacognitive monitoring, which are observed during the preschool years.

Between the ages of 3 and 5 years, children demonstrate age-related improvements in their ability to monitor more fine-grained distinctions among their knowledge states. Children's ability to monitor their comprehension improves during this period (Cosgrove & Patterson, 1977; Pratt & Bates, 1982), as does their ability to introspect on their mental imagery (Estes, 1998) and monitor the vividness and detail of their memory representations (Ghetti & Castelli, 2006; Ghetti et al., 2002). Furthermore, there is initial evidence that children's ability to reflect upon their sense of certainty about the accuracy of their memory states improves during this period (Cherney, 2003) as well, suggesting that during the preschool years children may develop the ability to reflect upon cues to the subjective experience of high and low certainty (e.g., how fluently one processes information or retrieves a response, how much time is required to process information or retrieve a response). Consistent with this notion are recent findings indicating that between the ages of 3 and 5 years, preschoolers' confidence judgments increasingly differentiate between accurate and inaccurate perceptual decisions (although even 3-year-olds report higher confidence for accurate vs. inaccurate responses), and that during this period uncertainty monitoring ability begins to extend to additional types of decision making–for example, linguistic decisions (Lyons & Ghetti, in press).

There are, of course, limitations in young children's monitoring abilities, perhaps most strikingly demonstrated by Flavell et al. (1995, 2000) who documented that introspecting on the contents of one's thoughts poses great challenges for 5-year-olds. Nevertheless, the results of several studies provide compelling evidence that young children are in many ways quite capable of monitoring their ongoing mental activity. This ability is critical as it provides the foundation for future action aimed at improving one's knowledge or the accuracy of one's performance. Of course, this benefit can only occur to the extent that children are able to act on the basis of monitoring evaluations to metacognitively control their cognitive activity. Several lines of research suggest that this ability may be observed even in very young children.

3 Metacognitive Control in Early Childhood

Relatively little research has directly examined early self-regulation from a metacognitive perspective. However, findings from several lines of research suggest that young children readily engage in specific behaviors aimed at improving the precision of their understanding or the accuracy their performance, including studies of children's regulation of their knowledge states, children's use of rudimentary memory strategies, and children's control of goal-directed activity. This work has been conducted in a rather fragmented fashion; therefore, in the following sections we review each of these lines of research individually and then provide an integrative discussion.

3.1 Control of Knowledge States

To improve their knowledge states, children as young as 12 months of age seek information from adults (through expressions and gestures), and preschoolers ask an average of more than 70 information-seeking questions per hour; it has been argued that children ask these questions when they are faced with an experience of uncertainty (Chouinard, 2007), although this proposal has yet to be tested empirically. Furthermore, recent research indicates that even 3-year-olds seek information selectively and are more likely to trust individuals who were previously correct compared to individuals who declared their ignorance (Koenig & Harris, 2005), demonstrating that preschoolers actively seek information from reliable sources in order to improve their knowledge.

The results of comprehension monitoring studies offer more direct evidence that young children act deliberately to overcome limitations in their knowledge states. For example, even young preschoolers can be trained to request clarification in response to ambiguous messages to improve the accuracy of their performance (Pratt & Bates, 1982). However, older preschoolers respond to comprehension failures in a more problem-focused way (e.g., requesting clarification or elaboration that would be helpful in resolving the failure) to a greater extent than younger children (Revelle et al., 1985; see also Walters & Chapman, 2000), suggesting that the ability to overcome limitations in one's current knowledge state by seeking information does appear to improve with age. Taken together, these findings suggest that young children actively seek additional information from reliable sources in order to improve the accuracy and clarity of their states of knowledge and understanding.

3.2 Control of Memory Performance

The majority of research on metacognitive regulation of performance in early childhood has been conducted in the context of memory tasks (Acredolo, Pick, &

Olsen, 1975; Deloache & Brown, 1983). Several studies have found that very young children selectively engage in "studying" behaviors in the context of a memory task (but not on parallel tasks in which the memory demands are removed). For example, Wellman, Ritter, and Flavell (1975) found that 3-year-old children spent more time looking at and pointing at the location of a hidden object when they were told to "remember" the location of a toy hidden under one of several opaque cups than when children were simply told to "wait" during the delay period; the amount of time spent engaged in these behaviors positively correlated with memory performance.

Although younger children (2-year-olds) failed to demonstrate such an effect in Wellman et al. (1975), a similar pattern of selective "studying" behavior was observed in toddlers as young as 18 months of age, when the task involved searching for a stuffed toy hidden in one of several naturalistic hiding places (e.g., under a pillow, behind a chair) (Deloache, Cassidy, & Brown, 1985). Again, these behaviors (e.g., talking about the toy and/or its location) were associated with subsequent memory performance and were rarely observed when the memory demands were removed (e.g., when the child was told that the experimenter would search for the toy at the end of the delay, or the child was told that their task would be to "wake up" the toy, which was placed in plain sight, after the delay). Furthermore, these behaviors were more readily engaged in when the task was performed in a novel (laboratory) setting than a familiar (home) setting, offering a preliminary indication that these behaviors are under strategic control, and engaged in more readily when children are less certain about their performance ability (as would be the case in a novel setting). Experimental research corroborates this interpretation, as similar increases in studying behaviors have been documented as a result of incentive manipulations (i.e., preschoolers engage in more studying behaviors when the incentives for accurate performance are increased; see O'Sullivan, 1993).

Further evidence of control has been obtained by examining children's behaviors at retrieval. For example, Deloache and Brown (1984, 1987) observed that at retrieval, upon failure to locate a hidden toy (either because children forgot the location or because the toy was surreptitiously moved), children as young as 2 years of age directed their subsequent search to related locations (i.e., adjacent locations or analogous locations such as behind the pillow on the opposite side of a couch). Moreover, children were more persistent in searching when the toy had been surreptitiously moved (i.e., when they should be highly confident in their memory for the location), compared to when they forgot the location (i.e., when they should be less confident in their memory for the location). Taken together, these results suggest that, when faced with retrieval failure, children continue to act in a strategic fashion in order to produce optimal performance.

More recent findings indicate that preschoolers also strategically regulate the accuracy of their performance by withholding inaccurate information, when they are given the option to refrain from responding. For example, in an eyewitness memory interview study, Mulder and Vrij (1996) found that 4- to 5-year-old children were more accurate (i.e., provided fewer incorrect responses) when they were instructed to say "I don't know" when they were uncertain about their answers. Impressively, the advantage observed in this age group was similar to that of older children (8- to 10-year-olds). In another study, 3-year-olds evinced higher accuracy rates on a paired-associates memory test when they were given the option to skip trials, compared to when they were required to provide an answer on all trials (Balcomb & Gerken, 2008). However, clear limitations in metacognitive control were observed, as more than one-third of the children never used the opt-out option or used the opt-out response on all trials (with the former being significantly more common, exhibited by 89% of the children who consistently used or failed to use the opt out response).

This finding is consistent with the well-documented challenges that young children evince with inhibitory control (Gerstdat, Hong, & Diamond, 1994; Reed, Pien, & Rothbart, 1984). From this perspective, young children's challenges with metacognitively regulating their accuracy may be due to limitations in executive function, rather than an inability to act on the basis of monitoring evaluations to guide strategic responding. Extant procedures have thus far been unable distinguish between the contributions of metacognitive control and inhibitory control to preschoolers' performance on these tasks. Although research on the early development of metacognitive control has largely been conducted in parallel to research on the development of cognitive control and executive function during the preschool years, these domains are conceptually, theoretically, and developmentally intertwined (Fernandez-Duque, Baird, & Posner, 2000; Shimamura, 2000; see also Whitebread et al., this volume). Thus, the early development of metacognitive control needs to be examined while accounting for global developments in executive function and self-regulation skills in preschoolers (Carlson, 2005).

Following this approach, a recent study explored the relation between individual differences in inhibitory control and preschoolers' ability to take advantage of a withholding strategy to improve the accuracy of their performance on a perceptual-identification task (Lyons & Ghetti, 2010). Replicating previous findings, preschoolers across age groups evinced higher accuracy rates when the option to withhold was present compared to when it was not. Critically, participants reported higher confidence in association with responses that were provided rather than withheld, offering clear evidence that metacognitive control was influenced by the results of metacognitive monitoring evaluations. Further evidence consistent with this notion was the finding that individual differences in uncertainty monitoring ability positively predicted regulation of accuracy (via withholding of incorrect responses).

However, an additional path to metacognitive control was observed. Specifically, inhibitory control and uncertainty monitoring ability were independent positive predictors of accuracy improvements (due to strategic withholding). These findings are notable because they suggest that control of accuracy may be achieved if children approach tasks cautiously, treading carefully when there are any cues to indicate that the task at hand may be difficult to perform or that the risk of making an error is high (independently of uncertainty monitoring ability).

3.3 Self-Directed Metacognitive Control

Historically, the topic of self-directed control has received substantial attention in the developmental literature, at least in part stemming from the Vygotskian view that young children talk to themselves as a form of early self-regulation (Manning, White, & Daugherty, 1994). Consistent with this view is research from naturalistic settings (namely, preschools) indicating that older preschoolers engage in self-talk in a systematic way, for example, more frequently doing so while they are working alone than while in groups or working with adults, and more frequently when they are engaged in focused and prolonged goal-directed activities. While self-talk is also observed in younger preschoolers, it is not systematically associated with the same situations as it is for older preschoolers, suggesting improvement in the ability to self-regulate one's thoughts and actions during the preschool years (Winsler, Carlton, & Barry, 2000).

More recent evidence indicates that in addition to verbal self-regulation, preschoolers also evince nonverbal metacognitive regulation, including error detection and checking (e.g., that one is doing a task appropriately), redirection of activities (e.g., directing attention back to the main task), help-seeking (e.g., from peers or adults), and transferring previously successful strategies to new situations (Dermitzaki, Leondari, & Goudas, 2009; Whitebread, Bingham, Grau, Pino Pasternak, & Sangster, 2007). These behaviors are also more readily engaged in when children work more independently and are less frequently observed when adults are involved in interactions (Whitebread et al., 2007), suggesting that these behaviors may be under conscious control and more likely to be engaged in when the need for self-regulation is greater.

3.4 Conclusion

In sum, several studies have documented metacognitive control (or metacognitive control-like) behaviors in preschoolers. Children begin to seek information from adults around their first birthday (through gestures and expressions; Chouinard, 2007). During the preschool years children's ability to metacognitively control their knowledge states through information-seeking develops substantially, as they become increasingly likely to request information from reliable sources (Koenig & Harris, 2005), and their questions become more problem-focused, and directed towards gaining information that will be helpful in improving their knowledge and understanding (Revelle et al., 1985).

In addition to these strategic actions aimed at improving their knowledge states, preschoolers also act to improve the accuracy of their performance, engaging in deliberate strategies to improve memory retention (Deloache et al., 1985; Wellman et al., 1975), and retrieval accuracy (Deloache & Brown, 1984, 1987). These rudimentary strategies are observed very early in life (well before the preschool years, under some conditions; Deloache et al., 1985); however, the ability to engage in more advanced forms of strategic control of performance (e.g., through the strategic withholding of incorrect responses) improves with age during the preschool years (Balcomb & Gerken, 2008; Lyons & Ghetti, 2010; Mulder & Vrij, 1996), although there is evidence that metacognitive monitoring evaluations guide strategic responding even in young preschoolers. These age improvements in strategic control of performance accuracy develop in parallel with age improvements in children's ability to use verbal and nonverbal strategies for self-regulation when engaged in goal-directed behaviors (Whitebread et al., 2007; Winsler et al., 2000), consistent with the well-documented age-increase executive function observed during the preschool years (Carlson, 2005).

4 Towards a Comprehensive Model of Early Metacognitive Development

Overall, the findings reviewed above provide compelling evidence that critical milestones in metacognitive monitoring and control are achieved in early childhood. These early abilities likely provide the foundation for learning in a host of domains as well as subsequent metacognitive development. Nevertheless young children have exhibited some clear limitations in their ability to monitor their mental activity (e.g., being greatly challenged at reporting about what they were thinking about only a moment earlier; Flavell et al., 1995, 2000), as well as in their ability to strategically regulate their mental activity and behavior (e.g., in directing their attention during study only to items which they will be tested about later; Miller, 1990), underscoring the complex nature of early metacognitive development.

Thus, although important (but often unrecognized) steps have been made towards developing a comprehensive understanding of early metacognitive development, several fundamental questions remain unanswered. In the following sections we raise some of these questions about the early development of metacognition and offer some thoughts on how these issues may be resolved. Answering these questions will elucidate the mechanisms underlying early metacognitive development, and will help lay the foundation for a comprehensive model of the emergence of metacognition in early childhood.

4.1 How Do Metacognitive Monitoring and Control Develop in Early Childhood?

An important first step in developing a comprehensive model of early metacognitive development will be to characterize the exact mechanisms through which metacognitive monitoring and control develop. It has been proposed that awareness and regulation of mental activity may progress through a series of increasingly complex levels of self-reflection. For example, Zelazo and colleagues (Zelazo, 2004) have proposed an information processing account of the development of conscious awareness. This model contends that at birth children are minimally conscious of their world; they are aware only of the stimulus which they are currently encountering and whether this stimulus gives rise to a pleasurable or a negative feeling (i.e., “minimal consciousness”). Around the first birthday, infants achieve the ability to bring back to mind stimuli which are no longer in the environment and consider them in relation to one another (i.e., “recursive consciousness”); that is to say, they are able to bring to mind and reflect upon the contents of “minimal consciousness”. Around the second year, children achieve a third level of conscious awareness (i.e., “self-consciousness”), at which they are able to bring to mind the contents of “recursive consciousness” and are able to explicitly reflect on them. Additional levels of conscious awareness are achieved as children progress through further iterative recursions, bringing to mind and reflecting upon the contents of their mental activity from lower levels of consciousness. It is proposed that children’s ability to control their behavior increases as a function of achievements in the highest level of reflection that children can engage in (Zelazo, Gao, & Todd, 2007). Support for this proposal comes from a series of experiments demonstrating that in the first several years of life, children are increasingly able to follow complex embedded rules (e.g., on the dimensional change card sort task, “If we are playing the color game the red truck goes with the red cards, but if we are playing the shape game, the red truck goes with the truck cards”) (Zelazo, 2004). Although a number of experiments have provided evidence in support of components of the model, to date, the model has not been tested in its entirety within a single study.

Similarly, Flavell (2003) speculated that children may develop awareness of uncertainty in a four-stage process. At the first stage of development (i.e., at birth), infants may not have any experience of uncertainty. During this period, children would not exhibit any behavioral differences in the face of certain or uncertain situations. At the second stage of development, children may have a subjective experience of uncertainty but fail to be consciously aware of it. During this period, children may have slower reaction times when asked to predict the outcome of uncertain compared to certain situations, but they would not report feeling any differently about responding to the two kinds of prompts.

At the third stage of development, children may be consciously aware of the subjective experience of uncertainty, but may not attribute it as such. Although for adults and older children, the subjective experience of uncertainty is readily identified as such, one could imagine that when awareness of these feelings first emerges, young children may not identify them as indicating uncertainty. Even adults sometimes misattribute the sources of their subjective experiences; for instance, they may misattribute their arousal (actually resulting from experimental manipulations) to feelings of romantic attraction (Dutton & Aron, 1974; White, Fishbein, & Rutstein, 1981). Moreover, there is evidence for young children’s wishful thinking and biases towards overconfidence (Schneider, 1998). Therefore, it seems probable that young children may progress through a phase in which they may experience the feelings associated with subjective uncertainty but they may not associate these feelings with uncertainty. However, as children gain more experience (perhaps as they learn to associate their subjective feelings of uncertainty with instances in which they produced the wrong answer or response), they would progress to the final stage of uncertainty monitoring, at which they are consciously aware of their subjective experience of uncertainty and recognize it as uncertainty. This proposal has yet to be empirically tested.

To fully characterize the emergence and early development (and indeed development throughout the lifespan) of metacognition, future research should examine how the ability to monitor and regulate mental activity develops over time and across domains, using procedures that can be used with individuals from a wide age-span and across tasks. Developing such procedures may be challenging, but it would prove invaluable in helping to elucidate the reasons why young children are able to reflect on some types of mental activity (e.g., images, word knowledge; Estes, 1998; Marazita & Merriman, 2004) but not others (e.g., thoughts; Flavell et al., 1995, 2000), as well as why young children are capable of engaging in some aspects of metacognitive control (e.g., strategic withholding; Balcomb & Gerken, 2008), but not others (e.g., selectively attending to relevant information during encoding; Miller, 1990).

4.2 What Is the Relation Between Monitoring and Control in Early Childhood?

Research with young children to date has established a set of conditions under which young children can monitor their thinking and control their performance. However, it remains unknown to what extent young children's metacognitive control operations are guided by the results of monitoring evaluations. What level of awareness is required for children to be able to control their behavior? Do children act on the basis of implicit metacognition (e.g., asking questions to resolve ambiguity, despite not being explicitly aware that they are uncertain)? If so, how should this level of "awareness" be characterized, and how does acting on the basis of such monitoring influence awareness?

The question of whether monitoring evaluations guide control processes, and whether control can occur in the absence of monitoring has been debated in recent years in the adult metacognitive literature. Although traditional models of metacognition hold that control occurs as a consequence of monitoring, recent findings (from the adult behavioral and patient literature) suggest that control may occur in the absence of conscious monitoring (Moulin, Perfect, & Fitch, 2002), and that monitoring may even occur under some circumstances as a consequence of control operations (Koriat, Ma'ayan, & Nussinson, 2006).

Studying the emergence and early development of these processes may provide critical insight into the nature of the relation between metacognitive monitoring and metacognitive control. Recent findings suggest that children as young as 3-years-old engage in control operations on the basis of monitoring evaluations of certainty (refraining from responding when their confidence is low and providing responses when their confidence is high; Lyons & Ghetti, 2010). However, there may be a period earlier in development in which a dissociation is observed between monitoring and control (such that young children are capable of metacognitive control before they are capable of consciously reflecting on their ongoing mental activity), which would raise questions about how control is achieved in the absence of monitoring. One possibility, certainly worth investigating, is that children may proceed more cautiously when environmental cues suggest that the risk of making an error is high (e.g., if a task is unfamiliar).

4.3 Do the Monitoring and Control Processes Differ Across Domains in Early Childhood?

A comprehensive understanding of the early development of metacognition, must include a description of (and account for) the domain specificity or domain generality of metacognitive operations across cognitive functions. That is to say, are monitoring and control skills similar across domains in young children? Do young children develop the ability to monitor and regulate their mental states similarly for different kinds of cognitive functions (e.g., memory, perception, language) and/or for different aspects of cognitive functioning (e.g., judgments of learning, confidence ratings, feelings of knowing)?

The complex patterns of research findings to date suggest that the development of metacognitive monitoring and control are unlikely to develop uniformly across domains. Instead, it seems that children may develop the ability to monitor certain kinds of mental content (e.g., monitoring of one's mental imagery; see Estes, 1998; see also Lyons & Ghetti, in press) before they develop the ability to monitor other forms of mental content (e.g., monitoring of thoughts in one's stream of consciousness; Flavell et al., 1995, 2000). Similarly, children may develop the ability to engage in some forms of metacognitive regulation earlier than others–for example, strategies aimed at preventing forgetting (Wellman et al., 1975) may precede strategies aimed at remembering specific items (Miller, 1990). By characterizing the development of monitoring and control processes across domains, and identifying the characteristics of monitoring and control processes which emerge earlier and later in development, future research could provide critical insight into the mechanisms underlying age-related improvements in metacognition. Such information would be invaluable in helping to develop a comprehensive model of metacognitive development, and may provide insight into the nature of metacognitive processes generally.

4.4 What Are the Neural Bases of Monitoring and Control and How Do They Develop in Early Childhood?

A comprehensive model of early metacognitive development must take into account the neural underpinnings of monitoring and control, and how these change with age and experience. To date, very few studies have examined the neural correlates of procedural metacognition in young children. However, there is a growing body of cognitive neuroscience research (from patient and neuroimaging studies) which provides some insight into the neural bases of metacognitive operations in adults and older children.

There is good evidence, from several lines of research, that monitoring and regulation are supported by the prefrontal cortex (Bunge & Zelazo, 2006; Shimamura, 2000) and the anterior cingulate cortex (Fernandez-Duque et al., 2000). With regard to the former, studies with confabulating patients indicate that damage to ventromedial cortex causes impairments in individuals' ability to monitor the output of their retrieval (for accuracy), resulting in confabulation, that is, the reporting of false information (Schnider, 2003). Patients with prefrontal lobe lesions also exhibit deficits in the feeling-of-knowing evaluations (Shimamura & Squire, 1986), source monitoring (Janowsky, Shimamura, & Squire, 1989), and feelings of subjective recollection (Ciaramelli & Ghetti, 2007). Research with healthy adults also indicates that the prefrontal cortex is involved in monitoring of retrieved information on memory tasks (Ranganath, Johnson, & D'Esposito, 2000; Ranganath & Paller, 2000) and assessments of feeling of knowing (Paynter, Reder, & Kieffaber, 2009). Neuroimaging studies of executive functioning in typically developing adults and older children have also identified midfrontal circuits (involving the frontal cortex and the anterior cingulate cortex) as supporting error monitoring and cognitive control (Fernandez-Duque et al., 2000).

Taken together, these findings suggest that age-related changes in the structural and functional maturation of the frontal lobes and anterior cingulate cortex may underlie age-related improvements in monitoring and control. Integrating neuroimaging methods with behavioral assessments of metacognitive development in early childhood in future research will provide an important foundation for developing a model of early metacognitive development.

5 General Conclusions

An effective system must be able to monitor and regulate itself. Although young children are often assumed to have extremely limited metacognitive skills, there is good evidence to indicate that young children may be more metacognitively skilled than previously assumed. Nevertheless, the extant research is fraught with apparently discrepant findings, with young children demonstrating competence in some domains of metacognitive monitoring and control, while exhibiting striking limitations in others. These seemingly contradictory findings may be resolved by future studies addressing several remaining fundamental questions about the origins and early development of metacognition. However, the extant literature provides good evidence that critical achievements in metacognition are achieved in early childhood, providing the foundation for learning in a variety of domains.

References

Acredolo, L. P., Pick, H. L., & Olsen, M. G. (1975). Environmental differentiation and familiarity as determinants of children's memory for spatial location. *Developmental Psychology, 11*, 495–501.

Balcomb, F., & Gerken, L. (2008). Three-year-old children can access their own memory to guide responses on a visual matching task. *Developmental Science, 11*, 750–760.

Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science, 15*, 118–121.

Butterfield, E. C., Nelson, T. O., & Peck, V. (1988). Developmental aspects of the feeling of knowing. *Developmental Psychology, 24*, 654–663.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*, 595–616.

Cherney, I. D. (2003). Young children's spontaneous utterances of mental terms and the accuracy of their memory behaviors: A different methodological approach. *Infant and Child Development, 12*, 89–105.

Chouinard, M. M. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development, 72*(1, Serial No. 286), 1–129.

Ciaramelli, E., & Ghetti, S. (2007). What are confabulators' memories made of? A study of subjective and objective measures of recollection in confabulation. *Neuropsychologia, 45*, 1489–1500.

Cosgrove, J. M.., & Patterson, C. J. (1977). Plans and the development of listener skills. *Developmental Psychology, 13*, 557–564.

Cultice, J. C., Somerville, S. C., & Wellman, H. M. (1983). Preschoolers' memory monitoring: Feeling-of-knowing judgments. *Child Development, 54*, 1480–1486.

DeLoache, J. S., & Brown, A. L. (1983). Very young children's memory for the location of objects in a large-scale environment. *Child Development, 54*, 888–897.

DeLoache, J. S., & Brown, A. L. (1984). Where do I go next? Intelligent searching by very young children. *Developmental Psychology, 20*, 37–44.

DeLoache, J. S., & Brown, A. L. (1987). Differences in the memory-based searching of delayed and normally developing young children. *Intelligence, 11*, 277–289.

DeLoache, J. S., Cassidy, D. J., & Brown, A. L. (1985). Precursors of mnemonic strategies in very young children's memory. *Child Development, 56*, 125–137.

Dermitzaki, I., Leondari, A., & Goudas, M. (2009). Relations between young students' strategic behaviours, domain-specific self-concept, and performance in problem-solving situation. *Learning and Instruction, 19*, 144–157.

Dunlosky, J., & Connor, L. T. (1997). Age differences in the allocation of study time account for age differences in memory performance. *Memory & Cognition, 25*, 691–700.

Dutton, D. G., & Aron, A. P. (1974). Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of Personality and Social Psychology, 30*, 510–517.

Estes, D. (1998). Young children's awareness of their mental activity: The case of mental rotation. *Child Development, 69*, 1345–1360.

Estes, D., Wellman, H. M., & Woolley, J. D. (1989). Children's understanding of mental phenomena. In H. W. Reese (Ed.), *Advances in child development and behavior* (pp. 41–87). San Diego, CA: Academic.

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition, 9*, 288–307.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911.

Flavell, J. H. (2003). Varieties of uncertainty monitoring. *Brain and Behavioral Sciences, 26*, 344–344.

Flavell, J. H., Green, F. L., & Flavell, E. R. (1995). Young children's knowledge about thinking. *Monographs of the Society for Research in Child Development, 60*(1, Serial No. 243), 1–96.

Flavell, J. H., Green, F. L., & Flavell, E. R. (2000). Development of children's awareness of their own thoughts. *Journal of Cognition and Development, 1*, 97–112.

Gerstdat, C. L., Hong, J. Y., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3½ to 7 years old on a Stroop-like day/night task. *Cognition, 53*, 129–153.

Ghetti, S. (2008). Rejection of false events in childhood: A metamemory account. *Current Directions in Psychological Science, 17*, 16–20.

Ghetti, S., & Castelli, P. (2006). Developmental differences in false-event rejection: Effects of memorability-based warnings. *Memory Special Issue: Memory Editing Mechanisms, 14*, 762–776.

Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology, 38*, 705–718.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology, 56*, 208–216.

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Memory and metamemory: Comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology, 17*, 3–11.

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*, 1261–1277.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490–517.

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 34–53.

Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children's testimony: Can children control the accuracy of their memory reports? *Journal of Experimental Child Psychology, 79*, 405–437.

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General, 135*, 36–69.

Kuhn, D., & Pease, M. (2006). Do children and adults learn differently? *Journal of Cognition and Development, 7*, 279–293.

Lockl, K., & Schneider, W. (2002). Developmental trends in children's feeling-of-knowing judgments. *International Journal of Behavioral Development, 26*, 327–333.

Lockl, K., & Schneider, W. (2004). The effects of incentives and instructions on children's allocation of study time. *European Journal Developmental Psychology, 1*, 53–169.

Lyons, K. E., & Ghetti, S. (2010). "I don't want to pick!": Metacognitive monitoring and inhibitory control underlie early strategic behavior. Manuscript under review.

Lyons, K. E., & Ghetti, S. (in press). The development of uncertainty monitoring in early childhood. *Child development*.

Manning, B. H., White, C. S., & Daugherty, M. (1994). Young children's private speech as a precursor to metacognitive strategy use during task engagement. *Discourse Processes, 17*, 191–211.

Marazita, J. M., & Merriman, W. E. (2004). Young children's judgment of whether they know names for objects: The metalinguistic ability it reflects and the processes it involves. *Journal of Memory and Language, 51*, 458–472.

Miller, P. H. (1990). The development of strategies of selective attention. In D. F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 157–184). Hillsdale, NJ: Erlbaum.
Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development, 60*, 167–171.
Moore, C., Furrow, D., Chiasson, L., & Patriquin, M. (1994). Developmental relationships between production and comprehension of mental terms. *First Language, 14*, 1–17.
Moulin, C. J. A., Perfect, T. J., & Fitch, F. (2002). Metacognitive processes at encoding. In P. Chambers, M. Izaute, & P. Marescaux (Eds.), *Metacognition: Process, function, and use* (pp. 35–48). Dordrecht, The Netherlands: Kluwer.
Mulder, M. R., & Vrij, A. (1996). Explaining conversation rules to children: An intervention study to facilitate children's accurate responses. *Child Abuse & Neglect, 20*, 623–631.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation, 26*, 125–173.
Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: The MIT Press.
O'Sullivan, J. T. (1993). Preschoolers' beliefs about effort, incentives, and recall. *Journal of Experimental Child Psychology, 55*, 396–414.
Patterson, C. J., Cosgrove, J. M., & O'Brien, R. G. (1980). Nonverbal indicants of comprehension and noncomprehension in children. *Developmental Psychology, 16*, 38–48.
Paynter, C. A., Reder, L. M., & Kieffaber, P. D. (2009). Knowing before we know: ERP correlates of feeling-of-knowing. *Neuropsychologia, 47*, 796–803.
Plude, D. J., Nelson, T. O., & Scholnick, E. K. (1998). Analytical research on developmental aspects of metamemory. *European Journal of Psychology of Education, 13*, 29–42.
Pratt, M. W., & Bates, K. R. (1982). Young editors: Preschoolers' evaluation and production of ambiguous messages. *Developmental Psychology, 18*, 30–42.
Ranganath, C., & Paller, K. A. (2000). Neural correlates of memory retrieval and evaluation. *Cognitive Brain Research*, *9*, 209–222.
Ranganath, C., Johnson, M. K., & D'Esposito, M. (2000). Left anterior prefrontal activation increases with demands to recall specific perceptual information. *Journal of Neuroscience*, *20*, 108.
Reed, M. A., Pien, D. L., & Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill-Palmer Quarterly*, 30, 131–147.
Revelle, G. L., Wellman, H. M., & Karabenick, J. D. (1985). Comprehension monitoring in preschool children. *Child Development, 56*, 654–663.
Rieser, J. J., Garing, A. E., & Young, M. F. (1994). Imagery, action, and young children's spatial orientation: It's not being there that counts. It's what one has in mind. *Child Development, 65*, 1262–1278.
Roebers, C. M. (2002). Confidence judgments in children's and adults' event recall and suggestibility. *Developmental Psychology, 38*, 1052–1067.
Roebers, C. M., & Howie, P. (2003). Confidence judgments in event recall: Developmental progression in the impact of question format. *Journal of Experimental Child Psychology, 85*, 352–371.
Roebers, C. M., von der Linden, N., & Howie, P. (2007). Favourable and unfavourable conditions for children's confidence judgments. *British Journal of Developmental Psychology, 25*, 109–134.
Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science, 1*, 291–297.
Schneider, W., Vise, M., Lockl, K., & Nelson, T. O. (2000). Developmental trends in children's memory monitoring: Evidence from a judgment-of-learning task. *Cognitive Development, 15*, 115–134.
Schnider, A. (2003). Spontaneous confabulation and the adaptation of thought to ongoing reality. *Nature Reviews Neuroscience, 4*, 662–671
Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition, 14*, 301–321.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701–703.

Shields, W. E., Smith, J. D., Guttmannova, K., & Washburn, D. A. (2005). Confidence judgments by humans and rhesus monkeys. *Journal of General Psychology, 132*, 165–186.

Shimamura, A. P. (2000). Toward a cognitive neuroscience of metacognition. *Consciousness & Cognition, 9*, 313–323.

Shimamura, A. P., & Squire, L. P. (1986). Memory and metamemory: A study of the feeling-of-knowing phenomenon in amnesic patients. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 452–460.

Skarakis-Doyle, E. (2002). Young children's detection of violations in familiar stories and emerging comprehension monitoring. *Discourse Processes, 32*, 175–197.

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences, 26*, 317–373.

Son, L. K. (2005). Metacognitive control: Children's short-term versus long-term study strategies. *Journal of General Psychology, 132*, 347–363.

Walters, D. B., & Chapman, R. S. (2000). Comprehension monitoring: A developmental effect? *American Journal of Speech-Language Pathology, 9*, 48–54.

Wellman, H. M., Ritter, K., & Flavell, J. H. (1975). Deliberate memory behavior in the delayed reactions of young children. *Developmental Psychology, 11*, 780–787.

White, G. L., Fishbein, S., & Rutstein, J. (1981). Passionate love and the misattribution of arousal. *Journal of Personality and Social Psychology, 41*, 56–62.

Whitebread, D., Bingham, S., Grau, V., Pino Pasternak, D., & Sangster, C. (2007). Development of metacognition and self-regulated learning in young children: Role of collaborative and peer-assisted learning. *Journal of Cognitive Education and Psychology, 6*, 433–455.

Winsler, A., Carlton, M. P., & Barry, M. J. (2000). Age-related changes in preschool children's systematic use of private speech in a natural setting. *Journal of Child Language, 27*, 665–687.

Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences, 8*, 12–17.

Zelazo, P. D., Gao, H. H., & Todd, R. (2007). The development of consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 405–432). New York: Cambridge University Press.

Chapter 13
Children's Metacognition and Theory of Mind: Bridging the Gap

Plousia Misailidi

1 Introduction

Metacognition and theory of mind (ToM) have evolved over the past 20 years as two distinct and unconnected research fields. Nevertheless, as Flavell (2002) maintains, the two fields share the same overall objective, namely "to investigate the development of children's knowledge and cognition about mental phenomena" (p. 106). Whereas metacognition researchers are interested in children's developing capacity for thinking about – i.e., monitoring (or controlling) their own thoughts – ToM investigators address the ability to think about or make inferences about the thoughts and feelings of another person (Kuhn, 2000a, 2000b; Lockl & Schneider, 2006). Thus, the gap between the two research traditions may be more apparent than real.

The aim of the present chapter is to review some recent initiatives to bridge the gap between the metacognition and ToM research fields. The chapter is organized into three sections. The first section briefly introduces the metacognition and ToM constructs, and reports on some findings with regard to the age the two abilities begin to develop. The second asks why the gap between the metacognition and ToM research fields really exists. The suggestion that this gap is more apparent than real is discussed in the third section of this chapter, firstly by outlining two theoretical models attempting to incorporate ToM ability within a larger metacognition framework, and secondly by presenting some recent research that aimed to explore the relationship between children's ToM and metamemory as well studies investigating the relationship between children's ToM and metacognitive language.

P. Misailidi (✉)
Department of Primary Education, University of Ioannina, 451 10 Ioannina, Greece
e-mail: pmisaili@uoi.gr

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_13,

2 Metacognition and ToM defined

Metacognition has been broadly defined as awareness and management of one's own cognitive processes and products, or more simply as "thinking about thinking" (Flavell & Ross, 1981; Kuhn & Dean, 2004). Metacognition is generally considered to be a multidimensional construct (Schraw, 1998). A popular model (Flavell, Miller, & Miller, 2002) describes two related but conceptually distinct dimensions of metacognition: metacognitive knowledge and metacognitive processes. Metacognitive knowledge refers to a person's awareness and deeper understanding of cognitive processes and products (Vanderswalmen, Vrijders, & Desoete, this volume). Metacognitive processes, on the other hand, refer to an individual's ability to monitor and/or self-regulate her/his cognitive activities during problem solving (Flavell et al., 2002). Besides these two dimensions, Flavell's (1979) theoretical model also features metacognitive experiences (that is, "conscious or affective experiences that accompany and pertain to any intellectual enterprise", p. 906) as a prominent aspect of metacognition. Accordingly, a recent conceptualization of metacognition (Efklides, 2001, 2006, 2008) emphasizes metacognitive experiences as distinct from metacognitive processes, because the former are a manifestation of monitoring whereas the latter of control.

For many years it was believed that the various manifestations of metacognitive ability emerge around the age of 8–10 years (Veenman, Van Hout-Wolters, & Afflerbach, 2006). However, this view has been challenged recently on both theoretical and methodological grounds (see, e.g., Whitebread et al., 2009; Whitebread et al., this volume). From a theoretical standpoint, recently advanced models put forward that metacognition emerges from a very young age (Balcomb & Gerken, 2008; Koriat, 1993, 1994; Nelson & Narens, 1990, 1994). For example, Koriat's (1993, 1994) "trace accessibility" model suggests that metacognition may be available as a cognitive tool for learning in the form of implicit access to knowledge states that can drive behaviour, long before it is well differentiated and verbalizable. From a methodological standpoint, researchers have recently begun to recognise that past studies, using experimental paradigms that made heavy demands on young children's verbal and working memory abilities, underestimated young children's performance on metacognitive tasks (Van Hout-Wolters, 2000; Whitebread et al., 2005). Aiming to overcome this methodological drawback, Whitebread and his associates (Demetriou & Whitebread, 2008; Whitebread et al., 2005) observed preschool children in naturalistic settings (their kindergarten) and found evidence of source memory and other forms of metacognitive knowledge even in these young children. Similarly, Balcomb and Gerken (2008), using a non-verbal task, originally developed for work with non-human animals, demonstrated memory-monitoring skills in toddlers.

A related area of cognitive development is ToM, which has been broadly defined as knowledge about the existence of the mind and its contents (e.g., beliefs, desires and intentions) as well as the ability to use this knowledge for the prediction and explanation of human action (Premack & Woodruff, 1978). Based on this definition, ToM is considered as a valuable social tool; assumptions about other people's mental states guide children's actions in their social environment and frame their inferences and interpretations of other people's behaviour. The effects of ToM are also spread across cognitive, communication, and emotional development (Hughes & Leekam, 2004; Lalonde & Chandler, 1995).

Children begin to develop a ToM from a very early age, but this ability undergoes major shifts during the course of development. One such shift occurs around the age of 4 years, when children develop the capacity to recognize that other people's as well as their own beliefs can be false. For example, whereas 3-year-olds do not understand that another person could hold a false belief most 4-year-olds understand that beliefs are representations of reality that can be mistaken (Gopnik & Wellman, 1994; Perner, 1991). Once the false-belief concept is fully developed, children are claimed to be mind-readers. In other words, the acquisition of an awareness of other people's false beliefs is considered to be the benchmark accomplishment signaling the emergence of a ToM[1] (Dennett, 1978).

The study of the development of metacognitive and ToM abilities in young children has made rapid progress in recent years. Nevertheless, the two research fields have curiously remained isolated from one another. Very few attempts have been made to investigate developmental interrelationships between these two abilities.

3 Why Is There a Gap Between Metacognition and ToM Research?

Flavell (1997, 2000) and Kuhn (1999, 2000a, 2000b) have speculated reasons as to why the metacognition and ToM research fields have been kept separate for so many years. Three reasons have been put forward:

1. Investigations in the two research fields target different age groups. ToM research has been largely confined to children aged up to 6 years. On the other hand, metacognitive development research has focused – at least until very recently, as explained earlier – on developments that occur during the elementary-school and the adolescent years.
2. Metacognition research investigates task-related mental activities, including strategies for improving performance on tasks or attempts to monitor these improvements. In contrast, ToM research is interested in children's knowledge about the contents of the mind. For example, do young children appreciate that

[1]False belief understanding is just one of the multiple facets of the ToM construct. ToM encompasses a range of reasoning abilities besides the ability to comprehend false beliefs. Other manifestations of ToM include the ability to distinguish appearance from reality, the ability to comprehend the distinction between desire and intention, and the ability to understand knowledge. Studies have shown that all these abilities are mastered before the age of 6 years (e.g., Flavell, Flavell, & Green, 1983; Pillow, 1999; Schult, 2002).

false beliefs typically cause mistaken actions or can they distinguish between the concepts of desire and intention?

3. Metacognition research is concerned more with what children know about their own mental processes (Flavell, 2000). In contrast, ToM research is mainly interested in children's ability to think about or make inferences about the contents of other people's minds. For example, the classical false belief task assesses young children's understanding of other people's false beliefs.
4. An additional, fourth, reason that may explain the discontinuity between the metacognition and ToM research approaches is related to their domains of application. Applications of ToM research have predominantly been located in the social arena, examining mostly the implications of children's emerging ToM skills on their social interactions[2]. Metacognition research, on the other hand, has been predominantly located in the academic (educational) arena focusing mainly on the impact of metacognitive abilities on educational outcomes.

In short, the gap between the metacognitive and ToM development research traditions exists because their foci and applications are different. Nevertheless, despite these divergences, the conceptual connection between these two bodies of research is self-evident: Both imply activities involving thinking about thinking or the formation of cognitions about cognition (Flavell, 1997). Moreover, as it was indicated earlier, there is evidence that both capacities emerge during the same period of development, the preschool years. So, the gap between these two literatures may be more a matter of emphasis rather than a genuine divide.

4 Attempting to Bridge the Gap Between Metacognition and ToM Research

In the last few years a number of theoretical and empirical initiatives have been made to bridge the gap between the metacognition and ToM research fields.

4.1 Theoretical Models Attempting to Link Metacognition and ToM

Two recent theoretical models have proposed links between metacognitive and ToM competencies. Both these models describe ToM as one of the multiple dimensions of the construct of metacognition. The first model (Kuhn, 1999, 2000a, 2000b)

[2]It is worth noting that in recent years, several researchers have abandoned the term "theory of mind" in favour of the term "social understanding" (e.g., Nelson, Plesa, & Henseler, 1998), in an attempt to emphasize the central role that children's inferences about the mind hold for social interaction.

locates ToM within a broader 'meta-knowing' framework, whereas the second model (Alexander & Schwanenflugel, 1996; see also Efklides, 2008) describes ToM as metacognitive knowledge about mental activity concepts. More specifically the two models are described below.

Kuhn's (2000a) model describes metacognition or "meta-knowing" as "any cognition that has cognition…as its object" (p. 302). Meta-knowing consists of three components: (a) metacognitive knowing, (b) metastrategic knowing, and (c) epistemological knowing. The distinction between metacognitive and metastrategic knowing is based on the widely employed dichotomy in cognitive psychology between the concepts of declarative and procedural knowledge. Metacognitive knowing refers to one's base of declarative knowledge; that is, knowledge a person may have about cognition (knowing that). Metastrategic knowing, on the other hand, involves procedural knowledge; that is, a person's knowledge about her/his own cognitive processes and of their impact on performance (*knowing how*). Epistemological knowing is the more abstract component of meta-knowing and it has to do with an individual's broader understanding of what knowledge and knowing are in general (how does anyone come to know).

How does the concept of ToM fit into Kuhn's (1999, 2000a, 2000b) theory of meta-knowing? According to Kuhn, ToM corresponds to the metacognitive knowing component of meta-knowing and includes children's knowledge about the mind (i.e., knowledge that mental states exist). This declarative knowledge about the contents of the mind can be according to Kuhn both personal and impersonal. Personal metacognitive knowing is knowledge about one's own mental states, whereas impersonal metacognitive knowing is knowledge about others' mental states. In this respect, young children's ToM refers to their ability to view themselves as well as other people as cognizers – both abilities are clearly metacognitive processes.

One of the core claims in Kuhn's (1999, 2000a, 2000b) theoretical model is that ToM serves as the foundation for the development of both metastrategic and epistemological knowing. This means that children need to acquire a ToM first, before they begin to develop the other two dimensions of meta-knowing. Kuhn's claim is conceptual rather than empirical: She assumes that having concepts of mental states, such as beliefs, is prerequisite to thinking about the strategies available in one's repertory and appreciating the nature of epistemological beliefs. In other words, Kuhn considers the acquisition of mental state concepts as a necessary initial step before the development of the other components of metacognition.

Similarly to Kuhn's (2000a, 2000b), Alexander and Schwanenflugel's (1996) model distinguishes three components of metacognition: (a) declarative metacognitive knowledge, (b) cognitive monitoring, and (c) regulation of strategies. The former component corresponds to the individual's knowledge about the contents of the mind (i.e., her/his ToM). The two latter metacognitive components respectively refer to the individual's ability to read one's own mental states and to predict how these states will affect present and future performance on a mental activity task as well as the ability to use metacognitive knowledge strategically to achieve goals (Alexander, Fabricius, Fleming, Zwahr & Brown, 2001).

According to Alexander and Schwanenflugel (1996), declarative metacognitive knowledge comprises three bodies of knowledge: (a) knowledge about mental activity concepts; (b) declarative metacognitive knowledge about these concepts; and (c) strategy-specific metacognitive attributions. Knowledge of mental activity concepts in this model refers to the comprehension of the language (or terms) that one uses to describe the mind or mental activities as well as to the organization (the intentional relations) of mental activity concepts. Declarative metacognitive knowledge entails knowledge of the task and situational variables that may influence an individual's performance in different cognitive tasks. Finally, strategy-specific metacognitive attributions refer to children's understanding of the operation of specific strategies (i.e., why specific strategies operate the way they do).

In sum, Kuhn's (1999, 2000a, 2000b) and Alexander and Schwanenflugel's (1996) models conceptualize ToM as one of the several dimensions of the broader construct of metacognition. In Kuhn's model, ToM is a particular kind of meta-knowing, whereas in Alexander and Schwaneflugel's model ToM constitutes declarative knowledge of one's cognitive content. Kuhn maintains that the early ToM achievements (e.g., false belief understanding) serve as the foundation for the metacognitive (metastrategic and epistemological) competencies that appear later in development. In contrast to Kuhn's, Alexander and Schwanenflugel's theoretical model does not make any specific predictions with regard to the developmental trajectory of ToM in relation to the other dimensions of metacognition.

4.2 Empirical Studies Relating Metacognition and ToM

There are at least two points of contact between research on metacognitive and ToM development: (a) One is research investigating developmental interrelationships between children's metamemory and ToM competencies. (b) The other is studies investigating associations between ToM and understanding of metacognitive language.

4.2.1 Metamemory and ToM

Whereas metacognition refers to a broad range of activities and processes (knowledge, strategies, regulation), metamemory refers specifically to "one's knowledge of memory, how it works in general and what one's own memory is like in particular" (Uhlfelder, 1985, p. 6). Two recent studies (Demetriou, 2009; Lockl & Schneider, 2007) investigated whether there is a developmental relationship between metamemory and ToM. The research hypothesis that guided both these studies was that ToM skills constitute a precondition for the development of metamemory competencies. Children in the first study (Lockl & Schneider, 2007) were tested longitudinally, that is, at three different time points separated by a testing interval of approximately 1 year. At each time of testing, children completed a set of false belief tasks. At Time 3 (when the children were 5 years old), children's metamemory was also assessed in an interview which contained examples from everyday memory tasks as well as from laboratory-like situations. In an attempt to test whether the relationship between false belief and later metamemory competencies was mediated by differences in children's verbal mental age or nonverbal mental abilities, the authors also included tests of verbal and nonverbal mental age.

The results of Lockl and Schneider's (2007) study demonstrated that false belief comprehension and metamemory were strongly related. Importantly, the correlation between false belief performance and metamemory remained significant even when the contributions of verbal and non-verbal mental age scores were partialled out. A series of hierarchical regression analyses revealed that false belief performance at Time 1 (when the children were 3 years old) and at Time 2 (when the children were 4 years old) made independent contributions to performance on the metamemory tasks at Time 3. Overall, then, Lockl and Schneider's (2007) study presented evidence supporting that early development of ToM competencies facilitates the development of metamemory.

Metamemory, broadly defined, encompasses a number of judgments, including source monitoring (see Johnson, Hashtroudi, & Lindsay, 1993). Within the metamemory framework, source monitoring refers to the ability of individuals to determine the origins of their memories. The study by Demetriou (2009, see also Whitebread et al., this volume) utilized a longitudinal approach to examine developmental interrelationships between children's memory source monitoring and ToM. Children in Demetriou's (2009) study were assessed longitudinally on three time points, separated by a 6-month time interval (children had a mean age of 4 years at the beginning of the study). At each time point, children were tested on two false belief tasks as well as on a range of measures assessing the development of both cognitive (verbal and non-verbal mental age, working memory, inhibitory control, language skills and so forth) and memory source monitoring competencies. The results revealed significant correlations among children's false belief task performance, verbal mental age and other cognitive abilities. However, the strongest relationship was that between false belief performance and source memory competencies, and this relationship became stronger with the advancement of age. Further and more importantly, the results showed that the relationship between false belief performance and source memory monitoring remained unchanged even after the effects of verbal mental age and other cognitive abilities (e.g., inhibition control and working memory) were partialled out. Nevertheless, the expected direction of the relationship was not proved. Contrary to her predictions, the results of Demetriou's (2009) longitudinal study showed that earlier source memory monitoring significantly predicted later false belief task performance and vice versa (see Whitebread et al., this volume). This could be interpreted as evidence that ToM does not solely predict metamemory competencies; the reverse can also be true. ToM and source memory monitoring are bidirectionally related.

In sum, taken together the findings of the two above longitudinal studies suggest that the development of ToM is strongly interrelated with that of metamemory. Nonetheless, the data are contradictory with respect to whether ToM competencies constitute an early precursor of subsequent metamemory development or, conversely, whether earlier metamemory predicts later developments in ToM skills. Lockl and Schneider's (2007) results support that ToM is an earlier developmental achievement that predicts later metacognitive abilities, whereas Demetriou's (2009) findings provided evidence that the relationship between ToM and source memory competencies may be bidirectional. This means that the view advanced by Kuhn's (2000a, 2000b) model, representing ToM as an earlier socio-cognitive achievement and metacognition as a later competency may have been an overstatement which needs further investigation. More empirical evidence (ideally longitudinal or even training studies) is necessary to construct and evaluate more detailed and comprehensive accounts of the relationship between ToM and metacognition.

4.2.2 Metacognitive Language and ToM

Another attempt to interrelate the development of metacognition with ToM has been made by researchers whose interest is the relationship between metacognitive language and mental state understanding. Metacognitive language is language that describes the mind's contents and cognitive processes in general (see Olson & Astington, 1993). In Scholnick and Hall's (1991) terms, "it is language [*or terminology*] by which we signal to ourselves and others that we are engaged in some form of internal processing of events, and it's the language by which we identify that others are engaging in internal processing" (p. 402, italics added). Metacognitive language is considered as a crucial part of metacognition for two reasons. First, language about cognitive states and processes allows individuals "to gain access to our internal states [*and processes*], to monitor and transform them" (Scholnick & Hall, 1991, p. 402, italics added). Second, such terminology about cognition is subject to reflection that allows individuals "to understand and inter-relate aspects of mental functioning to one another" (Scholnick & Hall, 1991, p. 402).

Research has documented an increasing capacity to understand metacognitive language during the preschool years (Lyon & Flavell, 1994; Moore, Bryant, & Furrow, 1989). More specifically, research has shown that between the ages of 4 and 5 years children begin to comprehend the meaning of metacognitive terms, such as "know" and "guess". For example, if children are uncertain about the place where an object is hidden and a doll tells them that "she *guesses that* the object is hidden in place A", whereas a second doll tells them that "she *knows that* the object is in place B", children select place B as the place where they should search for the object (Moore et al., 1989). Around the same age, children also begin to comprehend the semantic differences between specific metacognitive terms, including the terms "learn", "remember" and "forget". More specifically, 4-year-olds, but not younger children, comprehend that the terms "remember" and "forget" do not simply refer to the knowledge or the ignorance of an individual but moreover imply the existence of previous knowledge (Lyon & Flavell, 1994) and, yet, that these terms differ from the term "guess" which implies absence of knowledge (Astington, 2000).

To what extent do the changes that take place in children's understanding of metacognitive terms during the preschool period are related with changes in the development of ToM competencies? So far only a handful of studies have investigated the relationship between the development of ToM and children's understanding of metacognitive terms (Astington & Pelletier, 1998a; Charman & Shmueli-Goetz, 1998; Moore, Pure, & Furrow, 1990). In one of these studies, Moore et al. (1990) examined 4- to 6-year-old children's abilities to: (a) distinguish between the verbs *believe* and *guess* with the verb *know*, and (b) to attribute false beliefs to others. The results showed that the two abilities were significantly related. That is, children who were successful in the false belief task were those who also correctly answered questions about the semantic differences between the verbs believe/know and guess/know.

In another study by Astington and Pelletier (1998a) the developmental relationship of ToM with metacognitive language was examined with two groups of children aged 4 and 5 years respectively. For the investigation of children's ability to understand mental states, these researchers used four false belief tasks, whereas for the investigation of children's ability to understand metacognitive terms the "Metacognitive Vocabulary Test" (Astington & Pelletier, 1998b) was employed. This test examines children's ability to understand the semantic differences between the metacognitive terms "know", "guess", "remember", "forget", "wonder", "figure out", "explain", "understand", "learn", "teach", "predict", and "deny". The results showed that children with higher scores on the false belief tasks were also more successful on the Metacognitive Vocabulary Test, whereas children who had low scores on the ToM tasks also scored lower on the Metacognitive Vocabulary Test.

Taken together, the findings of the above studies suggest important links between children's ability to attribute false beliefs (i.e., their ToM) and their acquisition of the ability to comprehend metacognitive terms. The cross-sectional nature of these findings, however, limits their interpretability, because it does not allow causal inferences to be drawn. A more recent longitudinal study by Lockl and Schneider (2006) provides some insight into the causal pathway of this relationship. Its results showed that early performance on ToM tasks significantly predicted later metacognitive vocabulary, even when individual differences in children's non-verbal mental age and general vocabulary were taken into consideration. These longitudinal data suggest that developmental changes in children's ToM predict changes in metacognitive language, a finding which lends evidence in support of Kuhn's (2000a, 2000b) claim that early ToM competencies precede advancements in metacognition.

5 Conclusions

In this chapter, we have identified some reasons why the metacognition and ToM research fields have been kept separate from one another for so many years. We have emphasized that this separation has probably been more a matter of emphasis than a genuine divide. The two theoretical models that we outlined in this chapter represent some first sketchy attempts to bring closer together these two research approaches. Importantly, the view that ToM constitutes one of the first manifestations of the construct of metacognition (Kuhn, 2000a, 2000b) has provided the impetus to investigate developmental interrelationships between these two abilities.

Further attempts to bridge the gap between the metacognition and ToM research should be considered imperative for a number of reasons: First, current theoretical models of the relationship between metacognition and ToM (Kuhn, 1999, 2000a, 2000b) favour ToM as an earlier developmental milestone. However, the available findings (Demetriou, 2009; Lockl & Schneider, 2006, 2007) are contradictory with respect to which of these two abilities precedes the other. While some evidence supports that ToM is an earlier achievement that predicts later metacognitive competencies, other data indicate that the two abilities are intercorrelated in a reciprocal manner. Further research is, clearly, needed to identify the direction of the relationship between metacognition and ToM. If, in accordance to Kuhn's (2000a, 2000b) claims, it is proved that ToM precedes metacognition, a further critical question to be considered is whether there is anything special about metacognition that requires ToM and, conversely, whether there is anything special about ToM that allows metacognition to develop.

Second, besides the question concerning the causal direction of the association between metacognition and ToM, further research is necessary to determine the relationship of different dimensions of ToM with specific components of the metacognition construct. Both ToM and metacognition are broad, multidimensional constructs: false belief is just one aspect of the capacity to reason about mental states; metamemory and metacognitive language are, similarly, just two of the multiple components of the metacognition construct. In essence, what this means is that, there exists a possibility of different relations among different dimensions or components of metacognition and ToM, and also the possibility of change in these relations over developmental time, as well as of individual differences in these relations. To understand the developmental links between metacognitive and ToM competencies more fully, researchers need to investigate the growth of different aspects of children's ToM in relation to the development of different aspects of metacognition.

Third, further research on the association of metacognition with ToM is warranted to shed light on the nature and course of each of these two aspects of cognitive development separately. The majority of studies that have been conducted on the development of ToM have so far focused on the competencies of children up to 6 years of age. ToM developments beyond the preschool period have been rarely investigated (e.g., Carpendale & Chandler, 1996; Fabricius & Schwanenflugel, 1994). More research into the relationship between metacognition and ToM is expected to provide more detailed information about the older child's concepts of thinking processes and about the development of their understanding of particular mental states, which will aid to adequately describe the developmental course of ToM beyond the preschool years (see Flavell, Green, & Flavell, 1995; Schwanenflugel, Fabricius, & Alexander, 1994).

Likewise, as it was explained in the first section of this chapter, research conducted on the metacognitive development of preschoolers is relatively scarce

(Whitebread et al., 2009). Further research on the relationship between metacognitive and ToM competencies will make possible a more accurate description and measurement of metacognitive development in young children. This research is clearly important and advantageous, both in relation to revising the existing models of metacognitive development as well as for charting the effects of the early (preschool) metacognitive achievements on later developments in metacognition (as well as in ToM). Evidently, it is necessary to know more about the components of metacognition that develop first and of the conditions under which this development occurs (Veenman et al., 2006). Moreover, it is important to investigate how earlier metacognitive competencies prepare the way for the development of the later ones.

To conclude, metacognition and ToM have been viewed, until recently, as two separate domains of cognitive development. Yet, the theoretical and empirical work reported in the previous sections suggests that these two abilities are developmentally inter-related. Further research on the association between metacognition and ToM must be done, before we can draw any conclusions about the causal nature of this relationship.

References

Alexander, J. M., & Schwanenflugel, P. J. (1996). Development of metacognitive concepts about thinking in gifted and non-gifted children: Recent research. *Learning and Individual Differences*, *4*, 305–326.

Alexander, J. M., Fabricius, W. V., Fleming, V. M., Zwahr, M., & Brown, S. A. (2001). The development of metacognitive causal explanations. *Learning and Individual Differences*, *13*, 227–238.

Astington, J. W. (2000). Language and metalanguage in children's understanding of mind. In J. W. Astington (Ed.), *Minds in the making* (pp. 267–284). Oxford, UK: Blackwell.

Astington, J. W., & Pelletier, J. (1998a). The language of mind: Its role in teaching and learning. In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development* (pp. 593–619). Oxford, UK: Blackwell.

Astington, J. W., & Pelletier, J. (1998b). *Metacognitive Vocabulary Test. Institute of Child Study.* Unpublished manuscript, University of Toronto, Canada.

Balcomb, F. K., & Gerken, L. A. (2008). Three-year-old children can access their own memory to guide responses on a visual matching task. *Developmental Science*, *11*, 750–760.

Carpendale, J. I. M., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development*, *67*, 1686–1706.

Charman, T., & Shmueli-Goetz, Y. (1998). The relationship between theory of mind language and narrative discourse: An experimental study. *Current Psychology of Cognition*, *17*, 245–271.

Demetriou, D. (2009). *Inter-relationships in children's socio-cognitive development: The case of false belief understanding.* Unpublished doctoral dissertation, University of Cambridge, UK.

Demetriou, D., & Whitebread, D. (2008). *Theory of mind and metacognitive knowing: Have we been investigating similar constructs without realizing it?* Paper presented at the 3rd Biennial Meeting of the EARLI Special Interest Group 16: Metacognition, Ioannina, Greece.

Dennett, D. (1978). Beliefs about beliefs. *Behavioural and Brain Sciences*, *1*, 568–570.

Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation, and self-regulation. In A. Efklides, J. Kuhl, & R. M. Sorrentino (Eds.), *Trends and prospects in motivation research* (pp. 297–323). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review*, *1*, 3–14.
Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self- and co-regulation. *European Psychologist*, *13*, 277–287.
Fabricius, W. V., & Schwanenflugel, P. J. (1994). The older child's theory of mind. In A. Demetriou & A. Efklides (Eds.), *Intelligence, mind, and reasoning: Structure and development* (pp. 111–132). Amsterdam: Elsevier.
Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new era of cognitive-developmental inquiry. *American Psychologist*, *34*, 906–911.
Flavell, J. H. (1997). *Recent convergences between metacognitive and theory-of-mind research in children.* Paper presented at the Biennial Meeting of the Society for Research in Child Development: Constructing Metacognitive Knowledge, Washington, DC.
Flavell, J. H. (2000). Development of children's knowledge about the mental world. *International Journal of Behavioral Development*, *24*, 15–23.
Flavell, J. H. (2002). Development of children's knowledge about the mental world. In W. W. Hartup & R. K. Silbereisen (Eds.), *Growing points in developmental science: An introduction* (pp. 102–122). Hove, UK: Psychology Press.
Flavell, J. H., & Ross, L. (1981). *Social and cognitive development.* New York, NY: Cambridge University Press.
Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, *15*, 95–120.
Flavell, J. H., Green, F. L., & Flavell, E. R. (1995). The development of children's knowledge about attentional focus. *Developmental Psychology*, *31*, 706–712.
Flavell, J. H., Miller, P. H., & Miller, S. A. (2002). *Cognitive development* (4th ed.). Upper Saddle River, NJ: Prentice Hall.
Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. Hirschfield & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257–293). New York, NY: Cambridge University Press.
Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*, 590–619.
Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.
Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*, 609–639.
Koriat, A. (1994). Memory's knowledge of its own knowledge: The accessibility account of the feeling of knowing. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 115–135). Cambridge, MA: MIT Press.
Kuhn, D. (1999). Metacognitive development. In L. Balter & C. S. Tamis-LeMonda (Eds.), *Child psychology: A handbook of contemporary issues* (pp. 259–286). Philadelphia, PA: Psychology Press.
Kuhn, D. (2000a). Metacognitive development. *Current Directions in Psychological Science*, *9*, 178–181.
Kuhn, D. (2000b). Theory of mind, metacognition, and reasoning: A life-span perspective. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 301–326). Hove, UK: Psychology Press.
Kuhn, D., & Dean, D. (2004). Metacognition: A bridge between cognitive psychology and educational practice. *Theory into Practice*, *43*, 268–273.
Lalonde, C. E., & Chandler, M. J. (1995). False belief understanding goes to school: On the social-emotional consequences of coming early or late to a first theory of mind. *Cognition and Emotion*, *9*, 167–185.
Lockl, K., & Schneider, W. (2006). Precursors of metamemory in young children: The role of theory of mind and metacognitive vocabulary. *Metacognition and Learning*, *1*, 15–31.

Lockl, K., & Schneider, W. (2007). Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development*, *78*, 148–167.
Lyon, D. L., & Flavell, J. H. (1994). Young children's understanding of 'remember' and 'forget'. *Child Development*, *65*, 1357–1371.
Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development*, *60*, 167–171.
Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development*, *61*, 722–730.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–141). San Diego, CA: Academic.
Nelson, T., & Narens, L. (1994). Why investigate metacognition. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.
Nelson, K., Plesa, D., & Henseler, S. (1998). Children's theory of mind: An experiential interpretation. *Human Development*, *41*, 7–29.
Olson, D. R., & Astington, J. W. (1993). Thinking about thinking: Learning how to take statements and hold beliefs. *Educational Psychologist*, *28*, 7–23.
Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT press.
Pillow, B. H. (1999). Children's understanding of inference and knowledge. *Journal of Genetic Psychology*, *160*, 419–428.
Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*, 515–526.
Scholnick, E. K., & Hall, W. S. (1991). The language of thinking: Metacognitive and conditional words. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 397–439). Cambridge, UK: Cambridge University Press.
Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, *26*, 113–125.
Schult, C. A. (2002) Children's understanding of the distinction between intentions and desires. *Child Development*, *73*, 1727–1747.
Schwanenflugel, P. J., Fabricius, W. V., & Alexander, J. (1994). Developing theories of mind: Understanding concepts and relations between mental activities. *Child Development*, *65*, 1546–1563.
Uhlfelder, H. F. (1985). Metamemory and problem solving: The search for connections. Unpublished doctoral dissertation, Georgia State University, Atlanta, GA.
Van Hout-Wolters, B. (2000). Assessing active self-directed learning. In R. Simons, J. van der Linden, & T. Duffy (Eds.), *New learning* (pp. 83–101). Dordrecht, The Netherlands: Kluwer.
Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, *1*, 3–14.
Whitebread, D., Anderson, H., Coltman, P., Page, C., Pino Pasternak, D., & Mehta, S. (2005). Developing independent learning in the early years. *Education 3–13*, *33*, 40–50.
Whitebread, D., Coltman, P., Pino Pasternak, D., Sangster, C., Grau, V., Bingham, S., Almeqdad, Q., & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning*, *4*, 63–85.

Chapter 14
Self-Confidence and Academic Achievements in Primary-School Children: Their Relationships and Links to Parental Bonds, Intelligence, Age, and Gender

Sabina Kleitman and Tanya Moscrop

1 Introduction

Decision-making has occupied the minds of many esteemed researchers in different disciplines. What has attracted researchers to the study of decision-making is its potential to provide theoretical and diagnostic frameworks, with application to areas as diverse as psychology, education, economics and law. Within the field of decision-making, *knowledge calibration* is a major paradigm. It concerns the self-monitoring, in terms of confidence judgments, that people assign to events (answers to questions, decisions, predictions) and their correspondence to the accuracy of those events (see Harvey, 1997 for a review). In fields of education and psychology, these confidence judgments have been referred to as self-confidence ratings and are argued to initiate an essential component of *metacognition*, that is, self-monitoring (for a review, see Stankov, 1999).

This chapter focuses on an important aspect of knowledge calibration, test-taking situations, where people are given multiple-choice questions and are asked to quantify the level of their confidence in each answer. The decision-making paradigm stresses general tendencies and views confidence ratings as a reflection of certain decision-making processes that are supposed to follow the normative laws of different theories of probability. At the same time, the individual differences approach, while acknowledging general tendencies in the way people assess their confidence, emphasises person-driven factors that predispose people to give higher or lower confidence judgments. Findings from different fields of research are overviewed and the results of studies coming from our laboratory are presented in light of metacognitive theory. The individual differences approach is used to provide the framework for an integrative model of confidence judgments. Their predictive validity in school achievements and their determinants are discussed.

S. Kleitman (✉)
School of Psychology, The University of Sydney, Sydney, NSW 2006, Australia
e-mail: sabinak@psych.usyd.edu.au

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_14,

1.1 Metacognition and Knowledge Calibration Paradigm

Metacognition refers to the executive processes involved in reflecting on one's own thinking; that is, "thinking about thinking" (Flavell, 1979) or "knowing about knowing" (Metcalfe & Shimamura, 1994). Most theories distinguish between two major components of metacognition – knowledge about cognition and regulation of cognition; the latter consisting of monitoring and control of cognition (Brown, 1986; Nelson & Narens, 1994; Schraw & Dennison, 1994; Schraw & Moshman, 1995). Monitoring of cognition, the focus of this chapter, is defined as the ability to watch, check and appraise the quality of one's own cognitive work in the course of doing it (Schraw & Moshman, 1995). Confidence judgments reflect these instances by deliberately evoking subjective feelings of certainty that one experiences in connection with decision-making and action-regulation (Allwood & Granhag, 1999; Stankov, 1999).

The main emphasis of the knowledge calibration paradigm is placed upon the different aspects of the association (or calibration) between confidence ratings and actual performance (for a review see Harvey, 1997). Metacognitive self-monitoring is reflected in the different measures illustrating this correspondence (for reviews see Stankov, 1999; Stankov & Kleitman, 2008). However, all these calibration-type measures are initiated by confidence scores, which are the essential ingredient of such calculations. Moreover, while accuracy of performance and overall confidence levels can be manipulated by environmental factors (see Harvey, 1997), the systematic individual differences reflect that the Self-Confidence factor remains stable and unaffected by such manipulations. For instance, Kleitman and Stankov (2001) employed representative and non-representative (or misleading, so called "tricky") general knowledge items (Gigerenzer, Hoffrage, & Kleinbolting, 1991; Juslin, 1994) as well as items capturing diverse domains of cognitive range, namely reasoning, perceptual, and general knowledge. These manipulations resulted in important differences in measures, reflecting the calibration matters. However, the Self-Confidence factor still emerged despite the experimental manipulations. This finding suggests the habitual nature of the process of assessing one's own competence to deal with uncertainty in test-taking situations. Thus, understanding the psychological factor that underlies the stability of self-confidence could provide a powerful window into metacognitive self-monitoring and knowledge calibration.

Although there is an important conceptual overlap between self-confidence and self-efficacy judgments, there is a major distinction between the two in terms of broadness. Self-efficacy refers to a belief that if one is engaged in a particular behaviour, one will achieve a positive and desired outcome (Bandura, 1997). It is a form of confidence in one's own ability to perform on a specific task or within a particular domain. A closer examination of the empirical evidence which supports the constructs of self-confidence and self-efficacy suggests that self-efficacy, in comparison to self-confidence, tends to be domain specific – it is limited to a particular task or a domain (e.g., mathematics, verbal and/or physical domain). Self-confidence, however, has consistently been shown to be a general factor that extends across different tasks and domains (for a review see Kleitman, 2008). Moreover, while self-efficacy is modifiable (Bandura, 1997), confidence ratings are shown to remain unaffected despite many experimental manipulations (Allwood & Granhag, 1999; Kleitman & Stankov, 2001).

Some of the important questions needing to be addressed are "When does this self-confidence trait begin to stabilise in the course of our lives?", "Who or what is responsible for this stability: Genetics? Parents? Teachers? Peers?", "Which individual or collective experiences contribute to it?", "Do levels of confidence foster educational achievements *incrementally* to cognitive abilities (and other traditional factors in education such as age and gender)?" This chapter aims to answer three questions: (a) Do primary school-aged children already display the habitual general levels of self-confidence across different cognitive domains? (b) Do family dynamics predict confidence levels? (c) What is their predictive validity in the school setting, incremental to the traditional factors?

1.2 Self-Confidence as an Aspect of Metacognitive Self-Monitoring

This work relies on a definition that captures the main purpose of self-monitoring, that is, the ability to judge the quality of one's own performance in the course of doing it. In such an instance, immediately after responding to an item in a test, participants are instructed to give a confidence (or "sureness") rating indicating how confident/sure they are that their chosen answer is the correct one (see Fig. 14.1). It is important to distinguish this assessment of self-confidence from the putatively similar personality trait(s) that is/are presumed to arise from the responses to items such as "I feel self-assured when I have to give a speech to a large group of people", "I'm self-confident" and "I'm self-assured". The confidence rating procedure follows the cognitive act of providing a response to a typical cognitive test item, rather than relying on a general perception of one's own habitual way of acting. We now have overwhelming psychometric evidence that this numerical method – which probes the *actual* cognitive act rather than relying on a subjective perception of it – is a more accurate measure of the self-confidence trait

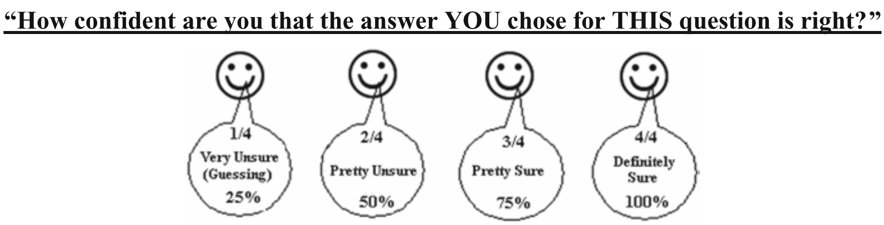


Fig. 14.1 A four-point confidence rating scale used in the Class test

(Stankov, 1999; Kleitman, 2008). This trait reflects the *habitual* way in which adults assess the accuracy of their cognitive decisions.

1.3 Confidence Judgments[1]

The procedure is simple. Specifically, immediately after responding to a question, people are asked to rate on a percentage or probability scale, how confident they are that their answer is correct. The level of confidence is expressed in terms of percentages and/or verbal statements. The starting point (the lowest confidence) on a rating scale is defined in terms of the number of alternative answers (k) given to a question (100/k). Thus, there are different starting points for questions with two and five alternative answers (50% and 20%, respectively). That is, in multiple-choice questions with five alternative answers, 20% is a starting point because 20% is the probability of answering the question correctly by chance. This is explained to a participant and often indicated on the rating scale (Allwood, Granhag, & Jonsson, 2006; Allwood, this volume). Consequently, the confidence rating scales may include both percentages and labels (e.g., "guessing", "fairly sure", "absolutely certain"), respectively. The confidence ratings for all attempted test items are averaged to give an overall confidence score.

The scales for this type of confidence ratings – including both percentages and labels – could take several forms (Allwood et al., 2006; Allwood, this volume). Importantly, it has been demonstrated that the outcomes of research remain stable regardless of the type of scale used for confidence ratings (Allwood et al., 2006; Allwood, this volume).

1.4 Empirical Findings

There are numerous findings in relation to confidence judgments, especially in the adult population. In this chapter, we review the research findings that stem from psychological and educational traditions, using the individual differences framework

[1] Moore and Healy (2008) provide a comprehensive review of different types of confidence judgments, such as unique confidence judgments that people provide immediately after responding to a test's item, and general ratings of one's state of knowledge/performance in comparison to the others. Immediate confidence judgments could be given in two broad formats: (a) in terms of unique probabilistic numbers along a "confidence scale" or/and as a verbal category along a typical Likert-type scale (e.g., ranging from "unsure" to "very sure") and (b) as confidence intervals asking participants to estimate for instance 90% confidence intervals around their answers. The former judgments are more prevalent than the latter, comprising 64% of research on knowledge calibration (Moore & Healy, 2008), and in a series of studies were demonstrated to reflect a thought level higher than knowledge – the metacognitive level – within the taxonomy of cognitive/metacognitive analysis (Kleitman, 2008). The confidence judgments of the first broad format are at the focus of this chapter.

of research. However, a comprehensive review of all the findings gathered using the knowledge calibration paradigm is outside the scope of this chapter, and is available elsewhere (Harvey, 1997).

1.4.1 Self-Confidence Trait

Confidence judgments have high internal consistency (reliability estimates are typically higher than 0.90) (see Kleitman, 2008 for a review) and robust test–retest estimates (Jonsson & Allwood, 2003). There is overwhelming empirical evidence showing individual differences in confidence ratings (for a review see Kleitman, 2008). The correlations between accuracy and confidence scores from the *same* test are significant (average between 0.40 and 0.50). Nevertheless, correlations between confidence ratings from a broad battery of cognitive tests reflecting diverse cognitive abilities have been consistently high enough to define a strong Self-Confidence factor. That is, people who are more confident on one task, relative to other people, also tend to be more confident across other tasks. Thus, when measured across different items, cognitive tests, and knowledge domains, a Self-Confidence factor emerges to reflect the *stability* of confidence judgments.

Table 14.1 summarises such results from a study by Kleitman and Stankov (2007). Specifically, some tests sample several different scores, namely accuracy of performance, confidence and the so-called speed scores. That is, in addition to the typical correct/incorrect scoring (accuracy measure for each item), at least on some tests (here Verbal Reasoning, Syllogisms, Esoteric Analogies, and General Knowledge tests) people were asked to indicate their confidence levels in each answer. In addition, the time taken to answer each item was collected on computerised tests (here Verbal Reasoning, Syllogisms, and Esoteric Analogies) and is referred to as test-taking speed, or ‘speed’ scores. These scores are averaged across the test to index test-taking speed for each test. When factor analytic techniques are used (either exploratory or confirmatory; here confirmatory), several latent traits or factors typically emerge. These are cognitive ability or intelligence factors defined by the relevant accuracy measures – here the results were separated into Fluid Intelligence (Gf) and Crystallized Intelligence (Gc) factors;[2] Test-taking Speed or Speed factor, defined by the speed scores; and the Self-Confidence factor, defined by the confidence scores. The fourth factor (its relevance will become apparent later) was defined by the self-report measures, that is, the Metacognitive Awareness Inventory (MAI; Schraw & Dennison, 1994) and our own Memory and Reasoning Competence Inventory (MARCI; Kleitman & Stankov, 2007), sampling different aspects of metacognition.

[2] The Horn-Cattell theory is a hierarchical model that defines intelligence in terms of independent broad abilities (Carroll, 1993). According to the model, fluid intelligence (*Gf*) reflects basic abilities in reasoning, while crystallized intelligence (*Gc*) reflects the effects of acculturation. The model regards *Gf* and *Gc* as second-order factors, while *g* refers to a general intelligence, a higher-order factor (Horn & Noll, 1994).

Table 14.1 The findings of the Kleitman and Stankov (2007) study regarding the structure of accuracy, confidence, test-taking speed, and metacognitive measures scores

Factors	Gf	Gc	Confidence	Speed	Metacognitive processes	h^2
Accuracy						
Quantitative switching	0.50					0.25
Verbal reasoning	0.34	0.42				0.39
Syllogisms	0.54					0.30
Esoteric analogies	0.53	0.21		0.26		0.47
General knowledge		0.82				0.67
Probability reasoning	0.48					0.23
Conditional reasoning	0.64					0.42
Confidence						
Verbal reasoning		0.22	0.57			0.42
Syllogisms			0.64			0.40
Esoteric analogies			0.90			0.81
Sureness	–0.31		0.39		0.22	0.28
General knowledge		0.57	0.40			0.58
Test-taking speed						
Verbal reasoning				0.46		
Syllogisms				0.69		
Esoteric analogies				0.93		
Metacognitive measures						
MAI					0.69	0.47
Memory inventory	0.19				0.37	0.17
Reasoning inventory	0.38				0.54	0.44
Factor correlations						
Gf	1	0.34	0.34	–	–	
Gc		1	0.20	–	–	
Confidence			1	–	0.41	
Speed				1	0.30	
Metacognitive processes					1	

Note: MAI = metacognitive awareness inventory; memory inventory = memory competence score of the memory and reasoning competence inventory; reasoning inventory = reasoning competence score of the memory and reasoning competence inventory; Gf = fluid intelligence factor; Gc = crystallised intelligence factor; confidence = self-confidence factor, speed = test-taking speed factor; metacognitive processes = metacognitive processes factor

The Self-Confidence factor is well established in Differential Psychology and is argued to reflect a latent trait which underlies processes higher than the ‘knowledge’ level of cognition, representing an essential component of a regulatory, self-monitoring aspect of metacognition (Stankov, 1999). Kleitman (2008) empirically demonstrated the veracity of such a claim. In a series of studies, the unique nature of the Self-confidence trait was determined. When a diverse number of cognitive tests was employed, the robust Self-Confidence factor always emerged, defined by the confidence ratings which people assign to their answers. The factor was broad enough to include Sureness judgements (see Table 14.1) – confidence ratings which participants assigned to a set of non-cognitive items, asking people to express their opinion on events that may or may never happen (e.g., a likelihood that a cure for AIDS will ever be found). This generality of the Self-confidence factor provides key evidence of broad, perhaps basic, human factors which predispose people to adopt a particular level of confidence across different cognitive acts (whether verifiable or not).

While sharing meaningful positive relationships with the *Gf* and *Gc* ability factors (here the *r* values are 0.34 and 0.20, respectively), the Self-Confidence factor extended beyond these factors. As evident from Table 14.1, the Self-Confidence factor also had a meaningful positive association with the Metacognitive Processes factor ($r=0.41$). This suggests that people who hold higher beliefs in the competence of their cognitive abilities (as captured by MARCI) and in the quality of their metacognitive awareness in general (as captured by MAI), assign higher confidence ratings to their answers and opinions. It is worth noting that people's beliefs regarding their reasoning competencies were related to the actual performance on a variety of tests that relied on reasoning abilities – it had a meaningful loading on the Gf factor (see Table 14.1). This highlights the veracity of such beliefs. Importantly, confidence judgments were predicted by the Reasoning score of MARCI and this prediction remained significant after controlling for relevant accuracy scores, the common factor for both, confidence levels and the Reasoning score of MARCI. Together, these findings are important as they attest to that confidence ratings reflect processes, meaningfully related to, but other than the 'knowledge' level of cognition, verifying the *metacognitive* nature of confidence ratings.

Many other established psychological constructs have been investigated as predictors of the self-confidence trait. These constructs include personality and a variety of *global* self-esteem and self-concept measures (see Kleitman & Stankov, 2007 for a review). However, no consistent associations with these constructs have been established. Nevertheless, there are known predictors of the Self-Confidence factor, such as intelligence (Stankov, 2000), age (Stankov & Crawford, 1996), gender (Pallier et al., 2002), *specific* self-concept measures (Efklides & Tsiora, 2002; Kröner & Bierman, 2007; Kleitman & Stankov, 2007), and parental rearing techniques (Want & Kleitman, 2006).

1.4.2 Intelligence

As a Predictor

As mentioned above, one of the well established predictors of self-confidence is performance accuracy (measured on the same cognitive task which is used to measure self-confidence levels) where greater accuracy has been shown to predict greater confidence (Kleitman & Stankov, 2007). Thus, individuals who perform better on a given cognitive test assign higher confidence ratings to their answers. This result typically extends to the performance on other cognitive tests, as performances on individual cognitive tests tend to correlate positively with one another, a phenomenon known as positive manifold (Carroll, 1993). Thus, an intelligence factor (or factors) shares a significant and psychologically meaningful positive relationship with self-confidence.

As a Control Variable

This relationship that self-confidence shares with intelligence may be falsely inflated; thus it requires clarification. In other words, accuracy of performance in cognitive tests employing confidence ratings is a common factor (a variable assumed to affect the influence and the outcome) which influences both confidence ratings and intelligence. Accordingly, Kleitman and Stankov (2007) argue that to accurately assess a relationship between self-confidence and any intelligence-related measure, the accuracy of cognitive performance needs to be controlled (its common influence must be partitioned out). Thus, it is necessary to control for performance accuracy when examining the influence that any psychological factor has on self-confidence.

Moreover, intelligence influences academic achievement measures (Veenman, Wilhelm, & Beishuizen, 2004). Thus, in the present investigation, intelligence is considered in two ways. Firstly, it is considered as an important predictor of both self-confidence and academic performance. Secondly, it is considered as a common factor needing to be controlled for, when the influence of self-confidence on academic achievements is examined.

1.4.3 Age

As a Predictor

Prior research has established that older adults tend to have higher levels of self-confidence than their younger counterparts (Stankov, 1999; Stankov & Crawford, 1996; Want & Kleitman, 2006). To date, no such research has been undertaken with children. However, if a similar trend exists in children, older children would be expected to exhibit greater levels of self-confidence than their younger counterparts. However, in self-concept research, younger children compared to older children are found to be "overoptimistic" when assessing their abilities, while older children have a better calibrated self-concept in relation to their academic performance (for a review, see Efklides & Tsiora, 2002). Therefore, if a similar trend existed in children, younger children could have greater (and less realistic) levels of self-confidence than their older peers. Thus, age was included as a predictor variable; however, no directional predictions in relation to self-confidence were made.

As a Control Variable

As any intelligence test manual will attest (e.g., Raven, Raven, & Court, 2003), on average, older children tend to achieve greater cognitive accuracy scores when given the same testing instrument (when data is used without an adjustment for norms, as it is in this study). Given that greater performance accuracy is a known predictor of self-confidence among adults, it is expected that greater performance will also influence children's levels of self-confidence. Thus, it is important to control for age when examining the relationships between intelligence, performance accuracy, and self-confidence. If the control is not exercised, older children may exhibit inflated levels of the Self-Confidence factor as a consequence of their superior test performance. Moreover, it is important to control for age when predicting academic achievement (Marsh & Kleitman, 2002). Therefore, the effect of age will be statistically controlled for when considering the effects of all other variables in the present study.

Age as a Developmental Factor in Shaping the Broadness of Metacognitive Processes

It is currently unknown at what age confidence ratings develop into a general, stable trait. However, modern research and theories of cognitive development allow predictions regarding developmental trajectories for metacognitive processes to be made. Metacognitive awareness is suggested to develop around the age of five, while metacognitive skill is not thought to develop until around 11 years of age (Veenman & Spaans, 2005). This view stems from cognitive theories stating that by age of 11, typically, children should be able to realise that their own thoughts can influence their performance on a task (Alexander, Carr, & Schwanenflugel, 1995; Miller & Weiss, 1982). Moreover, Flavell, Miller, and Miller (1993) suggest that at this age, thinking becomes a conscious and reflective metathinking – where a child begins to think about thinking itself, rather than about the objects of thinking (see also Veenman & Spaans, 2005). Thus, self-monitoring capacity, initiated by the Self-Confidence factor, is expected to be more finely developed in children by age 11.

The present study aims to clarify the existence of the self-confidence trait in children. Accordingly, the study examines children in Grades 4 and 6 (aged 9–11 and 11–12 years, respectively), that is, grades with endpoints at the age of 11 (the age suggested as being developmentally important for shaping metacognitive processes), thus allowing for a study of developmental trends of the self-confidence trait.

1.4.4 Gender

As a Predictor

Findings from previous research investigating the differences between males and females and their levels of the Self-Confidence factor have been mixed. While some researchers (Pallier et al., 2002) have found that females have lower levels of confidence than males, others (Stankov, 1999) argue that there are no gender differences in confidence judgments. Further research is needed to determine the link between gender and self-confidence in children. Thus, gender is included as a possible predictor variable in the present study. Yet, given that the existing evidence is mixed, no directional predictions were made.

As a Control Variable

Gender-specific patterns generally exist among intelligence and academic performance results, particularly within tests involving the application of mathematical skills, where boys tend to outperform girls (Geary, 2006). Thus, similarly to age, gender was statistically controlled for, allowing results examining predictive relationships between self-confidence and aptitude to be interpreted irrespective of gender.

1.4.5 Parent-Child Bonding

The bond between a parent and a child is the most common affectional bond in the human relationships. It significantly impacts upon many facets of human life, particularly childhood development (Bowlby, 2005). Parent-child bonds provide a child with a stable foundation, upon which they can confidently explore the world (Bowlby, 1970, 2005; Parker, 1990). Conversely, the disruption of a secure parent-child bond is known to have adverse affects on a child's development; research has consistently linked poor emotional attachment and lack of security between parent and child to psychiatric disorders in childhood (Berk, 2003; Bowlby, 1970, 2005; Parker, 1990) as well as more generalised dysfunctional cognition; whereby poor attachment precipitates the development of dysfunctional schemas about the self, in turn developing to negative cognitive self-statements that are ineffective when dealing with stressful life situations (Ingram, Overbey, & Fortier, 2001). Moreover, poor attachment to parents has been shown to have significant negative correlations with language development (Van IJzendoorn, Dijkstra, & Bus, 1995); communication and cognitive engagement (Moss & St-Laurent, 2001); and academic competence (Diener, Isabella, Behunin, & Wong, 2007). Thus, parent-child dynamics are a significant source of a child's attainment of social competence and an important factor in their cognitive development.

Given this link between parental bonds and a child's optimal cognitive development, it is expected that parent-child bonds will have a significant impact upon metacognitive development. To date, only one study has empirically linked metacognitive self-confidence ratings with parent-child bond dynamics. This study, by Want and Kleitman (2006), focused on parental levels of care and overprotection. *Care* reflects the level of warmth and affection a parent displays to their child, versus the level of coldness and rejection; and *overprotection* refers to the level of excessive control and intrusiveness a parent exhibits versus the level of autonomy or freedom a child has in the relationship (Parker, Tupling, & Brown, 1979). Low levels of care and high levels of overprotection are regarded as unhealthy in terms of optimal child development, as both are reported as predisposing factors in the onset of "most psychiatric conditions" (Parker, 1990, p. 281; see also Higgins & Silberman, 1998; Pomeranz & Ruble, 1998).

The Want and Kleitman (2006) study sampled the adult population and showed that individuals, who retrospectively reported higher maternal overprotection in their childhood, had lower levels of self-confidence. However, the link between parental bonding and self-confidence in children is yet to be investigated. The investigation presented in this chapter intended to examine this link.

1.4.6 Metacognition and Education

Knowing the limits of one's own knowledge and being able to regulate that knowledge, are two essential components of self-regulated and successful learning (Schraw, Crippen, & Hartley, 2006). If students are aware of their own strengths and weaknesses and can apply such knowledge to their learning, they have the means to improve their cognitive achievements. For example, if a student knows of being weak in a particular subject area, he/she could plan to spend more time studying it. In a test-taking situation, if a student is unsure that an answer is correct, he/she knows to come back and check it if time permits. In the realm of education, students who are aware of, control, and reflect about their own thinking, are referred to as self-regulated learners (Zimmerman, 1990).

1.5 The Present Study

It is consistently demonstrated that self-regulated learners outperform their non-reflective counterparts in academic performance measures (Butler & Winne, 1995; Pintrinch & De Groot, 1990). However, empirical evidence regarding the link between the trait of self-confidence and real-life academic achievements is scarce. In fact, research examining the predictive power of self-confidence on any psychological and educational factors is limited (for reviews, see Stankov, 1999; Stankov & Kleitman, 2008). Consequently, just as the predictive factors influencing levels of self-confidence in children are unknown, so too is the predictive nature of the Self-Confidence factor itself. The present study examined the predictive nature of the Self-Confidence factor on real-life, school-based achievements.

It was hypothesised that students with high levels of self-confidence will have greater school achievement outcomes. However, this relationship should be approached with caution, as it is also possible that good academic achievements result in having more self-confidence. In fact, it is quite possible that both of these relationships co-exist. While causal links may not be determined in the present study (in fact, this study focuses on predictions only), for purposes of data analyses it was hypothesised that higher levels of self-confidence predict greater school achievement levels, and not vice-versa (Hypothesis 1). This is expected due to the time precedence of development of the Self-Confidence factor to *current* school achievement marks.[3]

[3] If self-confidence is an intrinsic trait, thus, similar to personality and intelligence, it is a more stable characteristic than academic achievements at a given time.

In summary, the present study had three broad aims. First, it aims to test the factorial stability of the self-confidence construct in primary school-aged children in Grades 4 and 6. Second, the study aimed to clarify the predictive power that intelligence, age, gender and parent-child bonding patterns have on levels of self-confidence and academic achievement. Finally, the study aimed to examine the predictive nature of self-confidence on school academic achievements, whilst controlling for cognitive ability, gender and age.

1.5.1 Hypotheses

The respective hypotheses are listed below.

Hypothesis 1: The Self-Confidence factor would exist as a distinct broad factor in children across all ages; however, the stability of the factor is expected to be more apparent in children in Grade 6 rather than in Grade 4.

Hypothesis 2: Intelligence should positively predict self-confidence and academic performance.

Age is hypothesised to be an important predictor of levels of self-confidence and academic achievement.

Hypothesis 3a: No directional predictions are made in relation to self-confidence and they will be clarified in the present study.

Hypothesis 3b: Age is predicted to share a positive relation with performance on the test of Gf.

Hypothesis 3c: Age is predicted to share a positive relation with achievement.

Hypothesis 4: Gender is hypothesised to be a possible predictor variable for self-confidence and achievement. However, given that the existing evidence is mixed, no directional predictions are made.

Hypothesis 5: Higher levels of parental overprotection will predict lower levels of self-confidence and achievement.

Hypothesis 6: Higher levels of parental care will predict higher levels of self-confidence and achievement.

Hypothesis 7: Higher levels of self-confidence will predict higher achievement.

For each variable, the above relations are hypothesised to exist incrementally to the other variables considered in this study.

1.5.2 Statistical Analyses

To investigate the first aim (see Hypothesis 1), confirmatory factor (CFA) analysis was performed (see Fig. 14.2 in Sect. 3). The theoretical model predicted two latent factors, namely Accuracy and Self-Confidence. To investigate the two latter aims (see Hypotheses 2–7), path analysis was utilised. The path model is presented in Fig. 14.3 (see Sect. 3). In this model the independent (exogenous) variables were

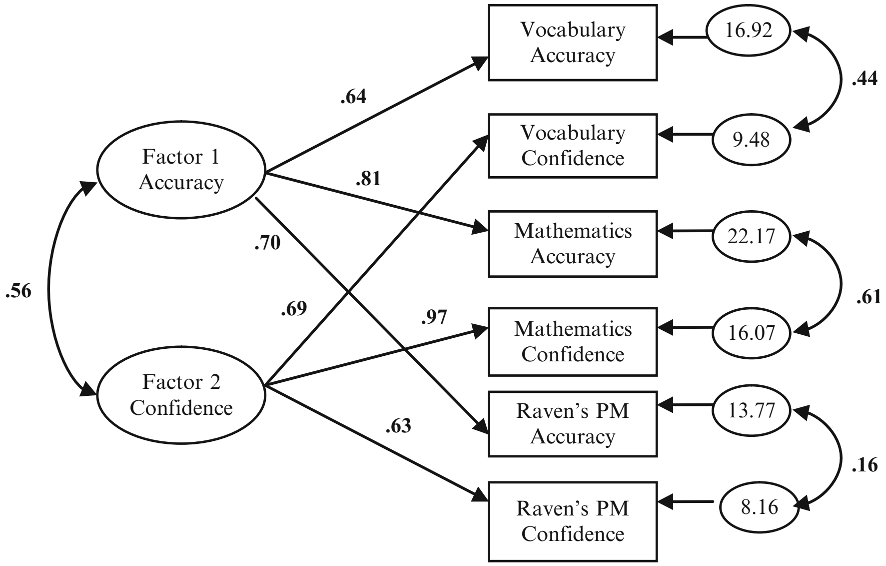


Fig. 14.2 Results of the confirmatory factor analysis (Model 3)

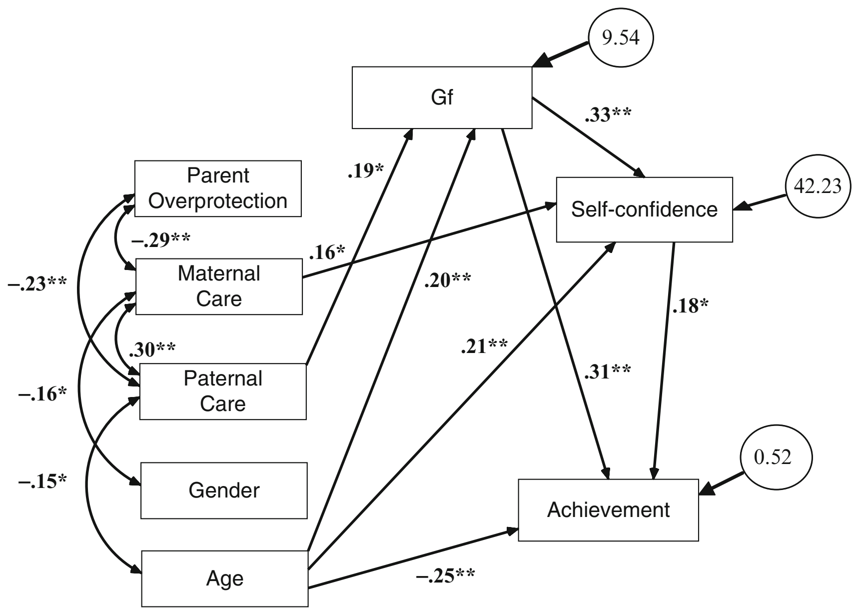


Fig. 14.3 The path analysis model. *Note:* Only significant regression weights (standardised) and correlations are presented. *Curved lines* with *double-headed arrows* represent correlations (Pearson *r*); *straight unidirectional lines* represent regression estimates (betas and gammas).The confidence variable is a total of the average of the confidence ratings from the Vocabulary and Mathematics tests only, and the achievement composite is the sum of teachers' ratings of standardised grades for mathematics, spelling and reading.*$p<0.05$; **$p<0.01$

parental rearing styles (care and overprotection assessed for each parent), gender, and age. The dependent (endogenous) variables were achievement (indexed by grades) and self-confidence (the first-order factor of the CFA model). In accordance with the outlined theoretical model and hypotheses, relationships between the variables achievement and self-confidence were considered while controlling for intelligence (see Fig. 14.3). That is, all possible relationships between exogenous variables and intelligence were built in the path model in addition to all possible relationships between exogenous variables, self-confidence, and achievement (see Fig. 14.3). Finally, as hypothesised, relationships between age and gender with achievement were investigated. Path analysis enabled the investigation of all the abovementioned relationships simultaneously. The word "effect" may be used only for the sake of simplicity, and referring only to the *predictive* nature of the relationships between the different constructs.

2 Method

2.1 Participants

Participants in the study were 197 primary school students; 93 students from Grade 4 and 104 from Grade 6. Participation was voluntary. There were three students who were absent from school for a whole day of testing, six students who had substantial incomplete or missing data, and five students who voluntarily withdrew from the study. These 14 students were eliminated from all the analyses producing the final sample of 183 participants in total (Grade 4 = 85; Grade 6 = 98, 101 males). Students ranged in ages from 9 years and 1 month to 12 years and 11 months.

Participants' age ranged from 9 years 1 month to 12 years 11 months ($M = 10.4$ years, $SD = 1.07$). In Grade 4 students' ages ranged from 9 years 1 month to 11 years ($M = 9.4$, $SD = 0.47$) and in Grade 6 it ranged from 11 years to 12 years 11 months ($M = 11.31$, $SD = 0.43$).

Each participant was enrolled in a mainstream (general ability) class at co-educational public school within the Western Sydney region (New South Wales Department of Education [NSW DET], 2007).[4] Ethics approvals for this research were gained from both Human Research Ethics Committee (HREC) of the University of Sydney and the State Education Research Approval Process (SERAP) for New South Wales, Australia.

[4] Schools within the same region were asked to participate in order to control for socioeconomic status (SES). Additionally, to control for fluency in English, schools with a high enrolment of NESB (Non-English Speaking Background) students were not approached to participate.

2.2 Measures

2.2.1 Parental Bonding Instrument: Brief Current

The Parental Bonding Instrument – Brief Current (PBI-BC) (Klimidis, Minas, & Ata, 1992a) is an 8-item version of the original 25-item PBI (Parker et al., 1979). The PBI has been extensively validated; however, there are important limitations to the retrospective reports. The PBI requires adults to report on rearing practices which occurred years ago, measuring only adults' recollections of the events that took place in their childhood. Thus, their objectivity and their accuracy are suspect. Recognising this limitation, the PBI-BC is a psychometrically validated brief current version (Klimidis et al., 1992a, 1992b). It measures the same two parenting dimensions, with high care and low overprotection reflecting healthy parent-child relations. The instrument allows for real-time reflections of parent-child bonding to be measured, rather than retrospective recollections measured on the original PBI. The authors report that these eight items have reasonable reliability indices. Specifically, Cronbach's alphas were 0.80 and 0.72 for the paternal care and overprotection subscales, respectively, and 0.75 and 0.72 for the equivalent maternal subscales. Thus, the PBI-BC was used to measure children's perception of current patterns of parental bonding behaviours, for each parent, over the past 3 months. The students had to evaluate the extent to which each statement described their concurrent family dynamics using a three-point Likert scale ranging from 1 (never) to 3 (usually). For example, "My mother/father tries to control everything I do"; "My mother/father makes me feel better when I am upset". Higher scores on each scale reflect greater perceived levels of that dimension within the respective parent-child relationship.

Thirteen participants provided answers based on a step-parent's behaviour rather than the indicated biological parent (step-father $n = 11$; step-mother $n = 2$). This data was included as a bonding source in the current study. Ten participants provided data for one parental figure only (no paternal figure $n = 9$; no maternal figure $n = 1$). Such responses were treated as valid data (Amato, 1993). The remaining participants provided responses for a maternal and paternal biological parent; of these cases, 49 participants had biological parents who lived apart.

2.2.2 Standard Raven's Progressive Matrices Test

The Raven's Progressive Matrices (RPM) (Raven, 1938; 60 items) test is a non-verbal test of abstract reasoning that has been consistently and reliably used as a measure of fluid intelligence (Gf) over the past 40 years, with reliability estimates generally ranging between 0.76 and 0.87 (Raven, Raven, & Court, 2003). The RPM test requires individuals to select the piece of puzzle that correctly completes a larger pattern. For the first two sets there are six possible options to choose from, while for the final three sets, the difficulty increases and there are eight options. The mean accuracy score calculated for the test represents the overall percentage of items answered correctly. The high reliability estimate for the RPM (Cronbach's $\alpha = 0.88$ for the overall sample) is consistent with its well established psychometric properties (Raven, Raven, & Court, 2003).

2.2.3 Class Test

It covered two subject areas, vocabulary (Synonym Vocabulary test) and mathematics. The tests were assembled by the researchers based on the NSW school curriculum. The spectrum of item difficulty was broad in order to accommodate for the achievement levels of both grades. Each question was multiple-choice with four-response alternatives. The mean accuracy score computed for each test represents the overall percentage of items answered correctly.

Synonym Vocabulary Test

It is a 16-item test. Students were asked to select, from four possible alternatives, which word is closest in meaning to the keyword. Example item is "The word SMART means the same as A (CLEVER; correct answer); B (SILLY); C (SLOW); D (NICE)". Items were a combination of mainstream curriculum and high ability items. The high-level items were taken from an academic selective test, designed to discriminate between high achieving Grade 5 students seeking placement in an advanced class for Grade 6 (NSW DET, 2003). The remaining ten items were designed aiming at an age-appropriate difficulty level determined by the school syllabus (Board of Studies of NSW [BOS NSW], 2003, 2007). Reliability (Cronbach's $\alpha = 0.71$) for the overall sample was reasonable.

Mathematics Test

It is a 19-item test that was designed around the mathematics curriculum outcomes for both Grades 4 and 6, and required the application of a broad range of mathematical skills, such as numerical and basic geometrical calculations. Nine items were adopted from an Opportunity Class test (NSW DET, 2003). Examples items are: "What number is missing in the number sentence $6 \times \square = 36$? Answers: A 4; B 10; C 6 (correct answer); D 2; "Julie buys some boxes of oranges for \$190 and sells them for \$220. If she makes profit of \$5 on each box, how many boxes did she sell?" A four; B six (correct answer); C eight; D nine. No calculators were permitted during testing sessions. Reliability (Cronbach's $\alpha = 0.69$) for the overall sample was reasonable.

2.2.4 Confidence Rating Scales

Confidence ratings were collected in the RPM, Synonym Vocabulary, and Mathematics tests. These tests contained multiple-choice questions with four, six or eight response choices. Immediately after completing each item, students were asked to rate how confident they were that they had chosen the right answer. The confidence rating scales included both numerical and verbal statements and were based on the culmination of the prior works of Allwood et al. (2006), Clarke (1990), Roebers and Howie (2003), and Schwarz and Roebers (2006) (see Fig. 14.1). Proportions were also included to highlight that due to the question format (multiple-choice), children had a chance of correctly answering the questions by guessing or eliminating some of the alternatives.

To ensure all participants received the same information, standardised instructions for confidence rating (CR) were used. Prior to testing, the researchers ensured that students understood confidence as being how sure they were, and then explained the meaning of each possible response option on the confidence rating scale, from "very unsure" to "very sure". To minimise socially desirable responding, the instructions reiterated that there was no one correct way to respond; that different people would have different levels of confidence, and that it was acceptable to be very confident, not very confident, or anywhere in-between. Averaged confidence scores were then calculated for each cognitive task, with higher values reflecting higher levels of confidence. This resulted in three confidence scores for each participant, one each from the Vocabulary, Mathematics, and RPM tests.

Reliability estimates for confidence scores were uniformly high (for the overall sample ranging from Cronbach's $\alpha = 0.84$–0.96) and were all consistently higher than the reliability estimates for accuracy scores from the same test. These results are consistent with research in adult populations (Kleitman & Stankov, 2007) and offer initial support for the stability of confidence ratings in children.

2.2.5 Achievement Scores

Standardised class marks for mathematics, reading, and spelling were collected from relevant class teachers who were naïve to the aims of the study. These marks reflected individual student achievements within the current school year, relative to their peers from the same grade level across New South Wales, Australia according to standards that are set by the NSW DET and the BOS NSW, Australia. Achievement scores were collected as either an A to E mark, based on New South Wales common rankings, or as a percentage. Both were then converted to a final score ranging from 1 to 5, such that higher scores reflected higher levels of achievement for all data analyses.

2.3 Procedure

All testing took place within the school and it was administered to small groups of 15–30 students during 3 days. This extended procedure was utilised to ensure minimal disruption to school activities, student learning and to avoid cognitive strain on the participants. All instruments were given in a pen-and-paper format and standardised instructions were given prior to each session. No time restrictions were applied, although 1 h was the maximum time required for any single testing session.[5] The PBI-BC was completed on the first day of testing, the Class test on day two, and the RPM test on the third day. The Class test was given before the RPM test to allow students to become familiar with the simpler four-point confidence rating scale, before introducing the more complex 6- and 8-item CR scales.[6]

3 Results

3.1 Missing Value Analysis (MVA)

Prior to all analyses, any other missing data within tests was imputed using the Expectation Maximisation (EM) method in the SPSS 15.0. The EM iterative algorithm provides estimates of imputed values for missing data on the basis of the Maximum Likelihood (ML) procedure; and is a superior method of imputation that offers minimal discrepancy from the original covariance matrix (Little & Rubin, 1989). For ML to be employed, the following three requirements must be met. First, the percentage of missing data needs to be small (less than 5%). Second, the missing data must be identified by the researcher as continuous and multivariate normal in the absence of missing data. Finally, the pattern of any missing data must be random (Byrne, 2001). This was the case with the current data. A small percentage of *meaningful* missing data was evident for the PBI for the participants who did not have a paternal ($n=9$) or maternal figure ($n=1$). These values were not imputed.

3.2 Descriptive Statistics and Reliabilities

Reliability estimates for each test (Cronbach's α) and descriptive statistics for both accuracy and confidence scores are reported in Table 14.2 for the overall sample, and by each grade.

For the overall sample, the mean accuracy for each of the cognitive tasks was high, namely 61.24% for RPM, 80.69% for Vocabulary, and 60.93% for the Mathematics component. Not surprisingly, Grade 6 students performed better than Grade 4 students, and their confidence levels were also higher. The average

[5] No prior research has established additional time needed to incorporate confidence scores. Thus, although the RPM test has time limits to enable the inclusion of confidence ratings in the test, they were not applied. Consequently, the norms of the test were not applicable.

[6] The copies of the response categories for 6- and 8-point confidence rating scales are available from the first author.

Table 14.2 Reliability estimates and descriptive statistics for cognitive tests

	Overall sample ($n = 183$)			Grade 4 ($n = 85$)			Grade 6 ($n = 98$)		
	Cronbach's α	*M*	*SD*	Cronbach's α	*M*	*SD*	Cronbach's α	*M*	*SD*
Accuracy									
RPM	0.88	61.24	13.24	0.91	58.77	13.47	0.89	63.38	12.73
Vocabulary	0.71	80.69	14.74	0.70	75.35	15.70	0.67	85.33	12.14
Mathematics	0.69	60.93	16.42	0.65	54.89	15.77	0.68	66.17	15.19
Confidence									
RPM	0.96	87.93	10.61	0.96	85.37	12.57	0.93	90.15	7.97
Vocabulary	0.84	85.39	11.28	0.87	83.48	12.95	0.76	87.05	9.35
Mathematics	0.91	80.66	13.98	0.92	76.99	15.88	0.88	83.85	11.23

Note: RPM = Raven's Progressive Matrices test; vocabulary = synonym vocabulary test; math = mathematics test

Confidence scores across tasks ranged from 80 to 88% for the overall sample, from 77 to 85% in Grade 4, and from 84 to 90% in Grade 6. These results indicate that the difficulty level of each test was within the participant's cognitive limits, and that children were adjusting their confidence levels to the level of their performance across the grades. Accordingly, the differences between the overall confidence and accuracy scores (Over-/Underconfidence Bias scores[7]) were reasonably stable across the grades. Specifically, the differences were 26.60, 8.13, 22.1 in Grade 4 and 26.77, 1.72, 17.68 in Grade 6 for RPM, Vocabulary, and Mathematics tests, respectively. That is, the difference between the grades in these Bias scores was negligible for the RPM test (−0.17, $p > 0.05$) and small, yet statistically significant, for the Vocabulary and the Mathematics tests (6.41, $p < 0.01$ and 4.42, $p < 0.05$, respectively).

3.3 Confirmatory Factor Analysis

To investigate the structure of cognitive and metacognitive measures, a confirmatory factor analysis (CFA) was carried out using the Maximum Likelihood (ML) method from the AMOS 7 program (Arbuckle, 2006). Analyses were based on the accuracy and confidence scores derived from the RPM, Vocabulary, and Mathematics tests. Previous research suggests that if tests of a similar nature are given to adults, when factor analysis is performed, there would be two separate factors – Accuracy and Self-Confidence. To investigate whether the same holds within a child sample, three models were examined. Model 1 was a one-factor model, in which all scores were combined to define one broad Accuracy/Self-Confidence factor. Model 2 was a two-factor model, in which one factor was defined by all accuracy scores (Cognitive Abilities factor), and the second factor was defined by all confidence scores (Self-Confidence factor). Model 3 was based on the two-factor model theory, with its error terms within the same cognitive test correlated.

Chi-square (χ^2) is one of the most commonly used fit indexes. Small values relative to the degrees of freedom indicate statistically nonsignificant differences between the actual and the implied matrixes, signalling no discrepancy between the hypothesised model and the data. However, this statistic is sensitive to sample size. Thus, following the current practice, the root-mean-square error of approximation (RMSEA) and its 90% confidence interval (90% CI) were used to assess approximate goodness of model fit in the population; values lower than 0.05, with a narrower confidence interval, suggested good fit (Hu & Bentler, 1999). The relative

[7] Over-/Underconfidence Bias score is a difference between mean of confidence ratings and percentage of correct responses across all test items. Overconfidence is reflected via a positive bias and underconfidence is reflected by a negative bias. Confidence judgments are considered to be more realistic when the bias score approaches zero. As a rule of thumb, if the bias score lies within a ±5 limit, it is assumed to have little psychological significance and is argued to reflect a reasonably good calibration (Stankov, 1999).

likelihood ratio of χ^2 to degrees of freedom (χ^2/df) statistic is also reported; values less than 2 are considered to indicate good fit. In addition, Goodness of Fit Index (GFI) was used to reflect the relative amount of covariance accounted by the model, where values 0.90 and above 0.95 suggest acceptable and good fits, respectively (Hu & Bentler, 1999). Finally, the Tucker-Lewis index (TLI) and Comparative Fit index (CFI) were used, which are incremental fit indexes that have been shown to be relatively independent of sample size (Marsh, Balla, & McDonald, 1988). Values greater than 0.90 and 0.95 are considered to reflect acceptable and good fits, respectively (Hu & Bentler, 1999).

When comparing two different models, two things are important: the overall improvement in the fit indices as well as the statistical significance of the changes in the χ^2 statistics ($\Delta\chi^2$) relative to changes in degrees of freedom (Δdf) or $\Delta\chi^2/\Delta$df. The statistically significant *p* value for the latter statistic indicates a significant improvement for the postulated nested model, hence signalling the model's better fit (Byrne, 2001).

3.3.1 Evidence for Broad Confidence and Cognitive Processes

At first, Models 1 and 2 were fitted to both Grade 4 and Grade 6 data separately. Results demonstrated near identical model fits for each grade; thus the data was then combined and a single overall model (Model 3) was applied.[8] Table 14.3 summarises the fit indices statistics for the three models.

As expected, the one-factor model (Model 1) did not adequately describe the self-monitoring data. While Model 2 represented a statistically significant improvement to Model 1, $\Delta\chi^2/\Delta df = 103.68$, $p < 0.01$, it still had a poor fit and was not an acceptable representation of the current data. Thus, the theoretical model (Model 3) was tested with correlations of error terms within each test. By employing this method, the fit of Model 3 was significantly improved, $\Delta\chi^2/\Delta df = 14.93$, $p < 0.01$. Moreover, the majority of the fit indices were within the ranges that signal a good model fit, $\chi^2/\text{df} = 3.23$, GFI = 0.97, TLI = 0.92, and CFI = 0.97. Note, however, that RMSEA and its CI were still greater than the desirable maximum (RMSEA = 0.11, $0.05 < 90\%\ \text{CI} < 0.17$). This demonstrates that although most of the goodness-of-fit statistics are within the ranges that signal a good model fit, the model might be

Table 14.3 Goodness of fit indices for the three CFA models

	χ^2	*df*	$\Delta\chi^2/\Delta df$	χ^2/df	GFI	TLI	CFI	RMSEA	90% CI
Model 1	134.62**	9	–	14.96	0.82	0.52	0.71	0.28	0.24–0.32
Model 2	60.94**	8	103.68**	7.62	0.91	0.77	0.88	0.19	0.15–0.24
Model 3	16.14**	5	14.93**	3.23	0.97	0.92	0.97	0.11	0.05–0.17

**$p < 0.01$

[8] Results of CFA models performed on each grade are available on request.

problematic when it is generalised to a different sample. This might be a direct result of having a limited number of tests employed in this study (only three). Overall, however, when the error terms from the same cognitive task were correlated, the two-factor accuracy/confidence model had a reasonable fit. Model 3 was accepted as the model with the best fit to the data (see Fig. 14.2).

All loadings were statistically significant ($p < .01$) and were high, ranging from 0.63 to 0.97. All communality statistics (available on request) ranged between 0.40 and 0.93, indicating that these variables share a meaningful percentage of variance in common with the extracted factors (Byrne, 2001). As expected, Model 3 supports the existence of two broad factors: (a) Factor 1: *General Ability*. As expected, this factor was defined by the Accuracy scores from the RPM, Vocabulary, and Mathematics tests. It is a broad factor in terms of the cognitive processes that are captured. Although the Vocabulary Accuracy score (the only marker of Gc) had a high loading on this factor, the loadings from the RPM and Mathematics tests were more pronounced, indicating that Gf was reflected more in this Ability factor due to the reasoning processes captured in the latter two tests. (b) Factor 2: *Self-confidence*. As with adults, a distinct Self-Confidence factor exists among the current sample of children. This factor is exclusively defined by the high loadings of the Confidence scores from all three cognitive tests.

3.3.2 Evidence from Parental Care and Overprotection

The reliability coefficients (Cronbach's α), descriptive statistics and correlation coefficients for the PBI-BC are displayed in Table 14.4.

Research based on adolescent populations report reliability coefficients (Cronbach's α) of at least 0.70 for the Care and Overprotection scales in the

Table 14.4 Reliability estimates, descriptive statistics and correlation coefficients (Pearson *r*) for the Parental Bonding Instrument (PBI) subscales

	Cronbach's α						
	PBI-BC Original	PBI-BC Adjusted	*M* (*SD*)	PC	MO	PO	Parent overprotection
Maternal care	0.49	0.71	8.23 (1.23)	0.30**	–0.34**	–0.14	–0.29**
Paternal care	0.59	0.70	7.84 (1.50)	1	–0.09	–0.30**	–0.23**
Maternal overprotection	0.40		6.88 (1.74)		1	0.57**	0.90**
Paternal overprotection	0.35		6.61 (1.67)			1	0.88**
Parental overprotection		0.62	6.73 (1.52)				1

PC paternal care; *MO* maternal overprotection; *PO* paternal overprotection

Possible scores for care range from 3 to 9. Possible scores for overprotection range from 4 to 12

$^{**}p < 0.01$

PBI-BC (Klimidis et al., 1992a). In this study, however, when the instrument was used with younger children, reliability estimates were affected, ranging from 0.35 to 0.59 (see Table 14.4). Notably, one question on the Care scale "My mother/father seems emotionally cold to me" was misunderstood by the present cohort of participants. During the testing procedure, children often asked the researchers to explain what the word "cold" meant. The statistics confirmed concerns associated with this item (Question 2), and it was deleted from the scale for all major analyses, resulting in reliability increases from 0.49 to 0.71 for the Maternal Care subscale and from 0.59 to 0.70 on the Paternal Care subscale (see Table 14.4). The Maternal and Paternal Overprotection subscales returned low reliabilities (0.35 and 0.40, respectively). Moreover, looking at the correlation coefficients for the PBI subscales, the correlation coefficient between the Maternal and Paternal Overprotection subscales was reasonably high ($r = 0.57$, $p < 0.01$). This finding identified a possible multicollinearity problem if both were to be used simultaneously in path analysis. To remedy both of these problems, the Maternal and Paternal Overprotection subscales were combined to create a Parental Overprotection scale. In doing so, the reliability coefficient of the composite 8-item scale improved to 0.62.

The mean levels of care reported for both mothers and fathers in the present sample were high ($M = 8.23$, $SD = 1.23$ and $M = 7.84$, $SD = 1.50$, respectively) and reflected greater perception of care rather than rejection within each parent-child relationship. Perception of levels of maternal care was higher than that of paternal care, indicating that mothers were perceived as more caring than fathers. Both of these results were consistent with previous research findings (Klimidis et al., 1992a; Parker et al., 1979). Overall, children reported low and similar levels of overprotection for mothers and fathers (Maternal $M = 6.88$, $SD = 1.74$; Paternal $M = 6.61$, $SD = 1.67$). This pattern was consistent with previous research findings (Klimidis et al., 1992a; Parker, 1983, 1990), and it indicated that students felt they were in autonomous rather than controlling relationships and this perception was similar for both parents. This pattern also confirmed the decision to combine Maternal and Paternal Overprotection subscales into the Parental Overprotection scale. From this point on, any reference to Overprotection refers to the Parental Overprotection composite scale.

3.3.3 Evidence from Achievement Measures

Table 14.5 summarises descriptive statistics and correlation coefficients for the achievement scores.

The levels of achievements ranged between 2.51 and 2.97 (out of 5) and were similar across grades and across different subject-matters. Importantly, there was a pattern of strong positive correlations present between achievements in Mathematics, Spelling, and Reading (ranging from 0.69 to 0.84, $p < 0.01$). Thus, the scores were combined into the single Achievement composite.

Table 14.5 Descriptive statistics and correlation coefficients (Pearson *r*) for achievement scores

	Descriptive statistics						Correlations	
	Overall sample ($n=184$)		Grade 4 ($n=86$)		Grade 6 ($n=98$)		Overall sample ($n=184$)	
Tests	*M*	*SD*	*M*	*SD*	*M*	*SD*	Spelling	Reading
Mathematics	2.83	0.91	2.66	0.90	2.97	0.90	0.71**	0.69**
Spelling	2.68	0.89	2.51	0.96	2.83	0.80	1	0.84**
Reading	2.66	0.95	2.49	0.98	2.82	0.90		1

**$p<0.01$

3.4 *Path Analysis*

To investigate the hypotheses two to seven of the present study, path analysis was conducted using the Maximum Likelihood (ML) method from the AMOS 7 program (Arbuckle, 2006) using the correlation matrix summarised in Table 14.6.

Relationship between independent and dependent variables is referred to as beta (β), while relationship between dependent variables is referred to as gamma (γ).[9] Prior to examination of the betas and gammas, correlations between independent variables that were not statistically significant ($p>0.05$) were fixed to zero; thus only significant correlations are reported (see Fig. 14.3).[10] Then, the relationships between independent and dependent variables were determined.

All possible regression coefficients (betas and gammas) were built into the model. This was done to insure that the effects of each variable on self-confidence and achievement were calculated while statistically controlling for known common causes (intelligence, age and gender; see hypotheses above). The only exception was the variable indexing parent-child relationship dynamics which in this model cannot be classified as the "common causes". However, as existent research data did not allow exact predictions in regards to these variables and self-confidence and achievement, the path model included all possible relevant regression paths. For the sake of clarity, only significant coefficients ($p<.05$) are displayed in Fig. 14.3. The focus is on the discussion of significant *direct* effects (the effect one variable has on another without any intervening variables). Path analysis also allows calculations and interpretations of the indirect effects (the effect a variable has on another via a third intervening variable within the model). Only most meaningful indirect effects will be discussed here.

In the path analysis model, the independent (exogenous) variables are parental rearing styles (maternal care, paternal care, and parental overprotection), gender, and age. The dependent (endogenous) variables are the Achievement and Self-confidence scores. Fluid intelligence (Gf) is a common factor affecting these

[9] The unstandardised estimates are available on request.

[10] This is a recommended procedure for complex models examined on a relatively small sample size as it maximises degrees of freedom without affecting the model parameters of fit indices (Byrne, 2001).

Table 14.6 Correlation coefficients (Pearson r) for the variables used in the path model

		1	2	3	4	5	6	7	8	9	10	11	12	13
1	Age	1.00												
2	Gender	0.03	1.00											
3	M care	0.10	−0.13	1.00										
4	P care	−0.10	0.06	0.28	1.00									
5	Parental OP	−0.13	−0.07	−0.28	−0.21	1.00								
6	Voc corr	0.28	−0.06	0.06	0.12	−0.09	1.00							
7	Maths corr	0.34	0.09	0.13	0.20	−0.16	0.54	1.00						
8	RPM corr	0.17	0.08	−0.04	0.14	−0.03	0.44	0.55	1.00					
9	Voc conf	0.17	0.00	0.12	0.17	0.01	0.42	0.38	0.25	1.00				
10	Maths conf	0.28	0.14	0.16	0.14	0.00	0.23	0.55	0.42	0.65	1.00			
11	RPM conf	0.29	0.13	0.13	0.01	0.08	0.08	0.25	0.33	0.42	0.61	1.00		
12	Achievement	−0.15	−0.01	0.03	0.19	−0.03	0.38	0.41	0.35	0.27	0.19	−0.02	1.00	
13	Confidence	0.25	0.08	0.16	0.17	0.00	0.34	0.52	0.38	0.89	0.93	0.58	0.25	1.00
14	Accuracy	0.36	0.02	0.11	0.19	−0.14	0.86	0.89	0.57	0.45	0.45	0.19	0.45	0.50

Note: M care = maternal care; P care = paternal care; Parental OP = parental overprotection; Voc corr = vocabulary test accuracy; maths corr = mathematics test accuracy; RPM corr = Raven's Progressive test accuracy; voc conf = vocabulary test confidence; maths conf = mathematics test confidence; RPM conf = Raven's Progressive test confidence; achievement = achievement composite based on mathematics, spelling and readings scores; confidence = confidence composite excluding Raven's Progressive Matrices confidence score; accuracy = accuracy composite based on vocabulary and mathematics accuracy scores. All correlations higher than 0.15, $p < 0.05$; all correlations higher than 0.19, $p < 0.01$

constructs, thus its influence needs to be statistically controlled for in the investigation of the relationship between achievement and self-confidence. Relationships between dependent variables were also considered to examine the predictive influence that self-confidence has on educational achievements while controlling for Gf (see Fig. 14.3). Similarly, given the assumptions of path analysis, when all variables are incorporated in the model, the impact of each variable represents the impact of the variable that exists after controlling (or partitioning out) the influence of all other variables in the model. The inclusion of achievement as an outcome variable means that the influence of care and overprotection levels as well as gender and age on a child's achievement levels can also be examined. As noted earlier, this study was *not* intended to investigate causality. The words "effect" and "influence" here are used only for the sake of simplicity, and referring only to the *predictive* nature of the relationships between the constructs.

Finally, the confidence score is the sum of the mean confidence judgments provided for the Vocabulary and Mathematics tests only. The confidence score from the RPM test was not included to prevent the problem of statistical dependency (as confidence judgments provided for the RPM test are conceptually and empirically related to the accuracy of actual performance on this test, $r=0.33$, $p<0.01$). Given that the RPM was used as the measure of Gf, if the confidence scores from RPM test were to be included, this would inflate the relationship between Gf and self-confidence. This would impose problematic and misleading interpretations.

3.4.1 Correlations Between the Independent Variables

Care scores were positively correlated ($r=0.29$, $p<0.01$, see Fig. 14.3). Parent overprotection scores had small, yet similar and significant negative correlations with both care scores ($r=-0.29$ and $r=-0.21$, $p<0.01$), indicating that children linked higher levels of parental control to a lesser degree of parental care. Maternal care had a small, yet significant negative relationship with gender ($r=-0.16$, $p<0.05$). There was also a small, yet significant, negative correlation between paternal care and age ($r=-0.15$, $p<0.05$). However, these two tendencies were not pronounced.

3.4.2 Direct Effects

As shown in Fig. 14.3, as expected (Hypothesis 2) Gf positively predicted both dependent variables. Higher levels of Gf positively predicted self-confidence and achievement, indicating that students with greater Gf have greater levels of self-confidence and are achieving better results at school.

Moreover, age significantly predicted all three dependent variables, namely Gf, self-confidence, and achievement (Hypotheses 3a, 3b, 3c). It exhibited a positive effect on both self-confidence and Gf. Thus, older students had greater levels of self-confidence and (as predicted) performed better on the same test of Gf. However, contrary to our expectations (see Hypothesis 3c), age had a negative effect on achievement levels. A separate correlation analysis within each grade was performed to investigate these results further. They revealed that the relationship between age and achievement was negligible, but *positive* within each grade ($r=0.04$ and $r=0.12$, $p>0.05$ in Grade 4 and Grade 6, respectively). Moreover, there was some overlap in ages in each grade, and unusually, older students within grades were performing at a lower level than their younger counterparts. That is, in Grade 6, several older students were judged to be performing, on average, at a lower standard level than expected for this grade. Furthermore, within a framework of a path analysis, the effect of age on achievements was examined after controlling for the Gf of a student. Thus, on the overall sample, this negative beta indicated that after taking into account students' Gf, older students within a grade were judged by their teachers as achieving at a lower "state standard" level than the younger students within the same grade. Thus, the negative relationship does not mean that, on average, the older students have an inferior level of cognitive ability, as the opposite was demonstrated by the positive relationship between age and Gf.

Gender was hypothesised to be a *possible* predictor variable for self-confidence and achievement (Hypothesis 4). However, as shown in Fig. 14.3, gender did not directly predict any of the dependent variables in the model.

We expected (Hypothesis 5) that parental overprotection will predict lower self-confidence and achievement scores. Our results did not support this prediction. In fact, parental overprotection score did not directly predict any of the dependent variables.

As expected in Hypothesis 6, greater levels of maternal care positively predicted higher levels of self-confidence. Thus, children receiving greater levels of care from their mother tend to have greater levels of self-confidence than those children receiving lower levels of maternal care. Contrary to expectations (Hypothesis 6) paternal care did not have the same influence on self-confidence levels. However, paternal care did positively predict Gf, indicating that children who report receiving higher levels of care from their father are exhibiting higher levels of cognitive ability than those students who report receiving lower levels of paternal care.

A notable finding here is the direct positive effect that self-confidence had on achievement. As predicted in Hypothesis 7, those students exhibiting greater levels of self-confidence tend to perform better at school. This prediction holds for both boys and girls of all ages, irrespective of their Gf and parenting bonds.

The path analysis model had a good fit, $\chi^2(5, n=183)=5.23$, $p=0.39$, $\chi^2/df=1.05$, RMSEA$=0.02$ ($0.01<90\%$ CI<0.11), GFI, TLI, and CFI$=0.99$. This model accounted for 6.7% of the variance in Gf, 22.7% in self-confidence, and 21.1% in achievement.

4 Discussion

Metacognition is one of the three fundamentals of self-regulated learning, along with cognition and motivation (Schraw et al., 2006). Efficient test-taking behaviour and test-taking outcomes signify academic success and the metacognitive confidence judgments students assign to their on-going performance are at the core of this test-taking behaviour. The present study sought to identify the crucial ages at which self-confidence judgments begin to emerge as a *habitual* response pattern, or a trait, which is stable across different cognitive tasks. It also aimed to determine predictors of self-confidence, while investigating the predictive validity of self-confidence in school settings.

Our results do not permit to draw definite conclusions as to whether confidence judgments are task- or domain-specific in early childhood and at what age do they develop into the more general, stable trait evident in adults. It was hypothesised that self-confidence would exist as a distinct broad factor in children across all ages, although the stability of the Self-Confidence factor was expected to be more apparent in children aged 11 and over, that is, in Grade 6 rather than in Grade 4. This expectation was rooted in the theories of metacognitive development which stress the importance of age 11 in the development of metacognitive skill when children begin to think about thinking itself realising that these thoughts can influence their performance (Alexander et al., 1995; Flavell et al., 1993; Miller & Weiss, 1982; Veenman & Spaans, 2005). The results indicate that children in each grade exhibit identical trends associated with confidence ratings. Results also demonstrate that self-confidence ratings have high reliability within each test (Cronbach's $\alpha > 0.80$); a level of internal consistency greater than that was found for performance accuracy measured on the same test. Confidence ratings separated clearly from performance accuracy scores, defining a distinct Self-Confidence factor. Thus, self-confidence exists as a stable and identifiable metacognitive factor in children as young as 9–12 years of age, just as it does in adults.

This novel finding signifies that metacognition, in the sense of self-confidence, is a *stable* component of a child's thinking repertoire by Grade 4. Thus, this study provides a foundation for the improvement of teaching at the classroom level. For example, metacognitive self-monitoring skills should be seen as appropriate additions to the classroom curriculum before Grade 4, with an aim to foster these skills before they become habitual. Moreover, knowledge that a child as young as nine is already habitually assessing their own thinking is a crucial and powerful tool, one which can undoubtedly assist both school counsellors and child psychologists. If a child has the capacity to be a self-regulated learner, perhaps he/she has the capacity to self-reflect upon one's feelings and thoughts preceding these feelings. Self-reflective thinking and awareness of one's cognitions are vital skills which can be developed and fostered in the realm of counselling. Future studies need to explore these directions.

The study also aimed to identify the determinants of the self-confidence trait, by examining a key external influence of the early social environment, that is, one's relationships with parents. Want and Kleitman (2006) demonstrated that retrospective reports of high levels of maternal overprotection during childhood negatively predicted self-confidence in the adult population. However, no such studies had previously been conducted with children. The present study was the first to examine the relationship between parental bonds with each parent and the Self-Confidence factor in primary school children. The study was also the first to analyse current, rather than retrospective, reports of parent-child bonding in relation to levels of self-confidence, thus strengthening the validity and reliability of bonding reports and their reported influences on metacognitive development. These influences were studied while controlling for age, gender and fluid intelligence of a child.

The results indicate that irrespective of a child's age, gender, and fluid intelligence, maternal care predicts positively the levels of self-confidence, with higher levels of maternal care associated with greater levels of self-confidence. Similarly, Want and Kleitman (2006) found that maternal bonds, and not paternal bonds, *directly* predicted self-confidence levels in adults. Consistency of these results may form the foundations to suggest that mother-child bonds have a greater influence on metacognitive development than father-child bonds. Perhaps, then, the prediction that maternal care has on levels of self-confidence is intertwined with the vulnerabilities of a child's self-evaluations. It should be also noted that these self-evaluations are more strongly influenced by maternal rather than paternal levels of care. Future studies need to assess a possible mediation that self-concepts may have within the relationships of parental bonds and self-confidence. In terms of predictions, it should be emphasised that paternal care directly predicts fluid intelligence, which itself predicts self-confidence. Therefore, paternal care *indirectly* predicts self-confidence, via its link with fluid intelligence. This result recognises the importance of healthy father-child bonds for a child's optimal cognitive and metacognitive development.

Consistent with the hypothesis and research completed in adult populations, age demonstrated a significant relationship with self-confidence; older children displayed higher levels of self-confidence than their younger counterparts, irrespective of fluid intelligence or gender. Moreover, age positively predicted fluid intelligence, which itself, positively predicted levels of self-confidence. Thus, age has direct and indirect influences on self-confidence levels. While current results demonstrate that children as young as nine have developed stable self-confidence levels, older children on average are more confident, and only some portion of variance is attributable to advances in fluid intelligence.

As predicted, fluid intelligence was a strong positive predictor of the self-confidence composite. This is consistent with prior research in adult populations, where cognitive ability (measured on the same test employing confidence ratings) has been found to predict self-confidence (Kleitman & Stankov, 2007). The present study accounts for this relationship and, therefore, parallel fluid intelligence ratings to confidence ratings were not included in the self-confidence score utilised in the path model. The results still suggested that greater fluid intelligence predicts greater self-confidence. This relationship was not falsely inflated as a result of concurrent achievement, which can often be a weakness of research employing concurrent measures.

Also, as expected, fluid intelligence exhibited a positive influence on school-based achievement reflected by standardised grades. Interestingly, this influence was as strong as the influence fluid intelligence exhibited on self-confidence. Although the present results do not clarify causality of this relationship, it nevertheless demonstrates the importance of studying the developmental link between fluid intelligence and metacognition.

Gender did not predict any of the dependent variables in the model. Previous research has demonstrated mixed results for the role of gender in self-confidence. The present results add weight to the argument that gender does not influence children's self-confidence and achievement.

The most notable finding of the study is the positive relationship between self-confidence and school achievement that is incremental to a child's age, gender, and levels of fluid intelligence. As mentioned earlier, space constraints prevented examination of the over/under-confidence bias scores which index self-monitoring. A separate paper is devoted to this construct and its link to academic achievements (Kleitman & Moscrop, 2009). However, preliminary findings indicate that a smaller discrepancy between confidence and accuracy scores predict better achievements. Together, these findings imply that irrespective of the gender, age, and intelligence characteristics of a child, greater and more realistic self-confidence maximises effective learning. Students exhibiting these trends not only possess the regulative capacity to know what they know and how well they perform, they also utilise their knowledge and skills to learn how to learn.

4.1 Limitations and Future Directions

Contrary to expectations and earlier findings (Want & Kleitman, 2006), overprotection levels within parent-child relationships did not predict self-confidence levels. This result may be attributed to the young age of the participants in the current study. Perhaps, as a young child, one perceives that he/she needs higher levels of overprotection, discipline and direction, as one is not yet engaged in an autonomous, independent lifestyle (Berk, 2003). Another possibility for these results might be linked to the complex nature of parental overprotection which can be classified into two components, namely psychological and behavioural control. Psychological control refers to "attempts to intrude into the psychological and emotional development of the child", while behavioural control refers to "parental behaviours that attempt to control…children's behaviour" (Barber, 1996, p. 3296). It is possible that these two different types of control may hold differential influences on cognitive and metacognitive development. In fact, Bean, Bush, McKearney, and Wilson (2003) found that behavioural control predicted an increase in academic achievement, whereas psychological control predicted their decrease. The use of the PBI-BC prevented us from delineating these control tendencies. There were other concerns with this measure. Reliability issues were raised in the use of the PBI-BC in young child populations, surrounding the complex wording of some of the questions. In fact, one question was removed from the Care scale due to poor reliability statistics. Thus, future studies should examine the two types of control separately, using a more reliable measure of concurrent parental-child bonds.

Although sampling technique in the present study was limited, the resulting sample of the primary school children exhibited the trends similar to those demonstrated in other research in regards to parent-child relationships, cognitive abilities, age and gender. Therefore, the current sample can be seen as an appropriate indicator of the wider population.

The exploratory nature of the present study limits the scope of conclusions drawn. The path analysis model used in the present study focused only on the *predictive* relationships between the variables. A longitudinal research study with a greater control for known common causes (e.g., previous achievement) could greatly assist in determining causal links between these variables. Future research would also benefit from a larger selection of variables to mark each construct to have more than only the bare minimum (three) of potential markers for each latent factor. Moreover, in this study we only controlled for students' fluid intelligence. Ideally, both fluid and crystallized intelligence should be controlled for. Finally, future studies should examine a possible mediation role that certain self-concept measures (see Efklides, & Tsiora, 2002; Kleitman & Stankov, 2007; Kröner & Bierman, 2007) could play in parent-child bonds and in self-confidence relationships.

4.2 Conclusion

While future studies need to investigate the causal nature of the relationships between different constructs examined in this research and earlier ages in an attempt to identify the key age at which decision-making processes become entrenched, this study provides the foundation for identifying the development of habitual self-confidence. The results from this study not only indicate that self-confidence exists as a stable construct in children as early as 9 years of age, they also shed light on the predictive validity of the self-confidence trait in school settings. The results also extend the understanding of the factors which predict children's cognitive and metacognitive development and academic outcomes from the family unit. This knowledge offers great promise to educators, psychologists and parents alike, providing them with the potential to foster growth of decision-making abilities of children with a broad aim to improve their educational outcomes.

References

Alexander, J. M., Carr, M., & Schwanenflugel, P. J. (1995). Development of metacognition in gifted children: Directions for future research. *Developmental Review, 15*, 1–37.

Allwood, C. M., & Granhag, P. A. (1999). Feelings of confidence and the realism of confidence judgements in everyday life. In P. Juslin & H. Montgomery (Eds.), *Judgement and decision making: Neo-Brunswikian and process-tracing approaches* (pp. 123–146). Mahwah, NJ: Erlbaum.

Allwood, C. M., Granhag, P. A., & Jonsson, A. C. (2006). Child witnesses' metamemory realism. *Scandinavian Journal of Psychology, 47*, 461–470.
Amato, P. R. (1993). Children's adjustment to divorce: Theories, hypotheses, and empirical support. *Journal of Marriage and the Family, 55*(1), 23–38.
Arbuckle, J. L. (2006). *AMOS 7.0 users guide.* Chicago, IL: SPSS Inc.
Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.
Barber, B. K. (1996). Parental psychological control: Revisiting a neglected construct. *Child Development, 67*(6), 3296–3319.
Bean, R. A., Bush, K. R., McKearney, P. C., & Wilson, S. M. (2003). The impact of parental support, behavioral control, and psychological control on the academic achievement and self-esteem of African American and European adolescents. *Journal of Adolescent Research, 18*, 523–541.
Berk, L. E. (2003). *Child development* (6th ed.). Boston, MA: Allyn & Bacon.
Board of Studies of New South Wales. (2003). *Stages 4 and 5 mathematics syllabus.* Sydney, Australia: Board of Studies of New South Wales.
Board of Studies of New South Wales. (2007). Common Grade Scale K-6. Retrieved 05/02/2007 from the World Wide Web http://arc.boardofstudies.nsw.edu.au/go/k-6/common-grade-scale/.
Bowlby, J. (1970). Disruption of affectional bonds and its effects on behaviour. *Journal of Contemporary Psychotherapy, 2*(2), 75–86.
Bowlby, J. (2005). *The making and breaking of affectional bonds.* New York: Routledge.
Brown, J. D. (1986). Evaluation of self and others: Self-enhancement bias in social judgements. *Social Cognition, 4*, 353–376.
Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245–281.
Byrne, B. M. (2001). *Structural equation modelling with AMOS: Basic concepts, applications, and programming.* Mahwah, NJ: Erlbaum.
Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytical studies.* New York: Cambridge University Press.
Clarke, D. (1990). Verbal uncertainty expressions: A critical review of two decades of research. *Current Psychology, 9*(3), 203–226.
Diener, M. L., Isabella, R. A., Behunin, M. G., & Wong, M. S. (2007). Attachment to mothers and fathers during middle childhood: Associations with child gender, age, and competence. *Social Development, 17*(1), 84–101.
Efklides, A., & Tsiora, A. (2002). Metacognitive experiences, self-concept, and self-regulation. *Psychologia: An International Journal of Psychology in the Orient, 45*, 222–236.
Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911.
Flavell, J. H., Miller, P. H., & Miller, S. A. (1993). *Cognitive development* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
Geary, D. C. (2006). Gender differences in mathematics: An integrative psychological approach. *British Journal of Educational Studies, 54*(2), 245–246.
Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models – A Brunswikian theory of confidence. *Psychological Review, 98*(4), 506–528.
Harvey, N. (1997). Confidence in judgment. *Trends in Cognitive Sciences, 1*(2), 78–82.
Higgins, E. T., & Silberman, I. (1998). Development of regulatory focus: Promotion and prevention as ways of living. In J. Heckhausen & C. S. Dweck (Eds.), *Motivation and self-regulation across life span* (pp. 78–113). Cambridge, UK: Cambridge University Press.
Horn, J. L., & Noll, J. (1994). A system for understanding cognitive abilities: A theory and the evidence on which it is based. In D. K. Detterman (Ed.), *Current topics in human intelligence* (Vol. IV, pp. 151–203). New York: Springer.
Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling: A Multidisciplinary Journal, 6*, 1–55.
Ingram, R. E., Overbey, T., & Fortier, M. (2001). Individual differences in dysfunctional automatic thinking and parental bonding: Specificity of maternal care. *Personality and Individual Differences, 30*, 401–412.

Jonsson, A. C., & Allwood, C. M. (2003). Stability and variability in the realism of confidence judgments over time, content domain, and gender. *Personality and Individual Differences, 34*(4), 559–574.

Juslin, P. (1994). The overconfidence phenomenon as a consequence of informal experimenter-guided selection of almanac items. *Organizational Behaviour & Human Decision Processes, 57*(2), 226–246.

Kleitman, S. (2008). *Metacognition in the rationality debate: Self-confidence and its calibration.* Saarbücken, Germany: VDM.

Kleitman, S., & Moscrop, T. (2009). *Metacognitive self-monitoring and academic achievements in primary-school children: Their relationships and links to parental bonds, intelligence, age and gender*. Manuscript in preparation.

Kleitman, S., & Stankov, L. (2001). Ecological and person-oriented aspects of metacognitive processes in test taking. *Applied Cognitive Psychology, 15*, 321–341.

Kleitman, S., & Stankov, L. (2007). Self-confidence and metacognitive processes. *Learning and Individual Differences, 17*(2), 161–173.

Klimidis, S., Minas, I. H., & Ata, A. W. (1992a). The PBI-BC: A brief current form of the Parental Bonding Instrument for adolescent research. *Comprehensive Psychiatry, 33*(6), 374–377.

Klimidis, S., Minas, I. H., & Ata, A. W. (1992b). Construct validation in adolescents of the brief current form of the Parental Bonding Instrument. *Comprehensive Psychiatry, 33*(6), 378–383.

Kröner, S., & Bierman, A. (2007). The relationship between confidence and self-concept – Towards a model of response confidence. *Intelligence, 35*, 580–590.

Little, R. J. A., & Rubin, D. B. (1989). The analysis of social-science data with missing values. *Sociological Methods and Research, 18*(2–3), 292–326.

Marsh, H., & Kleitman, S. (2002). Extracurricular school activities: The good, the bad, and the nonlinear. *Harvard Educational Review, 72*(4), 464–514.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*(3), 391–410.

Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). *Metacognition: Knowing about knowing.* Cambridge, MA: MIT Press.

Miller, P., & Weiss, M. G. (1982). Children and adults knowledge about what variables affect selective attention. *Child Development, 53*, 543–549.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502–517.

Moss, E., & St-Laurent, D. (2001). Attachment at school age and academic performance. *Developmental Psychology, 37*(6), 863–874.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

New South Wales Department of Education. (2003). *Opportunity class sample tests, 2001; 2002; 2003.* Retrieved 05/02/2007 from http://www.schools.nsw.edu.au/learning/k-6assessments/oc_sampletest.php.

New South Wales Department of Education. (2007). *Department of Education and Training Regions for New South Wales.* Retrieved 05/02/2007 from http://www.nsw.edu.au.

Pallier, G., Wilkinson, R., Danthir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgements. *The Journal of Educational Psychology, 129*(3), 257–299.

Parker, G. (1983). *Parental overprotection: A risk factor in psychosocial development.* New York: Grune & Stratton.

Parker, G. (1990). The parental bonding instrument: A decade of research. *Social Psychiatry and Psychiatric Epidemiology, 25*, 281–282.

Parker, G., Tupling, H., & Brown, L. B. (1979). A parental bonding instrument. *British Journal of Medical Psychology, 52*, 1–10.

Pintrinch, P. R., & de Groot, E. V. (1990). Motivation and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33–40.

Pomeranz, E. M., & Ruble, D. N. (1998). The multidimensional nature of control: Implications for the development of sex differences in self-evaluation. In J. Heckhausen & C. S. Dweck (Eds.), *Motivation and self-regulation across life span* (pp. 159–184). Cambridge, UK: Cambridge University Press.

Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence.* London: Lewis.

Raven, J., Raven, J. C., & Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General overview.* San Antonio, TX: Harcourt Assessment.

Roebers, C. M., & Howie, P. (2003) Confidence judgments in event recall: Developmental progression in the impact of question format. *Journal of Experimental Child Psychology, 85*(4), 352–371.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. Educational Psychology Review, *7*(4), 351–371

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research Science Education*, *36*(1–2), 111–139.

Schwarz, S., & Roebers, C. M. (2006). Age differences in the effects of social influence on children's eyewitness performance and their metacognitive monitoring. *Journal of Experimental Child Psychology, 94*, 229–248.

Stankov, L. (1999). Mining on the "no man's land" between intelligence and personality. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 315–367). Washington, DC: American Psychological Association.

Stankov, L. (2000). Complexity, metacognition, and fluid intelligence. *Intelligence, 28*(2), 121–143.

Stankov, L., & Crawford, J. D. (1996). Age differences in the realism of confidence judgements: A calibration study using tests of fluid and crystallized intelligence. *Learning and Individual Differences, 8*(2), 83–103.

Stankov, L., & Kleitman, S. (2008). Processes on the borderline between cognitive abilities and personality: Confidence and its realism. In G. J. Boyle, G. Matthews, & D. Saklofske (Eds.), *The handbook of personality theory and testing* (pp. 541–555). London: Sage.

Van IJzendoorn, M. H., Dijkstra, J., & Bus, A. G. (1995). Attachment, intelligence, and language: A meta-analysis. *Social Development, 4*(2), 115–28.

Veenman, M. V. J., & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences, 15*, 159–176.

Veenman, M. V. J., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction, 14*(1), 89–109.

Want, J., & Kleitman, S. (2006). Imposter phenomenon and self-handicapping: Links with parenting styles and self-confidence. *Personality and Individual Differences, 40*(5), 961–971.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17.

Chapter 15
Metacognition and Reading Comprehension: Age and Gender Differences

Svjetlana Kolić-Vehovec, Igor Bajšanski, and Barbara Rončević Zubković

1 Metacognition and Reading

The question of whether metacognition is general by nature, or rather task- and domain-specific has not yet been unequivocally answered. Veenman, van Hout-Wolters, and Afflerbach (2006) argued that findings of some studies support the idea of general metacognitive skills, but the domain-specific nature of metacognitive skills is also indicated in studies that use quite different tasks. These studies show that metacognitive skills may initially develop in separate domains and later on become more integrated and applicable to a variety of different tasks and domains. Therefore, an exploration of metacognition in different domains is still required, at least in terms of extending the insight into the metacognitive activities of younger students. Skillful reading is fundamental for children's education. Thus, the investigation of the development of metacognition in the domain of reading remains very important, especially during the period of secondary education when learners progress from acquisition of basic skills to proficiency or expertise at higher levels of schooling.

In the first part of this chapter, we present an overview of the research that has examined the processes involved in reading with emphasis on metacognitive ones. Developmental trends in metacognition and reading comprehension are presented as the basis for the authors' recent work on the differences in children's reading during elementary and high school (9–17 years) in Croatia; gender differences are also taken into account. We argue that age is associated with changes in the efficiency of reading comprehension, as well as of metacognitive knowledge and metacognitive skillfulness, as shown by the interplay between the different aspects of reading. The focus was placed on upper elementary school as a crucial period for the development of metacognition in the domain of reading comprehension, and fully developed reading skill as a consequence. Metacognitive development in reading

S. Kolić-Vehovec (✉)
Department of Psychology, University of Rijeka, Croatia, Slavka Krautzeka b.b, 51000 Rijeka, Croatia
e-mail: skolic@ffri.hr

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*, DOI 10.1007/978-1-4419-6546-2_15, © Springer Science+Business Media, LLC 2010

during this period is related to gender, as was shown in faster metacognitive development of girls compared to boys. However, the pattern of differences varies across components of metacognition. At the end of the chapter, implications for education are discussed.

1.1 Metacognitive Knowledge and the Regulation of Reading

Walczyk (1994) has described three aspects of the complex skill of reading. The first aspect concerns subcomponent processes, which refer to the lexical processes of word recognition, and the post-lexical processes of word, sentence and text comprehension. The second aspect refers to the limited resources of attention and working memory that can be allocated to reading processes, and the third aspect concerns the executive metacognitive component of reading. The majority of researchers in the field of metacognition agree that the metacognitive component includes metacognitive knowledge about reading that relates to a person's knowledge about his or her reading, about the different types of reading tasks and about reading strategies (Baker & Brown, 1984; Paris, Lipson, & Wixson, 1983).

However, there is no such clear consensus when it comes to the other metacognitive component that refers to metacognitive activities. Some authors have conceptualized this component as metacognitive regulation that involves setting appropriate reading goals, choosing a strategy to attain the set goals, monitoring to see if the goal is being met, and taking remedial action if it is not (Baker & Brown, 1984). Others differentiate metacognitive monitoring that refers to comprehension checking from metacognitive control that refers to strategy use in order to improve comprehension (Nelson, 1996; Winne, 1996). Using metacognitive strategies that eventually could become skills is critical in the regulation of reading behavior (Meijer, Veenman, & van Hout-Wolters, 2006; Pintrich & De Groot, 1990).

Knowledge and the use of different (meta)cognitive strategies helps students to effectively learn from texts (Flavell, Miller, & Miller, 1993). Metacognitive knowledge of strategies contributes to an awareness of the ways to attain a learning goal, and the conscious use of these strategies while reading, helps to (a) identify the relevant information in a text; (b) retrieve the relevant background knowledge from long-term memory, and (c) monitor and direct the use of these strategies, so as to develop a situational model of text that supports comprehension.

Reading strategies include a broad variety of specific behaviors, which can be classified according to their goals (e.g., to activate or use prior knowledge, to infer information not explicitly stated in text, to identify the main idea of a text, to process a text additionally after reading it; see Pressley, 1995), and according to the phases of reading (strategies used before reading, during reading and after reading; see Paris, Wasik, & Turner, 1991).

Differences in metacognitive knowledge about reading and in strategy use have been consistently found between good and poor readers. Pazzaglia, De Beni, and Caccio (1999), as well as Roeschl-Heils, Schneider, and van Kraayenoord (2003), found that poor comprehenders have less understanding of which reading strategies are appropriate in different reading situations, such as for studying or reading for pleasure. Anderson and Armbruster (1984) found that poor readers tend to skim, reread, integrate information, plan ahead and make inferences to a lesser extent than more skilled readers. Other researchers (Garner & Kraus, 1982; Grabe & Mann, 1984; Paris & Myers, 1981; Swanson & Alexander, 1997) have showed that poor readers experience difficulties in identifying the inconsistencies in a text.

1.2 Development of Metacognition in Reading

There is more than a 25-year long history of research dealing with the developmental aspects of metacognition in reading. The literature on metacognitive development in reading suggests that metacognitive knowledge about reading develops first (Paris et al., 1991). Children learn a lot about reading before they even begin their formal education through their exposure to different kinds of printed matter, and through reading with their parents. They develop an initial awareness about reading that is critical for the effectiveness of early instruction and reading attainment. Reading awareness continues to develop past the age of seven. Myers and Paris (1978) have examined the metacognitive knowledge of children between 8 and 12 years old and found that older children knew more about text structure, various reading goals and reading strategies than younger children. Pazzaglia et al. (1999) investigated the relationship between metacognitive knowledge and reading comprehension in a sample of children from 8 to 13 years. Metacognitive knowledge of reading goals and of strategies showed different developmental trends: notable improvements for the former were detected between the age of 8 and 9 years, and 11 to 12; for the latter, changes were continuous from the age of 8 onwards till 12 years. During secondary education, metacognitive knowledge of the nature of reading becomes more refined but even 12-year-olds neither have a well-defined knowledge about reading nor effective strategies that enhance reading comprehension.

Metacognitive activities in reading emerge in the period from 8 to 10 years of age and develop in the years following, at least with respect to more sophisticated and academically oriented skills (Veenman & Spaans, 2005). Certain metacognitive activities, such as monitoring and evaluation, appear to mature later on than others (e.g., planning). Baker and Brown (1984) have claimed that any attempt to comprehend must involve comprehension monitoring. However, younger children have difficulties in detecting semantic inconsistencies in texts in comparison to older children (Baker, 1984; Garner & Taylor, 1982). It seems that young children do not monitor meaning while reading since all of their working memory capacity is engaged in word recognition.

Despite reading improvement with age, even good 12-year-old readers do not detect a large number of errors and inconsistencies inserted into a meaningful text (Winograd & Johnston, 1980). Pazzaglia et al. (1999) compared developmental trends in metacognitive skills in students from primary school to university.

They found a continuous positive trend in monitoring and text sensitivity in children aged 8 to 13 and a prolonged development that spans through high school and university.

Paris et al. (1991) have summarized factors that contribute to the effects of age on comprehension monitoring. First, young children might not believe that there are mistakes in a text. Second, attentional capacity is engaged primarily in word recognition, and there are not enough remaining cognitive resources to construct text meaning and monitor comprehension. Third, many young readers do not understand the standards that can be used to evaluate comprehension. Fourth, reporting comprehension failure is substituted by making inferences in order to construct text meaning.

In summary, the development of metacognition in reading starts before the beginning of schooling, but more intensive developmental changes take place after school starts. Metacognitive knowledge about reading develops first, and later on the development of metacognitive activities becomes more pronounced, especially during upper elementary school and high school.

2 Gender Differences in Reading Comprehension

When considering developmental trends in the relations between metacognition and reading comprehension, the issue of gender differences should not be ignored because gender differences in text comprehension have been consistently found (Lietz, 2006). If gender differences in the development of metacognition exist, then this study could provide more detailed insight into the possible mechanisms that are involved in the development of reading difficulties and could motivate educational programs to be more gender-sensitive. Although there is a scarcity of studies addressing this topic, gender differences in reading have long been investigated. One of the first epidemiological studies on reading disability (Berger, Yule, & Rutter, 1975), which was conducted almost 35 years ago, showed that it was predominantly male students who had a reading disability. Other studies consistently show that boys are more prone to reading disabilities than girls (Flannery, Liederman, Daly, & Schultz, 2000). Rutter et al. (2004) conducted a detailed analysis of four independent epidemiological studies, and also concluded that reading disability was substantially more common in boys than in girls.

These results are in line with the results of the Programme for International Student Assessment (PISA survey; Organisation for Economic Cooperation and Development, 2007) which was conducted on 400,000 15-year-old students from 57 countries. In all the countries that participated in PISA 2006, on average, girls performed better in reading than boys. Likewise, in every country, boys were more likely than girls to be poor readers (Chiu & McBride-Chang, 2006). A meta-analysis of results obtained in large scale studies between 1975 and 2002 (Lietz, 2006) has also shown that girls, on average, score a 1.19 standard deviation higher than boys in text comprehension.

There are several possible explanations for the lower reading performance of boys across countries and different educational settings put forth by Alloway, Freebody, Gilbert, and Musprat (2002). The first possible reason that they identified concerns biological factors that might make boys more prone to reading disabilities. Another possible reason could be that boys lack male role-models as they learn to read because there are less male teachers and fathers rarely read with boys. The last explanation they gave is a sociological and educational one. In the majority of cultures, literacy practices themselves are gendered and considered to be feminine activities. Also, there is something in the materials relating to reading and writing lessons, which puts boys and their interests at a disadvantage, which is not the case for girls.

Although research results show that some of the differences do indeed stem from neuropsychological and other biological processes (Habib, 2000; Sauver, Katusic, Barbaresi, Colligan, & Jacobsen, 2001), Chiu and McBride-Chang (2006) found that reading enjoyment is the most important variable that differentiates between boys and girls, because it mediated 42% of the gender effect in the PISA study. The important roles of motivation and interest in reading comprehension, especially for boys, has also been found in other studies (Ainley, Hillman, & Hidi, 2002; Oakhill & Petrides, 2007).

The relationship between motivation, metacognition and reading has been examined previously (see Paris et al., 1991), but rarely in the context of gender differences. Although it is known that the use of reading strategies reflects both metacognitive knowledge about strategies and a willingness to use those strategies, the question remains as to which of these aspects poses a problem for boys.

One of the studies that has explored the relationship of both metacognition and motivation with reading comprehension while considering developmental trends and gender differences was a study by Roeschl-Heils et al. (2003). In one of their earlier studies (van Kraayenoord & Schneider, 1999), these researchers examined German children's metacognitive knowledge in relation to reading motivation and reading comprehension, in the 3rd and 4th grades. They found significant correlations between reading motivation, metacognitive variables and reading comprehension. Three years later Roeschl-Heils et al. (2003) obtained similar results. However, although most of the relations between the variables of the study were found to be stable over this period, significant correlations among the metacognitive variables were only found for those variables that were measured with the same type of measures of metacognition (e.g., think aloud or perceived use of reading strategies). In the 3rd and 4th grades there were no significant gender differences in any of the variables examined, and in the 7th and 8th grades gender showed only a significant effect on reading self-concept, with boys scoring higher than girls. On other reading and metacognition measures there was a tendency for girls to outperform boys, but none of these differences were statistically significant.

Although the gender differences in reading comprehension are well documented, gender differences in metacognition in reading were not often examined. It is of interest to explore whether the reading backwardness in boys is related to their metacognitive knowledge and metacognitive activities. The question also needs to be answered as to whether there is a difference in the pattern of the relationships between different aspects of metacognition and reading comprehension in girls and boys at different developmental levels. Our research attempted to answer these questions in order to contribute to the understanding of the relationship between reading comprehension and metacognition in different age groups.

3 Age and Gender Differences in Metacognition in Reading in Croatia

Our research expands on previous studies on the development of metacognition in reading. We conducted several cross-sectional studies in order to further broaden the findings about the relationship between metacognition and reading comprehension with regard to gender differences. The participants were Croatian elementary and high-school students (9–17 years). In Croatia, children attend compulsory elementary schooling from the age of 7 to the age of 14. All state schools follow a national curriculum. After 8 years of elementary schooling, they move to high school that lasts 3 or 4 years.

The overarching aim of our research was to explore age and gender differences in metacognitive knowledge and comprehension monitoring in students of elementary and high school in Croatia. Two aspects of metacognitive knowledge were explored: metacognitive knowledge about reading strategies and metacognitive awareness of reading strategy use. The relation between metacognitive knowledge, comprehension monitoring and reading comprehension in girls and boys at different developmental levels is also examined.

In what follows, the most important findings are presented**,** obtained in our six studies listed in Table 15.1. Some additional results, exceeding the findings presented in the published papers will be offered in order to draw attention to some gender and age differences that were not stated overtly in the papers.

In the first study, Kolić-Vehovec and Bajšanski (2003) explored the relations between metacognitive knowledge, comprehension monitoring tasks and reading comprehension on a sample of 3rd, 5th and 8th elementary school graders (9, 11, 14 years of age, respectively). As a measure of metacognitive knowledge of reading, a modified Croatian version of the Informed Strategies for Learning questionnaire (ISL; Paris, Cross, & Lipson, 1984) was used. Comprehension monitoring was assessed with sentence detection and cloze tasks. In the sentence detection task, students read a story consisting of six passages, each containing one semantically inappropriate sentence that they had to detect. In the cloze task, students were required to fill in the missing words in a text. Although the cloze task is a complex measure that taps different aspects of reading, in order to be successfully completed, it requires an ongoing monitoring of comprehension and careful checking for text consistency. A significant difference in metacognitive knowledge was obtained between the 8th graders compared to the 3rd and 5th graders, suggesting that an

Table 15.1 List of studies presented in the present chapter

Studies	Grades	Metacognitive knowledge	Metacognitive awareness	Comprehension monitoring
Kolić-Vehovec and Bajšanski (2003)	Elementary school: 3rd, 5th, 8th grade	x		x
Kolić-Vehovec et al. (2008, 2009), Pečjak et al. (2009)	Elementary school: 4th, 8th grade	x		
Kolić-Vehovec and Bajšanski (2001, 2006a)	Elementary school: 5th, 8th grade High school: 3rd grade		x	x
Kolić-Vehovec and Bajšanski (2006b)	Elementary school: 5th, 6th, 7th, 8th grade		x	x
Kolić-Vehovec and Bajšanski (2007)	Elementary school: 5th, 6th, 7th, 8th grade		x	x
Rončević Zubković (2008)	Elementary school: 7th grade		x	x

important improvement in metacognitive knowledge of reading occurs at upper elementary school after the age of 10.

Similar improvements in various aspects of metacognitive knowledge of reading during upper elementary school and high school have been reported by Pazzaglia et al. (1999). However, in this study we used different versions of the comprehension monitoring tasks and different texts in the text comprehension task for the three age groups due to the large difference in the reading skills between 3rd and 8th graders. Thus, it was not possible to examine age differences in comprehension monitoring. What is important to note, however, is that in all three age groups, no significant gender differences in metacognitive knowledge were found. However, gender differences in both monitoring tasks were marginally significant in fifth-graders ($p=0.06$) and significant in eighth-graders, and significant gender differences in text comprehension were also found between boys and girls in the 8th grade. There were low to moderate positive correlations (from $r=0.22$, $p<0.05$, to $r=0.53$, $p<0.01$) between metacognitive knowledge and performance on the comprehension monitoring tasks in all three age groups. A metacognitive knowledge of reading and performance on comprehension monitoring tasks also moderately positively correlated with reading comprehension in all age groups (from $r=0.32$ to $r=0.56$; $p<0.01$), and explained 17% of text comprehension in the 3rd and more than 25% in the 5th and 8th grades. Furthermore, hierarchical regression analysis revealed that the effect of metacognitive knowledge on reading comprehension was mediated by comprehension monitoring tasks. This finding indicates that the active use of monitoring strategies is more important for reading comprehension than merely a passive knowledge about these strategies.

In a recent study (Kolić-Vehovec, Pečjak, Ajdišek, & Rončević, 2008; Kolić-Vehovec, Pečjak, & Rončević Zubković, 2009; Pečjak, Kolić-Vehovec, Rončević Zubković, & Ajdišek, 2009), we wanted to examine whether similar findings would be obtained in Slovenia, a neighboring country, with a similar language and schooling system. The age and gender differences in metacognitive knowledge of reading strategies, as well as reading motivation, were examined in samples of 4th- and 8th-grade students from Croatia and Slovenia. In both samples, 8th graders showed higher metacognitive knowledge than 4th graders, and, in the main, girls demonstrated a higher metacognitive knowledge than the boys. However, the boys in the Slovenian sample had the same level of metacognitive knowledge as girls in the 8th grade, while boys in the Croatian sample still fell significantly behind girls. This difference can be attributed to a lack of motivation for reading exhibited by Croatian boys. Moreover, metacognitive knowledge of reading strategy use was consistently related to reading comprehension scores in both age groups ($r=0.25$ for the 5th and $r=0.40$ for the 8th graders), and in both samples, except for the 4th-grade Croatian boys, probably because of their significant deficit in metacognitive knowledge compared to all other students. The differences obtained between the Croatian and Slovenian samples could stem from more than a decade of different programs that aim to improve reading achievement in Slovenia.

To further explore reading strategies employment, we have examined students' awareness of the frequency of use of different reading strategies in several studies.

A high metacognitive knowledge of reading (including knowledge about reading goals, text structure and reading strategies) does not imply that readers will necessarily apply this knowledge in various reading situations (Baker, 2005). This is particularly the case for the use of reading strategies. In order to measure the awareness of reading strategy use, including strategies used before, during and after reading, Kolić-Vehovec and Bajšanski (2001) constructed the Strategic Reading Questionnaire (SRQ). The participants were asked to rate how often they use different reading strategies when reading a story. Factor analysis revealed a three-factor structure. Specifically, the first factor was Active Comprehension Strategies, which contained items related to the active construction of text meaning during reading (e.g., "When I read, sometimes I stop and think about what is important in the text."); the second factor was Comprehension Monitoring, which included items related to the monitoring of comprehension, detection of comprehension failure and the regulation of reading (e.g., "After reading, I try to assess if I understand what I have read."); and the third factor was Inference Generation, which included items related to active imagination and the anticipation of the content of a story (e.g., "During reading, I try to figure out what will happen next").

In the study by Kolić-Vehovec and Bajšanski (2006a), metacognitive awareness of reading strategy use was assessed by SRQ, in addition to comprehension monitoring tasks. With the aim of examining age differences in comprehension monitoring, we constructed uniform versions of cloze and sentence detection tasks. The students' age range was extended to high-school students, so the study was conducted in 5th and 8th elementary and 3rd high-school grades (11, 14, 17 years of age, respectively). Significant differences were obtained between all three age groups in both monitoring tasks: older students performed better than the younger students. These results are in line with previous findings showing that the development of reading-related metacognition continues after elementary education, and it is manifested in better comprehension monitoring (Baker, 2005). Aside from age differences, significant gender differences were obtained. The 5th- and 8th-grade girls outperformed the boys in the comprehension monitoring tasks, but no gender differences were obtained in the 3rd grade high-school students. Nevertheless, this lack of difference could have been partly due to ceiling effects. An alternative explanation could be that during high school boys improve their comprehension monitoring more intensively than girls, which results in the similar monitoring performance of female and male students at the end of high school.

An interesting pattern of age differences was obtained for all three SRQ subscales: students in the 5th grade rated that they use reading strategies more often than older students. No gender differences were found in the 5th grade on all three SRQ scales, but in the two older age groups girls scored higher than boys on SRQ scales.

We also examined whether SRQ was related to comprehension monitoring and reading comprehension. In the 5th grade, metacognitive awareness of reading strategy use was not related to monitoring tasks and reading comprehension, in the 8th grade the correlations were weak, and in the 3rd high-school grade, all three subscales of SRQ were related to cloze task and reading comprehension and the Inference subscale was related to performance on the sentence detection task.

These results suggest that an awareness of reading strategy use improves during upper elementary and high school and that comprehension monitoring tasks were stronger predictors of reading comprehension than an awareness of the reading strategy use.

In a further study (Kolić-Vehovec & Bajšanski, 2006b), we specifically explored age differences in comprehension monitoring from the 5th to 8th grades of elementary school, since our previous study (Kolić-Vehovec & Bajšanski, 2006a) showed that significant changes in comprehension monitoring occur during upper elementary education. We also extended our research to bilingual students (Kolić-Vehovec & Bajšanski, 2007) because it is claimed that bilingualism could positively contribute to metalinguistic awareness during elementary school (Francis, 1999). The participants were bilingual 5th to 8th grade upper elementary students from four Italian schools in Rijeka, Croatia. All the tasks were the same as for the monolingual Croatian sample, but in the Italian language.

In both studies (Kolić-Vehovec & Bajšanski, 2006b, 2007), in addition to the SRQ and the cloze task, 10 monitoring items from the Metacomprehension test (Pazzaglia et al., 1994) were used, in order to examine other aspects of comprehension monitoring, such as the detection of different types of errors in the text, and monitoring of text level comprehension. According to the results of a factor analysis, two subscales were formed. Specifically, (a) The Error Correction subscale, containing items related to the detection and correction of syntactic, spelling and semantic errors, and (b) The Text Sensitivity subscale, containing items related to text level comprehension monitoring. In the bilingual sample, the factor analysis yielded only one factor, so the scale in that group was treated as single.

Monitoring measures, that is, the Metacomprehension test, as well as the cloze task had positive significant correlations with reading comprehension in all four age groups in both samples. The correlations ranged from $r=0.26$ to $r=0.50$ (in all cases, $p<0.01$) in the Croatian sample and from $r=0.30$ to $r=0.64$ (in all cases, $p<0.01$) in the bilingual sample.

The relationship between monitoring measures and awareness of reading strategy use (SRQ) was also analyzed. In the Croatian sample no significant correlation in the 5th and 6th grades was obtained between the SRQ and monitoring measures. In the 7th grade only Error Correction was related to the SRQ. Finally, in the 8th grade, all monitoring measures were positively related to the SRQ. The correlations ranged from $r=0.22$ ($p<0.05$) to $r=0.34$ ($p<0.01$). In the bilingual sample, the SRQ subscales were not correlated with monitoring measures, except for a positive correlation between both monitoring measures and the Inference Generation subscale in the 8th grade (for cloze task, $r=0.34$, $p<0.05$; for Metacomprehension test, $r=0.36$, $p<0.01$).

In text comprehension, there were significant age and gender differences in both samples. The older students scored higher than the younger students, and the girls scored higher than the boys. However, in the Croatian sample differences in text comprehension were obtained between the 5th and the 7th grade, and the 6th and the 8th grade in girls, but only between the 5th and the 8th grade in boys, while in the bilingual sample a difference was obtained between the 8th graders and all others.

Gender differences were found in all but the 6th grade in the Croatian sample; in the bilingual sample, although the main effect of gender was significant, significant gender differences were not found in any of the grades.

Significant age differences were also found in the comprehension monitoring tasks. Differences in the performance on the cloze task in the Croatian sample were obtained between the 5th and the 6th graders, and between the 6th and the 7th graders. For error correction, significant differences were obtained between the 5th and the 8th graders, and for text sensitivity there were differences between the 5th and the 7th graders. These findings suggest that various aspects of comprehension monitoring develop at a different pace during upper elementary school. Gender differences were found in all three measures of comprehension monitoring: girls showed higher scores than boys on all measures, and there was no significant interaction of gender with age on any measure.

Although it seems that the development of metacognitive abilities and reading comprehension unfolds in a somewhat different manner in bilingual than in monolingual students, the most important findings obtained with the monolingual students were also replicated in the bilingual sample. There was a positive effect of age on comprehension monitoring measures (cloze task, Metacomprehension test), and there were significant gender differences, with the girls scoring higher than the boys. In both samples, the 5th-grade girls scored higher on the monitoring tasks than the boys, but this difference disappeared in the 6th grade, because the boys in the 6th grade scored higher than those in the 5th grade. On the other hand, in the Croatian sample gender differences were also obtained in the 7th and the 8th grades, whereas in the bilingual sample gender differences were obtained only for metacomprehension tasks in the 8th grade.

The pattern of age and gender differences in SQR in both samples was similar to the Kolić-Vehovec and Bajšanski (2006a) study. In the Croatian sample, students in the 5th grade rated that they use active comprehension and comprehension monitoring strategies more often than students in the 8th grade, and there were no age differences in perceived inference generation. In the bilingual sample, the effect of the grade level on the awareness of reading strategy use was not found.

There are several possible explanations for the obtained pattern of age differences in SRQ. First, it is possible that younger students overestimate the frequency of strategy use as a consequence of their inadequate self-assessment. Second, younger students might use strategies more often but in an inefficient way. Third, in older students some aspects of strategic reading become automatic and are no longer under conscious control. Thus, lower ratings reflect a lack of awareness due to automatized processing rather than a lack of the use of reading strategies. Fourth, this pattern of differences could be in part due to motivational factors, including social desirability and the perceived value of studying, that is emphasized in younger ages.

The analysis of gender differences in the perceived use of reading strategies offers another possible explanation for the obtained pattern of age differences. For all three SRQ subscales, a significant effect of gender was obtained in both Kolić-Vehovec and Bajšanski (2006a and 2006b) studies on Croatian samples; specifically, girls reported that they use strategies more often than boys. However, post hoc analyses of the results of the first study (Kolić-Vehovec & Bajšanski, 2006a) revealed that in the 5th grade there were no significant differences in the perceived use of reading strategies between girls and boys. This was in clear contrast to the older groups of students (the 8th elementary school and the 3rd high-school grades). Furthermore, girls in all three age groups had similar results, whereas older boys produced lower ratings than younger ones. Similar results were obtained in the second study (Kolić-Vehovec & Bajšanski, 2006b): there were no gender differences in the 5th, the 6th, and the 7th grade, but the boys in the 8th grade scored lower than the girls in all grades. In the bilingual sample (Kolić-Vehovec & Bajšanski, 2007), girls in the 5th and the 6th grade scored higher than the boys in the same grades in perceived reading strategy use, but there were no differences between grades. These results could reflect a lack of reading motivation in boys at the end of elementary school.

We gained further insights into the relation between gender and metacognition from the study by Rončević Zubković (2008) that aimed to differentiate between the profiles of the 7th grade students based on reading-relevant variables (sentence detection task, SRQ, verbal working memory). It was found that the group of the best comprehenders, who were predominantly girls, had the highest scores on both the SRQ and the sentence detection task, while the group of the poorest comprehenders, who were predominantly boys, had the lowest score on the sentence detection task and a below average result on the SRQ. However, two distinct profiles of students emerged amongst the moderate comprehenders. Both groups of students performed averagely on the sentence detection task, but students in the first group, predominantly boys, had very low SRQ scores whereas in the second group they had very high SRQ scores with no difference from the best comprehenders. There was no significant difference in working memory capacity between the best comprehenders (predominantly girls), and the moderate comprehenders who rarely used reading strategies (predominately boys). Therefore, differences in text comprehension, self-reported frequency of reading strategy use and sentence detection efficiency between the two groups could stem from motivational factors.

4 Discussion and Conclusions

The results of all our presented studies support the proposition that metacognition plays an important role in text comprehension, but different components of metacognition do not contribute equally to comprehension. Metacognitive monitoring had a stronger effect on text comprehension than metacognitive knowledge about reading. Furthermore, the effects of these metacognitive components on text comprehension depend on the age and gender of students. Obtained age effects on metacognition are consistent with the literature review on the development of metacognition (Alexander, Graham, & Harris, 1998; Paris & DeBruin-Parecki, 1999; Schneider & Sodian, 1997; Veenman, Wilhelm, & Beishuizen, 2004). Overall, the importance of metacognition for text comprehension intensifies with age with a noticeable increase in this influence in the transition between lower to upper elementary school (after about 10 years of age). Children's metacognition improves strikingly during upper elementary school, but this improvement continues during high school (Kolić-Vehovec & Bajšanski, 2006a). The most important findings showing a positive effect of age on comprehension monitoring measures were also replicated in the bilingual sample (Kolić-Vehovec & Bajšanski, 2007). However, metacognitive and text comprehension development unfolds at a somewhat slower pace in bilingual students than in monolingual students. It seems that children should reach a threshold in language proficiency to effectively monitor their comprehension and improve their reading.

The effect of gender on metacognition in reading is mainly evident from the better performance of girls on the monitoring tasks, while gender differences in metacognitive knowledge of reading were not so consistent, although girls showed better metacognitive knowledge of reading strategies. Additionally, boys showed slower metacognitive development than girls, and this seemed primarily related to their lack of reading motivation. This interpretation mainly stems from results showing that gender differences in metacognition exist in Croatian students during upper elementary school, whereas in Slovenian students gender differences existed in the 4th grade but not in the 8th grade. This finding indicates that the systematic endeavor to support reading improvements within regular school practice that was carried out in Slovenia might have remedied motivational deficits in boys.

A more detailed investigation of changes in the different components of metacognition during upper elementary school showed interesting findings, suggesting that various aspects of comprehension monitoring develop at a different pace (Kolić-Vehovec & Bajšanski, 2006a). The monitoring of local comprehension develops first, then the monitoring of global text comprehension, and finally the monitoring of grammar and spelling accuracy. Students' awareness of reading strategy use improved in accuracy from overestimating it in the 5th grade to a more accurate self-assessment at the end of elementary school and in high school. A higher accuracy of self-assessment of reading strategy use is related to better text comprehension. This could mean that older students are also more aware of the appropriateness of specific strategies regarding text features and specific reading aims. Over time, students gain experience in strategy use, and high school lessons provide students with opportunities to practice strategy use that could foster the transfer of metastrategic knowledge to real learning contexts. This is in line with Kuhn's (2000) claim that cognitive development entails a shifting distribution in the frequencies with which more adequate strategies are applied. Comprehension monitoring leads to an enhanced metacognitive awareness of the goal and the extent to which it is being met by different strategies, as well as enhanced awareness and understanding of the strategies themselves. The selection of a suitable strategy and skillfully applying the strategy results in better text comprehension.

In all the presented studies that have assessed metacognitive awareness of reading strategy use (Kolić-Vehovec & Bajšanski, 2006a, 2006b, 2007) girls reported a more frequent use of reading strategies than boys. Gender differences could emerge from educational practices that might not be equally beneficial for boys and girls (Alloway et al., 2002). As Baker (2005) has noted, students in upper elementary education are cognitively prepared for metacognitive improvement, but at the same time their intrinsic motivation for learning decreases. It is possible that boys are not motivated enough to use reading strategies, while girls persist with reading (and probably with the use of reading strategies, too) and show good reading performance even if the text is not interesting (Ainley et al., 2002; Oakhill & Petrides, 2007). In line with this proposition is also the finding of another study of ours (Kolić-Vehovec et al., 2009) showing that girls had a higher reading motivation than boys in the final grade of elementary school. The importance of motivation in reading strategy use for reading comprehension is also evident in the results of Rončević Zubković (2008). Students who reported that they frequently use reading strategies overcame working memory limitations and attained a similar level of comprehension as those students, who were predominantly boys, with larger working memory spans but a low frequency of reading strategy use. The finding that gender differences ceased to be significant in the 3rd grade of high school (17 years old) could be the result of the increased motivation of boys prompted by an awareness of the necessity to employ reading strategies in order to efficiently master the demands of the high school of their choice.

Developmental theories of self-regulated learning postulate that learners have to pass through different levels of self-regulation until they achieve a level of self-regulating their learning in an adaptive way and in changing conditions (Zimmerman, 2002). Following Zimmerman's (2002) model of the development of self-regulated learning, it is only at the higher developmental levels that the learners can control and regulate their own learning process independently of others. In these stages, metacognitive reflection about when and how to use which strategy can take place in an independent way. On the other hand, young children, whose metacognitive and metastrategic knowledge is still developing (Kuhn, 2000), might benefit more from pure instruction of metacognitive strategies by modeling and imitating, so that they are still dependent on external feedback.

The results of intervention studies by Kolić-Vehovec (2002a, 2002b) point to the importance of training children's self-monitoring for improving accuracy of reading as early as the second grade. The students who were encouraged to engage in self-correction during reading significantly improved their reading accuracy. The results also indicated that reading practice alone was not sufficient enough to make changes in reading accuracy. These results are in line with Malicky, Juliebo, Norman, and Pool (1997), who claim that reinforcing self-correction is necessary for the development of metacognitive awareness and control, which are considered to be essential functions of a self-improving system. It is also an indication that comprehension monitoring could be gradually built, first at the level of sentence comprehension, and later on having experienced the benefits of the comprehension monitoring strategy, students could be encouraged to monitor the meaning of larger units of text. At the beginning of their schooling, children were able to engage in metacognitive behavior when tasks were at an appropriate level of difficulty. Baker (2005) also argues that it is possible, and beneficial, to teach comprehension monitoring strategies to young children, even if they are not completely fluent readers, but she also points out that the types of metacognitive demands placed on children should increase with development. Some studies show that efficient comprehension monitoring and the employment of compensatory strategies like rereading and looking back could help overcome inefficient word recognition skills and working memory constraints even at an older age (Walczyk, Wei, Grifith-Ross, Goubert, Cooper, & Zha, 2007). According to Alexander et al. (1998), children's strategic behavior changes as they become more experienced and competent, and they might therefore benefit more from the instruction in metacognitive strategies in order to broaden their strategy repertoire, as is supported by our results on reciprocal teaching training (Kolić-Vehovec & Muranović, 2004).

In conclusion, metacognition is important in different phases of reading development, but especially during upper elementary and high school. Comprehension monitoring is especially important for the development of reading comprehension. An effort should be made in fostering comprehension monitoring already at the word identification level, but special attention should be paid to the stimulation of comprehension monitoring after the age of 10, and should be extended throughout high school with an emphasis on motivational support as well. For many students, metacognitive skills are unlikely to develop automatically just by virtue of reading texts in classroom activities. Instead, metacognitive skills should be explicitly taught within the context of authentic literacy environments, and students should be given sufficient practice in applying them. It is especially important for boys that instruction is more adapted to their interests and that students are provided with the opportunity to practice strategy use in stimulating contexts.

References

Ainley, M., Hillman, K., & Hidi, S. (2002). Gender and interest processes in response to literary texts: situational and individual interest. *Learning and Instruction 12*, 411–428.

Alexander P. A., Graham, S., & Harris K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review, 10*, 129–154.

Alloway, N., Freebody, P., Gilbert, P., & Muspratt, P. (2002). *Boys, literacy and schooling: Expanding the repertoires of practice.* Canberra, Australia: Commonwealth Department of Education, Science and Training.

Anderson, T. H., & Ambruster, B. B. (1984). Studying. In P. D. Pearson, M. Kamil, R. Barr, & P. Rosenthal (Eds.), *Handbook of reading research* (Vol. 1., pp. 353–395). White Plains, NY: Longman.

Baker, L. (1984). Spontaneous versus instructed use of multiple standards for evaluating comprehension: Effects of age, reading proficiency, and type of standard. *Journal of Experimental Child Psychology, 38*, 289–311.

Baker, L. (2005). Developmental differences in metacognition: Implication for metacognitively oriented reading instruction. In S. E. Israel, C. C. Block, K. L. Bauserman, & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction, and professional development* (pp. 61–79). Mahwah, NJ: Erlbaum.

Baker, L., & Brown, A. (1984). Metacognitive skills and reading. In P. D. Pearson, M. Kamil, R. Barr, & P. Mosenthal (Eds.), *Handbook of reading research* (Vol. 1., pp. 657–679). White Plains, NY: Longman.

Berger, M., Yule, W., & Rutter, M. (1975). Attainment and adjustment in two geographical areas: II. The prevalence of specific reading retardation. *British Journal of Psychiatry, 126*, 510–519.

Chiu, M. M., & McBride-Chang, C. (2006). Gender, context, and reading: A comparison of students in 43 countries. *Scientific Studies of Reading, 10*, 331–362.

Flannery, K. A., Liederman, J., Daly, L., & Schultz, J. (2000). Male prevalence for reading disability is found in a large sample of Black and White children free from ascertainment bias. *Journal of the International Neuropsychological Society, 6*, 433–442.

Flavell, J. H., Miller, P. H., & Miller, S. A. (1993). *Cognitive development.* Englewood Cliffs, NJ: Prentice Hall.

Francis, N. (1999). Bilingualism, writing, and metalinguistic awareness: Oral-literate interactions between first and second languages. *Applied Psycholinguistics, 20*, 533–561.

Garner, R., & Taylor, N. (1982). Monitoring and understanding: An ivestigation of attentional assistance needs at different grade and proficiency levels. *Reading Psychology, 3*, 1–6.

Garner, R., & Kraus, C. (1982). Good and poor comprehenders' differences in knowing and regulating reading behavious. *Educational Reading Quarterly, 6*, 5–12.

Grabe, M., & Mann, S. (1984). A technique for the assessment and training of comprehension monitoring skills. *Journal of Reading Behavior, 16*, 131–144.

Habib, M. (2000). The neurological basis of developmental dyslexia: An overview and working hypothesis. *Brain, 123*, 2373–2399.

Kolić-Vehovec, S. (2002a). Self-monitoring and attribution training with poor readers. *Studia Psychologica, 44*, 57–68.

Kolić-Vehovec, S. (2002b). Effects of self-monitoring training on reading accuracy and fluency of poor readers. *European Journal of Psychology of Education, 17*, 129–138.

Kolić-Vehovec, S., & Bajšanski, I. (2001). Konstrukcija upitnika strategijskog čitanja [The construction of Strategic Reading Questionnaire]. *Psihologijske Teme, 10*, 51–62.

Kolić-Vehovec, S., & Bajšanski, I. (2003). Children's metacognition as a predictor of reading comprehension. In G. Shiel & U. Ni Dhalaigh (Eds.), *Other ways of seeing: Diversity in language and literacy* (pp. 216–222). Dublin: Reading Association of Ireland.

Kolić-Vehovec, S., & Bajšanski, I. (2006a). Dobne i spolne razlike u nekim vidovima metakognicije i razumijevanja pri čitanju [Age and gender differences in some aspects of metacognition and reading comprehension]. *Društvena Istraživanja, 15*, 1005–1027.

Kolić-Vehovec, S., & Bajšanski, I. (2006b). Metacognitive strategies and reading comprehension in elementary school students. *European Journal of Psychology of Education, 21*, 439–451.

Kolić-Vehovec, S., & Bajšanski, I. (2007). Comprehension monitoring and reading comprehension in bilingual students. *Journal of Research in Reading, 30*, 198–211.

Kolić-Vehovec, S., & Muranović, E. (2004). Evaluacija treninga recipročnog poučavanja strategija čitanja [Evaluation of reciprocal teaching training of reading strategies]. *Suvremena Psihologija, 7*, 95–108.

Kolić-Vehovec, S., Pečjak, S., Ajdišek, N. & Rončević, B. (2008). Razlike med spoloma v (meta) kognitivnih in motivacijsko emocionalnih dejavnikih bralnega razumevanja [Gender differences in (meta)cognitive, emotional and motivational factors of reading comprehension]. *Psihološka Obzorja, 17*, 89–116.

Kolić-Vehovec, S., Pečjak, S., & Rončević Zubković, B. (2009). Spolne razlike u (meta)kognitivnim i motivacijskim čimbenicima razumijevanja teksta adolescenata u Hrvatskoj i Sloveniji [Gender differences in (meta)cognitive and motivational factors of text comprehension in adolescent in Croatia and Slovenia]. *Suvremena Psihologija, 12,* 229–242.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science, 9*, 178–181.

Lietz, P. (2006). A meta-analysis of gender differences in reading achievement at the secondary school level. *Studies in Educational Evaluation, 32*, 317–344.

Malicky, G. V., Juliebo, M. F., Norman, C. A., & Pool, J. (1997). Scaffolding of metacognition in intervention lessons. *Alberta Journal of Educational Research, 43*, 114–126.

Meijer, J., Veenman, M. V. J., & van Hout-Wolters, B. H. A. M. (2006). Metacognitive activities in text-studying and problem-solving: Development of a taxonomy. *Educational Research and Evaluation, 12*, 209–237.

Myers, M., & Paris, S. G. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology, 70*, 680–690.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102–116.

Oakhill, J. V., & Petrides, A. (2007). Sex differences in the effects of interest on boys' and girls' reading comprehension. *British Journal of Psychology, 98*, 223–235.

Organisation for Economic Cooperation and Development (2007). *PISA 2006: Science competencies for tomorrow's world. Executive summary.* Retrieved July 11, 2008, from http://www.oecd.org/dataoecd/15/13/39725224.pdf

Paris, S. G, Cross, D. R., & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76*, 1239–1252.

Paris, S. G., & DeBruin-Parecki, A. D. (1999). Metacognitive aspects of literacy. In D. Wagner, R. Venezky, & B. Street (Eds.), *Literacy: An international handbook* (pp. 74–80). Boulder, CO: Westview.

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*, 293–316.

Paris, S. G., & Myers, M. (1981). Comprehension monitoring in good and poor readers. *Journal of Reading Behavior, 13*, 5–22.

Paris, S. G., Wasik, B. A., & Turner, J. C. (1991). The development of strategic readers. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 609–639). Mahwah, NJ: Erlbaum.

Pazzaglia, F., De Beni, R., & Caccio, L. (1999). The role of working memory and metacognition in reading comprehension difficulties. In T. E. Scruggs, & M. A. Mastropieri (Eds.), *Advances in learning and behavioral disabilities* (Vol. 13, pp. 115–134). Greenwich, CT: JAI.

Pazzaglia, F., De Beni R., & Cristante, F. (1994). *Prova di metacompresione.* [Metacomprehension test]. Firenze, Italia: Organizzazioni Speciali.

Pečjak, S., Kolić-Vehovec, S., Rončević Zubković, B., & Ajdišek, N. (2009). (Meta)kognitivni i motivacijski prediktori razumijevanja teksta adolescenata u Hrvatskoj i Sloveniji [(Meta)cognitive and motivational predictors of text comprehension of adolescents in Croatia and Slovenia]. *Suvremena Psihologija, 12,* 257–270.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33–40.

Pressley, M. (1995). More about the development of self-regulation: Complex, long term, and thoroughly social. *Educational Psychologist, 30*, 207–212.

Roeschl-Heils, A., Schneider, W., & van Kraayenoord, C. E. (2003). Reading, metacognition and motivation: A follow-up study of German students in grades 7 and 8. *European Journal of Psychology of Education, 18*, 75–86.

Rončević Zubković, B. (2008). *Uloga radnog pamćenja i strategijskog procesiranja u razumijevanju pri čitanju kod djece* [The role of working memory and strategic processing in children's reading comprehension]. Unpublished PhD thesis. University of Zagreb, Croatia.

Rutter, M., Caspi, A., Fergusson, D. M., Horwood, L. J., Goodman, R., Maughan, B., Moffitt, T. E., Meltzer, H., & Carroll, J. (2004). Gender differences in reading difficulties: Findings from four epidemiology studies. *Journal of the American Medical Association, 291*, 2007–2012.

Sauver, J. L., Katusic, S. K., Barbaresi, W. J., Colligan, R. C., & Jacobsen, S. J. (2001). Boy/girl differences in risk for reading disability: Potential clues? *American Journal of Epidemiology, 154*, 787–794.

Schneider, W., & Sodian, B. (1997). Memory strategy development: Lessons from longitudinal research. *Developmental Review, 17*, 442–461.

Swanson, H. L., & Alexander, J. (1997). Cognitive processes as predictors of word recognition and reading comprehension in learning disabled and skilled readers: Revisiting the specificity hypothesis. *Journal of Educational Psychology, 89*, 128–158.

van Kraayenoord, C. E., & Schneider, W. (1999). Reading achievement, metacognition, reading self-concept and interest: A study of German students in grades 3 and 4. *European Journal of Psychology of Education, 14*, 305–324.

Veenman, M. V. J., & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences, 15*, 159–176.

Veenman, M. V. J., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction, 14*, 89–109.

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning, 1*, 3–14.

Walczyk, J. J. (1994). The development of verbal efficiency, metacognitive strategies, and their interplay. *Educational Psychology Review, 6*, 173–189.

Walczyk, J. J., Wei, M., Grifith-Ross, D. A., Goubert, S. E., Cooper, A. L., & Zha, P. (2007). Development of the interplay between automatic processes and cognitive resources in reading. *Journal of Educational Psychology, 99*, 867–887.

Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences, 8*, 327–353.

Winograd, P., & Johnston, P. (1980). Comprehension monitoring and the error detection paradigm. *Journal of Reading Behavior, 14*, 61–76.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*, 64–72.

Chapter 16
Metacognition-Based Reading Intervention Programs Among Fourth-Grade Hungarian Students

Csaba Csíkos and János Steklács

Children, do you like reading?
Hmm, eeerrrr, not so much.
What do you like to do in your free time?
Computer games, chatting, emailing friends…
And, while you are doing these, aren't you reading different texts on the screen?
Oh, yes, we are.
Then, let me ask again, do you like reading?
Why, we do.

[Segment from a classroom conversation]

1 Introduction

The term metacognition (Flavell, 1971) has been seminal in inspiring research on strategic processes in fields related to students' academic achievement. Research on mathematics and reading skills acquisition are two fields that are strongly influenced by theories stressing the importance of metacognitive processes (see, e.g., Campione, Brown, & Connel, 1988; Hacker, Dunlosky, & Graesser, 1998; Pressley, 2000). In what follows, first, some theoretical models of reading comprehension are outlined, and then two intervention programs that took place in Hungary and provided evidence about the role of metacognitive strategies in reading comprehension will be presented.

C. Csíkos (✉)
Institute of Education, University of Szeged, Petőfi S. sgt. 30-34, 6722, Szeged, Hungary
e-mail: csikoscs@edpsy.u-szeged.hu

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*, DOI 10.1007/978-1-4419-6546-2_16,

1.1 Models of Reading Processes and Metacognition

Carroll's (1993) factor-analytic studies of intelligence have revealed first-order factors related to reading, such as, reading comprehension and reading decoding. Other models of reading, however, advocate a two-level structure comprising of higher- and lower-level components. Stanovich (1980, p. 36), for example, claims that poor readers who have a deficit in lower-level reading processes "might actually rely more on higher-level contextual factors" for reading comprehension. This viewpoint posits an "interactive-compensatory" model and entails that reading processes of either level may compensate for deficiencies in the other level. Yet, this model does not explicitly refer to metacognitive processes.

Another compensatory-type model is Walczyk's (1995) compensatory-encoding model which deals with reading under time pressure. According to Walczyk (1995, p. 399), "reading involves the concurrent execution of several subcomponent processes arranged hierarchically". Compensatory mechanisms enable the reader to overcome deficiencies in subcomponents of the reading process. These compensatory mechanisms are metacognitive control strategies, such as slowing the reading rate or re-reading parts of the text.

Perfetti's (1985) verbal efficiency theory, on the other hand, emphasizes the automatization of reading subcomponents because automatization enables mature readers to allocate or reallocate attention to higher-level reading processes, such as schema activation and inference drawing. This implies that higher-level processes are not necessarily metacognitive in nature, although they might involve metacognition. Specifically, the differentiation between higher- and lower-level reading processes can be understood in terms of the two levels of the Nelson and Narens's (1992) model of metacognition (Nelson, 1996). Lower-level components are automatized and, often, nonconscious processes, whereas higher-level components monitor and control the lower level ones. Indeed, lower-level processes of reading can be regarded as object level processes that are being monitored and controlled by meta-level processes. In most cases, for mature readers, decoding a word does not require conscious monitoring and control processes; on the contrary, higher-level (meta-) processes are implicated in more complex reading tasks that require text comprehension.

The question whether meta-level processes are conscious or nonconscious is still a major theoretical issue. Whereas classical definitions considered metacognition as "potentially reportable" (see Hacker, 1998) cognition about cognition, some authors (Kentridge & Heywood, 2000; Koriat & Levy-Sadot, 2000) have posited that even nonconscious processes can be involved in metacognition. In the present chapter we restrict the use of the term "metacognition" to conscious and deliberate monitoring and control processes, and we use the term "metacognitive strategy" (see Kluwe, 1987) as procedural metacognitive knowledge.

From an educational point of view, it seems useful to adopt the distinction emphasized by Kluwe (1987) as regards the facets of metacognition. Declarative metacognition refers to declarative knowledge about one's own thinking or about thinking in general. Procedural metacognition refers to procedural knowledge used for planning, monitoring and controlling one's own thinking processes. Another useful categorization of metacognitive processes was provided by Efklides (2001, 2006). Following Flavell's (1979) distinction between metacognitive knowledge (MK) and metacognitive experiences (ME), and accepting Nelson's (1996) division between monitoring and control processes, she categorized MK and ME as monitoring processes, while metacognitive skills as control processes.

Following the above distinctions, one could say that metacognition in reading involves awareness of one's reading as it takes place (e.g., metacognitive experiences such as having made an error or not understanding the meaning of the text) and metacognitive knowledge including beliefs about reading. Procedural metacognition (or control processes) in reading take the form of reading strategies and metacognitive skills, that is, processes that involve planning, monitoring and controlling, as well as evaluating one's reading behavior. In contrast, Afflerbach, Pearson, and Paris (2008) proposed to reserve the term "reading strategy" for various types of control and monitoring processes, including awareness that "helps the reader select an intended path" (p. 368). Therefore, in reading research there is consensus that there are cognitive and metacognitive processes involved, but no consensus on the nature of the latter processes.

1.2 Metacognition in Reading Comprehension

There is a lot of empirical research on the role of metacognition in reading comprehension. We summarize the studies of van Kraayenoord and Schneider (1999), Cromley and Azevedo (2006), and Meneghetti, Carretti, and De Beni (2006), because they all distinguish cognitive from metacognitive processes.

Van Kraayenoord and Schneider's (1999) research was conducted among third and fourth graders. It aimed at determining the predictors of reading comprehension. Four main independent variables were selected as predictors: (a) grade (whether third or fourth), (b) motivation, (c) metacognition, and (d) decoding skills. The study showed that reading comprehension was predicted by decoding skills as well as by metacognition. Metacognition was measured with the Index of Reading Awareness questionnaire (see Jacobs & Paris, 1987), which taps metacognitive knowledge about reading. Motivation had only indirect effects on reading comprehension. The indirect effects were mediated by both metacognition and decoding skills. Similarly, pupils' grade (whether third or fourth) had a relatively low impact on reading comprehension, and this effect was mediated also by decoding skills and metacognition.

One of the most recent models of reading, the Direct and Inferential Mediation (DIME) model was proposed by Cromley and Azevedo (2006). The model was tested through hierarchical regression analyses in which reading comprehension was the dependent variable. The sample consisted of ninth grade students. There were five predictors: (a) background knowledge (related to the topic of the texts of the reading comprehension test), (b) strategies (measured with multiple-choice questions on the use of reading strategies), (c) word reading (measured with accuracy and fluency tests), (d) vocabulary and (e) inference drawing (measured with inferences drawing tests after reading at the sentence and paragraph level). The Strategies factor had an indirect effect on reading comprehension mediated by the Inference Drawing factor. This implies that metacognition in the form of awareness of strategy use may not be as important as the previous study (Van Kraayenoord & Schneider, 1999) had suggested. However, this finding may be due to the higher experience with reading the students of the study had compared to those of Van Kraayenoord and Schneider (1999).

Meneghetti et al.'s (2006) research aimed at grouping ten aspects of reading comprehension into sub-components. The sample of the study comprised students aged between 9 and 13 years. Confirmatory factor analysis revealed that a model containing two higher-order factors accounted for the data most appropriately. The two factors in this model can be interpreted as representing processes at the reading skill level and processes at the strategy level of reading, respectively. The latter are metacognitive control processes.

To sum up, all the three studies presented in this section provided empirical evidence about the presence of metacognitive components in reading. However, Cromley and Azevedo's (2006) study showed that the effect of reading strategies on reading comprehension is indirect. Therefore, it is not clear whether interventions aiming at the enhancement of reading comprehension through the use of reading strategies will be effective, particularly with young readers who have not developed yet their inference drawing abilities. Nevertheless, metacognition-based reading intervention programs in primary education may be effective, particularly in the case of poor readers when there is compensation of deficiencies in lower-level subcomponents of the reading process through higher-level metacognitive processes (Walczyk, 1995).

1.3 What Develops in Reading Comprehension Development?

The question raised here paraphrases Flavell's (1971) celebrated question about memory development; it also paraphrases Pressley's (2000) title of a seminal study on reading comprehension instruction. Having presented multi-component and two-level models of reading, we may split the question: (a) Do the different components involved in reading comprehension develop? (b) How can students' learning of higher-level processes, such as reading strategies, be promoted? In the studies reported here we aimed at enhancing the metacognitive components of reading, that is, we were seeking ways of improving higher-level (strategic) components of reading. Our assumption was that components from different levels of reading develop "in tandem", that is, from the beginning of learning to read, both lower-level and higher-level components work closely together (see Schellings, Aarnoutse, & van Leeuwe, 2006). Pressley (2000) provided an outline of different components of reading comprehension. He used the distinction between lower- and higher-level processes not only for describing components of reading comprehension, but also for formulating recommendations on how to promote these processes. Although he stated that higher-order processes (in his words: processes above the word level) cannot be applied before word-level processes are mastered, he reported successful teaching of comprehension strategies in students as early as Grade 2.

There are several approaches to instruction targeting reading comprehension strategies. In their reciprocal teaching approach, Palincsar and Brown (1985) suggested focusing on four main strategies: predicting, questioning, clarifying, and summarizing. They also emphasized the importance of the teacher modeling the strategy use. Brand-Gruwel, Aarnoutse, and Van den Bos's (1998) classroom-based intervention with fourth-grade struggling readers also focused on the four strategies, namely predicting, questioning, clarifying, and summarizing, and proved to be effective. However, according to Keene and Zimmermann (1997), the most relevant and important strategies for reading comprehension are activation of prior knowledge, prioritizing information, emphasis on most important ideas, questioning the author and the text, evoking sensory images, drawing inferences, retelling or synthesizing, and using fix-up strategies to repair comprehension. This implies that reading comprehension instruction may involve many more strategies than the four proposed by Palincsar and Brown (1985).

Pressley (2002), on the other hand, grouped reading strategies according to phases of information processing. Before reading decoding processes commence, strategies like clarifying goals, skimming the text for information about length and structure, and activating prior knowledge are important. In the next step, the process of meaning construction, where we can find the strategies of identifying main ideas, making predictions, monitoring of understanding, and drawing inferences are relevant. At the end of the reading process summarizing, drawing conclusions, and formulating self-questions for understanding take place. Almasi (2003) also described three clusters of reading strategies depending on phases of reading. Specifically, *text-anticipation strategies* include, among others, schema selection and (metacognitive) knowledge about the way texts are organized. *Text-maintenance strategies* help to focus attention while reading, and to monitor the incoming information that may or should match mental images. Finally, *fix-up strategies* involve processes like re-reading, slowing down, or consulting an outside source (e.g., a dictionary).

Reading strategies identified and described in the past decades constitute the basis for the two classroom interventions presented in this chapter. Strategy clusters defined by Pressley (2000, 2002) and Almasi (2003) proved to be insightful regarding classroom instruction, although reading intervention research has also provided data about the factors affecting the effectiveness of reading strategies instruction.

1.4 Reading Strategy Intervention

Research on instruction for improving reading strategies has resulted in several principles and a large body of empirical data. The necessity and possibility of early interventions (even for struggling readers) was emphasized by Gaskins (1994).

She argues that early strategy instruction may result in automatic strategy use later on which is an instructional goal in reading. The need for improving primary school children's reading strategies was also underscored by various studies conducted by Cross and Paris (1988), Garner (1987), Hall, Bowman, and Myers (1999), and Myers and Paris (1978).

One of the guiding principles in interventions for the promotion of reading comprehension is the explicit instruction of reading strategies. According to Van den Bos, Brand-Gruwel, and Aarnoutse (1998), in reading comprehension the "what" and "how" questions are necessarily connected to explicit reading instruction. Baker and Brown (1984) also pointed out the importance of students' knowing when and where to use reading comprehension strategies. Explicit strategy teaching entails that the teachers model and explain the strategies at hand and think aloud when demonstrating the use of a strategy (Almasi, 2003), so that students become aware of their own reading processes (see also Palincsar & Brown, 1985).

Meloth and Deering (1992) employed cooperative learning techniques in their classroom-based experiment, in which in-service teachers were trained in two different ways. All of them learnt about four clusters of strategies as represented in the four subscales of Jacobs and Paris's (1987) Reading Awareness Questionnaire, namely Evaluation, Planning, Regulation and Conditional Knowledge. One group of the teachers (i.e., strategy group) studied how to provide information to students about the cognitive and metacognitive aspects of reading comprehension, and how to plan lessons and design activities to support student discussion about the four target strategies. The other group learnt how to assist students to participate in group discussions, and to help them focus on lesson content. The results of the study showed the effectiveness of the "strategy" group, since their students performed better on a reading comprehension test. This study showed the importance of working not only with students, but also with teachers as mediators of students' learning the use of reading strategies.

1.4.1 Reading Intervention in Context: The Teaching of Reading in Hungary

International comparison surveys like The Reading Literacy Study (Elley, 1994) and the Programme for International Student Assessment (PISA) studies (Organisation for Economic Co-operation and Development [OECD], 2006) have made clear that changes are needed in the Hungarian reading instruction.

After the first IEA reading study (see Thorndike, 1973), much effort has been devoted to the teaching of reading decoding skills (Báthory, 1992). Various methods have been developed, often named after either their author or their main feature. The characteristics of the Hungarian language favor some methods of instruction, while others have been proven to be less successful. For example, Hungarian is an agglutinative language (i.e., words are usually formed by joining morphemes), consequently, there is a huge number of word forms, and words can be quite lengthy. Thus, when teaching decoding skills, it is almost impossible to use the so-called global methods rooted in Decroly's pedagogy (see Adamikné, 2001). In Hungary there is a hot debate over the methods of teaching decoding skills, although research has not found decisive differences between various methods (Cs. Czachesz & Vidákovich, 1994). Nowadays the most frequently chosen teaching method is the "syllabic", one emphasizing syllables both in speech and in written texts.

Having analyzed the strategic questions and problems of reading instruction, Nagy (2007) proposed a model of fostering text comprehension. His model is part of a wider theoretical framework called competency-based criterion-referenced formative evaluation. Nagy (2007) paid special attention to the components of cognitive competences necessary for fluent reading: speech-sound identification, reading sight-words, and reading vocabulary. He emphasizes the importance of achieving an adequate level of basic reading skills before students can benefit from developing text comprehension strategies. In this spirit, Nagy's collaborators (Pap-Szigeti, Zentai, & Józsa, 2006) developed a text-comprehension intervention program for fifth and sixth grade students. The program addressed several components of text-comprehension including self-regulatory elements of sentence- and text-level comprehension.

On the other hand, international survey studies (e.g., PISA) offer teachers and lay people a different conception of reading, that is, from a narrow "enthusiastic enjoyment of belles letters" (Molnár, 2006) to a definition of reading literacy as a goal-directed process that serves knowledge acquisition and developing one's potential. The PISA reading literacy definition – "*Reading literacy* is understanding, using, and reflecting on written texts, in order to achieve one's goals, to develop one's knowledge and potential and to participate in society" (OECD, 2006, p. 46.) – contains implicit reference to self-regulatory and metacognitive elements of reading, thus providing a convincing and supporting conceptual background for reading intervention programs.

The two intervention programs presented in this chapter emphasize the role of meta-level components in reading. This line of research can be considered as rather pioneering in Hungary and more in line with the international research community that focuses on metacognitive processes of reading comprehension. Moreover, since there is evidence for successful instruction of metacognitive strategies (Brand-Gruwel et al., 1998; Gaultney, 1995) the present research focused on fourth graders.

The first intervention program aimed at providing empirical evidence on the side of metacognition-based strategy instruction. Since mathematics and reading have been in the focus of previous international classroom intervention efforts, we conducted a pioneering experiment in our county in both fields. The results of Study 1 led to Study 2 which focused on reading strategies only, eliminating any interference and transfer that might have originated from mathematics strategy instruction.

2 Study 1

The rationale for the present intervention programs, which aimed at developing students' metacognitive strategies in reading comprehension, was the following: (a) Even at early elementary school years students possess a repertoire of reading strategies, consequently it is possible and, even necessary, to provide them with the opportunity to learn about reading strategies. (b) Intervention programs should be fully integrated into classroom settings in order to increase ecological validity. As Brown (1992) exemplified in the case of reciprocal teaching, research that began with a one-to-one laboratory-based setting became more complex when it was incorporated into the dynamics of a classroom.

The design of Study 1 involved two parallel learning units, one in mathematics and one in reading, taught over the same 2-month long period, and embedded in the Mathematics and Reading courses, respectively. Although there is empirical evidence supporting the domain-general character of metacognitive strategies (Veenman, Elshout, & Meijer, 1997), when it comes to the teaching of strategies in the classroom within a specific subject-matter it is necessary to embed them in the course content. Therefore, it was decided to use two different sets of metacognitive strategies, one in mathematics and one in reading. The idea that metacognitive activities can be matched in different subject-matters is in line with van der Stel and Veenman's (2008) study, in which similar domain-specific strategies in mathematics and in history were identified. In Study 1 conceptually similar metacognitive strategies were used in mathematics and reading. The metacognitive strategies targeted were clustered into three groups, similar to the categorization made by Almasi (2003) and Pressley and Gaskins (2006) with respect to reading, namely before, during, and after reading. The two sets of strategies are shown in Table 16.1.

The hypothesis of Study 1 was that the experimental group that participated in the metacognition-based intervention program in the fields of mathematics and reading would outperform the control group that attended the regular curriculum in both subject-matters.

Table 16.1 Structure of metacognitive strategies in the intervention program of Study 1

Lesson number	Mathematics	Lesson number	Reading
1	Interpretation of results	1	Text anticipation strategies
2		2	
3		3	
4	Creating mental models of the problem situation	4	Activating prior knowledge, skimming
5		5	
6		6	
7	Planning solution	7	Text maintenance strategies
8		8	
9		9	
10	Evaluation of solutions, searching for errors	10	Fix-up strategies
11		11	
12		12	
13	Integration phase	13	Integration phase
14		14	
15		15	

2.1 Method

2.1.1 Sample: Design

The study involved 244 Grade 4 students (their age ranged from 10 to 11 years) from eight schools from a pool of 20 schools which had an above the average rate of lower SES students. From each of four randomly selected schools one experimental class participated in the intervention; these four experimental classes formed the experimental group. The four control schools were randomly selected from the other 16 schools; all Grade 4 classes of the control schools formed the control group. The experimental group consisted of 86 students, and a matched sample was selected from the 158 students of the control group.

The experimental and control classes completed the same tests at the pre- and posttest on reading and mathematics. The two pretests that were re-administered at the posttest aimed at testing the near transfer effects of the intervention. Two additional instruments, one for mathematics and one for reading, were administered to test for far transfer effects.

2.1.2 Tasks and Measures

Mathematics Achievement Test

A paper-and-pencil instrument was developed that covered the major topics of the fourth-grade mathematics curriculum as defined by the Hungarian National Core Curriculum: arithmetic, geometry, and measurement. It had two versions sharing the same structure but with different data in the tasks (in order to assure individual work and avoiding cooperation between desk-mates), and it consisted of 36 items. An example for a simple geometry task: “The schoolyard is of square shape. Its perimeter is 348 meters. How long is its side?” At the pretest, the Mathematics Achievement Test had high internal consistency, with Cronbach’s alpha reliability coefficients ranging between 0.83 and 0.94, depending on the sub-sample (experimental vs. control) and the version of the test. At the posttest, the reliability coefficients ranged between 0.90 and 0.93.

Reading Test I

This test contained open-ended and true-false questions (26 items in total) tapping the content of four document-type texts. According to Elley (1994, p. 6), documents are “structured presentations of information, set out in the form of graphs, charts, maps, lists, or sets of instructions.” The first task contained the text of the wrapping paper of a chocolate bar. The second text was about geographical data, while the third text was borrowed from the sachet of a vanilla pod. The fourth text was the instructions for use of a children shampoo. All texts were followed by questions with different function: from testing the mere recall of information to testing the reflective understanding of the purpose and structure of the given text. Cronbach's alpha reliability coefficients of the 26-item test were 0.77 for the experimental, and 0.81 for the control group at the pretest, and 0.82 and 0.80 at the posttest, respectively.

Mathematical Word Problems Test

Ten "problematic" tasks from Verschaffel, De Corte, and Lasure (1994) were used at the posttest (see also Csíkos, 2003). Scoring of these tasks was based on whether students gave so-called "realistic reactions". In these tasks, a realistic reaction can be either an estimation of the solution or an explicit declaration that the task is unsolvable. Realistic reactions were considered as good solutions, whereas meaningless executions of arithmetic operations were considered wrong. Cronbach's alpha reliability coefficients of the Mathematical Word Problems Test were 0.82 for the experimental and 0.64 for the control group.

Reading Test II

It consisted of four texts (half page each): one narrative text on how to cook milk loaf, two explanatory texts on historical topics, and one document-type text on a newly developed type of wheel-chair. Each text was followed by open-ended questions, 16 questions altogether. For example, the document-type text was the recipe of the Advent milk loaf, and the text was followed by four questions consisting of 19 items, 13 of which required simple information retrieval, whereas the remaining six items required drawing inferences. Cronbach's alpha reliability coefficient was 0.96 for both the experimental and the control groups.

2.1.3 The Intervention Program

Two different sets of material addressing metacognitive strategies were prepared, one for mathematics, and one for reading. The strategies targeted and the lessons devoted to each type of strategy are given in Table 16.1. The first six lessons focused on orientation and planning strategies, and the next four on monitoring processes. Two lessons were devoted to fix-up strategies which addressed metacognitive evaluation processes. The final three lessons were devoted to the integration of the previously learnt strategies (see Table 16.1).

Although there were three clusters of strategies, namely planning, monitoring, and evaluation in both mathematics and reading, the domain-specific characteristics of each subject-matter dictated differentiation of the form and of the sequence of strategies. In reading, the sequence of text anticipation, text maintenance, and fix-up strategies reflected the usual phases of the (meta)cognitive strategies as described in previous sections. In mathematics, the sequence was determined based on data from a previous study (Csíkos, 2003), with the first three lessons focusing on interpretation of results of easy word problems that are solved with overly automatized strategies. Thus, orientation strategies were trained through reflection on what the problem requirements were and what was found.

All the training tasks of the intervention program in both mathematics and reading were developed so that they would facilitate the use of metacognitive strategies. In reading, lessons contained (a) "usual" narrative or explanatory texts followed by questions that required reflection on either text characteristics or students' own comprehension processes; (b) document-type texts that might shape students' beliefs about reading in general and about the variety of possible text comprehension phases and processes.

2.1.4 Procedure

The intervention program spanned over 15 lessons, with each lesson lasting 45 min as customary in Hungarian schools. In the experimental groups, the teachers proceeded with the regular curriculum topics in the first half of each lesson and then went on to use the intervention tasks in the second half (about 20 min). Consequently, in reading, the intervention program left less time for the usual loud-reading practice, since the latter emphasizes fluency and pronunciation–often at the cost of comprehension. Similarly, in mathematics, the regular 'drill' practice on solving numerous word problems was forced back.

The activities of the intervention program shared the most important characteristics of intervention programs conducted with young children in previous studies such as working on problem solving, working in small collaborative groups, and making metacognitive strategies explicit through self- and shared-regulation processes of planning, monitoring and evaluation of problem-solving activities (see Whitebread et al., 2009). The tasks of the intervention programs were developed to evoke and enrich different metacognitive strategies. For example, on the fifth lesson students worked in groups and discussed what kind of written information they would find on the wrapping paper of a chocolate bar, on a cinema ticket or on a poster advertising a football match. After that discussion the teachers did show students wrapping papers, cinema tickets and posters, and the whole class evaluated how precisely they could predict the information.

Teachers explicitly named the strategies practiced during task processing in each lesson and particularly in the closing (integration) phase.

The principles of the intervention program and its lesson plans were introduced to the teachers by means of personal communication (about 1 h). The written lesson plans were given to the teachers of the experimental classes 2 weeks before the intervention took place for them to study, and their questions and comments were discussed at a second meeting. The pretests were administered in the classrooms by the class teachers, and the intervention program started after a second meeting with the teachers. As regards the control classes, there was no personal contact with the teachers, and the pre- and posttests were posted to the schools at the same time as in the case of the experimental classes. In both the experimental and the control classes, both the pre- and posttests were completed during regular classes of Mathematics and Reading according to written instructions given to the teachers as to how the test administration procedure be accomplished. The instructions were to work on the tests as a diagnostic evaluation tool received from educational researchers.

2.2 Results

The students from the experimental group had higher scores than those from the control group on the pretests, $F(1, 242)=31.95$, $p<0.001$, partial $\eta^2=0.12$, and $F(1, 242)=19.69$, $p<0.001$, partial $\eta^2=0.08$, for the mathematics pretest and for the reading pretest, respectively. It was decided to keep only 86 participants from the control group in order to match the starting level and distribution of achievement scores at both groups. The basis for matching the experimental and the control group was their mathematics achievement. Based on scores on the mathematics achievement test, we formed five ability groups, and, using the unrestricted random sampling method, excluded students from the control group to have approximately the same number of students in each ability group. Thus, the experimental and control group formed did not differ in both mathematics and reading pretest scores, $F(1, 170)=3.80$, $p>0.05$, and $F(1, 170)=2.35$, $p=0.13$, respectively. The distribution of achievement scores in the five ability groups did not differ in the experimental and control groups, Kolmogorov-Smirnov's $Z=0.76$, $p=0.61$.

Table 16.2 shows the means and standard deviations of pre- and posttest achievement scores. Both the experimental and the control groups had higher posttest scores, as compared to pretest scores, on the Mathematics Achievement Test and the Reading Test I, as the paired-sample *t*-tests showed, $t(85)=5.10$, $p<0.01$, Cohen's $d=0.55$, and $t(85)=4.03$, $p<0.01$, Cohen's $d=0.43$, for the mathematics test and reading test, respectively, in the control group; and $t(85)=4.99$, $p<0.01$, Cohen's $d=0.54$, and $t(85)=4.91$, $p<0.01$, Cohen's $d=0.53$, in the experimental group.

2.2.1 Intervention Effects

To investigate the intervention effects on the near and far transfer tests, separately for reading and mathematics, MANCOVAs on the scores of the two posttests were applied with group as between subjects factor. The respective pretest score was the covariate.

On the reading posttests, the MANCOVA revealed significant differences between the experimental and control groups, Pillai's trace $=0.05$, $F(2, 169)=4.12$, $p=0.02$, partial $\eta^2=0.05$. The pretest effect was robust, with Pillai's trace $=0.60$, $F(2, 169)=123.47$, $p<0.001$, partial $\eta^2=0.60$. The group effect on the mathematics

Table 16.2 Pre- and posttest achievement scores in Study 1 ($n=86$ for each group)

	Experimental group		Control group	
	M	*SD*	*M*	*SD*
Pretest				
Mathematics achievement test	21.55	7.53	19.19	8.34
Reading test I	14.95	3.64	14.02	4.29
Posttest				
Mathematics achievement test	23.57	7.85	21.06	7.93
Mathematics word problems test	4.13	2.72	1.90	1.54
Reading test I	16.55	4.17	15.13	3.84
Reading test II	39.95	15.02	33.34	12.30

posttest scores was also significant, Pillai's trace $=0.23$, $F(2, 169)=24.95$, $p<0.001$, partial $\eta^2=0.23$ The pretest effect was also significant, with Pillai's trace$=0.81$, $F(2, 169)=359.51$, $p<0.001$, partial $\eta^2=0.81$.

The univariate *F*-tests showed significant differences between both reading tests. Specifically, for Reading Test I, $F(1, 170)=5.38$, $p=0.02$, partial $\eta^2=0.03$, and for Reading Test II, $F(1, 170)=9.99$, $p<0.01$, partial $\eta^2=0.06$. There were also significant differences in both mathematics tests with $F(1, 170)=4.36$, $p=0.04$, partial $\eta^2=0.03$ for the Mathematics Achievement Test, and $F(1, 170)=43.92$, $p<0.001$, partial $\eta^2=0.21$ for the Word Problems Test.

According to Cohen (1969), interpretation of effect sizes of the "proportion of variance" type (see Olejnik & Algina, 2000) can be based on certain milestones, that is, effect size of 0.01 is small, 0.06 is medium, and 0.15 is large. Therefore, in reading the effect sizes were small to medium, whereas for mathematics were low in the Mathematics Achievement Test and large in the Mathematical Word Problems Test, that is, in the far transfer test.

2.3 Discussion

Study 1 produced evidence supporting the idea of metacognition-based strategy instruction in reading and mathematics. Our brief intervention program contained 15 lessons in both Reading and Mathematics embedded in classroom settings. The lesson plans of the intervention program contained various problems often different from the 'usual' classroom tasks. Especially in reading there were several types of document texts provided with tasks and questions facilitating conscious use of reading strategies. The significant positive effects of the program raised questions about different factors that may have contributed to the findings. Possible transfer effects between the reading and mathematics program and motivational effects of the new types of texts and tasks might have positive effects on students' achievement.

In Study 2 the aim was to eliminate the likely transfer effects of the mathematics part of the program and the motivational effects gained from new types of texts. Moreover, an effort was made to enrich the intervention with more explicit metacognitive knowledge on reading strategies and with more time for practice. Thus, a new intervention program was developed only for reading.

3 Study 2

Building on the results and experiences gathered from Study 1, an improved intervention program was developed for reading comprehension. The main difference between the first and second program can be described in terms of four changes: (a) A longer "integration phase" that lasted for more than half of the time of the intervention; (b) there was a new cluster, namely declarative metacognitive knowledge that took place in the first half of the program; (c) new strategies of a given week were introduced on the first lesson of that week, and the next three lessons served for practicing those strategies; and (d) the program lasted 8 weeks including 32 lesson units. All these changes were expected to enhance the outcomes of the intervention program.

Hence, the hypothesis in Study 2 remained the same as in Study 1. Specifically, the experimental group that participated in the metacognition-based intervention program was expected to outperform the control group that attended the regular curriculum.

3.1 Method

3.1.1 Sample

Study 2 involved 158 students (79 boys and 79 girls) from nine fourth grade (their age ranged from 10 to 11 years) classes in Kecskemét, Hungary (county seat town with about 100,000 inhabitants) and its surrounding areas. There were five experimental and four control classes with 94 and 64 students, respectively. The distribution of students' gender and of their SES status was representative of the whole county population.

3.1.2 Tasks and Measures

Two different reading comprehension tests, developed by the second author, were used. One of them was used as pretest and the other as posttest. Both of them included 14 open-ended questions on two explanatory non-fiction texts. For example, the first part of the posttest contained a five paragraph description of the red kangaroo followed by seven questions. Four out of the seven questions required simple information recall and two other questions required drawing inferences. The last question addressed reflective evaluation of the text structure, that is, "To how many parts would you divide the text according to its content? Give titles to the parts." Cronbach's alpha reliability coefficients were 0.74 for the pretest, and 0.71 for the posttest. A student background questionnaire was also administered, containing questions about the students' academic achievement and socio-cultural background. Socio-cultural background was recorded because it has been shown to impact reading achievement in national and international surveys (see for example Báthory, 1992; Elley, 1994).

3.1.3 The Intervention Program

Table 16.3 summarizes the instruction given to the experimental group. Compared to Study 1, a more detailed repertoire of strategies was included in the program. We started the program with strategies belonging to the orientation and planning cluster of reading processes, then continued with monitoring and maintenance strategies, and finished with evaluation strategies. As it is indicated in the last row of Table 16.3, teachers in the first 4 weeks explicitly addressed students' metacognitive knowledge of strategies, that is, conditional knowledge of strategy use and, particularly, what they were doing and why those strategies were useful for them. Explicit naming of reading strategies was expected to enable students to elaborate their previous experiences about the efficiency of reading strategies, and connect them to the current learning situation.

3.1.4 Procedure

Teachers received instructions from the researchers before starting the intervention program, and were given lesson plans. Students' regular reading book texts were used during the intervention, avoiding any extra motivational or other interfering effect that new and not "canonized" texts may have caused. On the first lesson of each week, teachers explicitly named the newly introduced strategies, explained the use and benefits of those strategies, and asked the students about their previous experience of the strategy. In this way metacognitive knowledge of strategies was made explicit. The next three lessons of each week served the practicing of the strategies by means of collaborative group work.

In the first week children learnt and practiced pre-reading phase (see Pressley, 2002) strategies. From the second week on strategies of the process of meaning construction were also introduced. These constructively responsive strategies (see Pressley & Gaskins, 2006) – often labeled as strategies of "process reading" – have recently appeared in Hungarian reading instruction (Adamikné, 2001). In short, when using this method, students' expectations about the topic of the text are evoked before and during reading. The teacher stops the children at important points of the text (depending on its length and complexity) and asks them whether their expectations about the text are met and how they expect the text to continue. Two other

Table 16.3 The structure of the improved reading-strategy intervention program in Study 2

Strategies	1st week	2nd week	3rd week	4th week	5th week	6th week	7th week	8th week
Activating scheme	x	x	x	x	x	x	x	x
Preview	x	x	x	x	x	x	x	x
Scanning	x	x	x	x	x	x	x	x
Process reading		x	x	x	x	x	x	x
Text anticipation		x	x	x	x	x	x	x
Checking predictions		x	x	x	x	x	x	x
Creating sensory images			x	x	x	x	x	x
Summarizing				x	x	x	x	x
Conclusion				x	x	x	x	x
Synthesizing				x	x	x	x	x
Metacognitive knowledge	x	x	x	x				

strategies, text anticipation and its correction, were also taught during the second week. Children formed expectations about the passage at hand and then they checked whether or not the text fulfilled their expectations, as well as how and why the text continued in the way it did.

In the third week children practiced creating mental images related to the text they were reading. According to this method children are to imagine and elaborate what they have experienced or seen with their inner eyes. They then predict what is going to happen and after reading the passage they confirm, criticize or disagree with it (Wood, 2001). In the fourth week, strategies that belong to the "after reading" cluster (see Pressley, 2002) of strategies were introduced.

3.2 Results

The analyses were restricted to the 158 participants (out of the 182 students involved) who answered both the pre- and posttest. Table 16.4 shows the means and standard deviations of pre- and posttest scores.

The results of the pretest did not show significant differences between the control and the experimental groups, $F(1, 157)=0.72$, $p=0.40$. However, there were significant differences in the results of the posttest, $F(1, 157)=9.30$, $p<0.01$. Students in the experimental classes achieved better comprehension results after completing the program. The effect size was 5.7% (Cohen's $f=0.24$, $f^2=0.057$). The group effect on the posttest was analyzed by ANCOVA with pretest scores as covariates. The group effect was significant, $F(1, 157)=17.23$, $p<0.001$, partial $\eta^2=0.10$, and also the pretest effect, $F(1, 57)=64.21$, $p<0.001$, partial $\eta^2=0.29$. The comprehension scores of students in the experimental group surpassed those of the control classes. Gender differences did not play a significant role in student performance as shown by the 2(gender)×2(group) ANOVA, $F(1, 157)=0.03$, $p=0.88$ for the posttest. Finally, correlations between test scores and other background variables were not significant.

3.3 Discussion

Study 2 provided evidence for the effectiveness of metacognition-based strategy reading instruction. The intervention program lasted for 8 weeks embedded in classroom work. The lesson plans of the intervention program contained 'regular' reading book

Table 16.4 Descriptives of the pre-/posttest in Study 2

	Experimental group ($n=94$)		Control group ($n=64$)	
Testing phase	*M*	*SD*	*M*	*SD*
Pretest	22.14	5.77	22.86	4.32
Posttest	19.76	5.70	17.12	4.71

texts, explicitly addressing students' previous experience about using different reading strategies. The results supported the hypothesis that the students of the experimental group would perform better on the reading comprehension posttest than those of the control group. The low correlations between reading comprehension scores and different background variables and the lack of a gender effect suggest wide generalizability of the positive experimental effects of the intervention program.

4 General Discussion

The main aim of the present study was to develop ecologically valid and effective reading intervention programs that can provide empirical evidence on the importance of metacognitive strategy use in reading comprehension. Having reviewed some relevant theories of reading and aiming at connecting them to Nelson's conceptualization of metacognition, we posited that reading processes belong to two levels, namely the level of cognitive processes (e.g., decoding skills), and the meta-level of strategy use. Our metacognition-based reading intervention programs targeted the latter, the strategy level of reading processes. The development of the intervention programs was based on both theoretical considerations and on empirical evidence from studies on explicit reading strategy instruction.

Study 1 combined mathematical and reading intervention programs. The results on reading suggested that the training program was successful but the effect sizes on the two reading tests were small to medium.

Study 2 was a more elaborated version of the reading part of Study 1. Different structure and different emphases were introduced in an 8-week intervention program. A medium level of effect size was found again, with no relationship between the performance scores and background variables. In general, the intervention program proved to be effective irrespective of students' gender and academic achievement.

Our results suggest that in Grade 4, Hungarian elementary school students benefited from the explicit training of their planning, monitoring and control processes in reading. The results support the claim that the conscious and flexible use of different reading strategies will result in better achievement on reading comprehension tests. However, despite the greater emphasis on metacognitive knowledge and practice on reading strategies in Study 2, the effect size was moderate suggesting that other factors, such as the decoding or comprehension skills themselves are crucial for performance on reading comprehension test. This conclusion is supported by the fact that at the posttest the control group had also higher performance compared to the pretest. In essence, instruction on strategy use seems to improve students' control over their comprehension processes.

The findings and their possible implications are in line with previous research in the field of reading strategy instruction. Reading strategy instruction in classroom settings can be a powerful tool to enhance students' reading achievement irrespective of their gender and their previous achievement level. The latter finding points to the importance of possible future discussions on whether explicit strategy instruction would be helpful for struggling readers, as well as for normal ones.

Classroom-based experiments of the brief intervention type hold an optimistic point of view in that it is possible to reveal significant and, what is more, educationally relevant changes during a period of few weeks or months. The development of reading comprehension (from a less developed learner status to a mature reader level) requires several years of progress in both reading decoding and reading strategy use. Our assumption is that a short intervention program may not develop new procedural knowledge of reading strategies but may enable even struggling readers to be aware of their strategy use. It is well documented (see Myers & Paris, 1978) that even second graders have a lot of declarative metacognitive knowledge about reading, and according to our results with the Index of Reading Awareness questionnaire (Csíkos, 2008), Hungarian fourth graders do possess differentiated beliefs about reading and about themselves as readers. The role of and possible changes in beliefs during a short classroom intervention program requires further investigations.

Reading instruction is facing challenges in how strategic level components can contribute to the development of reading comprehension. Study 1 showed the possibility of effective intervention in reading and mathematics by means of strategy instruction. Study 2 provided evidence about the positive effect of explicit reading strategy instruction with emphasis on "how?" questions that facilitated use of reading strategies, while questions of the type "what to read?" were kept in the background. More evidence about the best instructional practices on reading strategies is needed, by taking into account the effect of text genre on the selection of the most appropriate reading strategy each time. In our studies, we followed an integrative approach to instructional practices and focused on strategies that are genre-free. More experimental research is needed on which aspect of instruction (e.g., collaborative work or reflection) is best fit for each specific strategy.

Acknowledgements The research presented in this paper was supported by the Hungarian National Science Fund (OTKA) projects F38222 and 63360 to the first author. Study 1 was partly supported by the Research Group on Abilities of the Hungarian Academy of Science. Research collaborators at Study 1 were Klára Tarkó, István Zsigmond, and Éva Molnár. The reading comprehension test of Study 1 was developed by Monika Dózsa.

We thank Edit Katalin Molnár for her helpful comments on a previous version of the manuscript.

Special thanks are due to the teachers who participated in the projects.

Parts of the present study have been presented at conferences (Csíkos, 2005a; Steklács & Csíkos, 2007). Also, part of the data have been published in Csíkos (2005b) and Steklács and Csíkos (2009).

References

Adamikné, J. A. (Ed.). (2001). *A magyar olvasástanítás története* [The history of reading education in Hungary]. Budapest: Osiris Kiadó.

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61*, 364–373.

Almasi, J. (2003). *Teaching strategic processes in reading*. New York: Guilford.

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, R. Barr, M. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353–394). New York: Longman.

Báthory, Z. (1992). *Tanulók, iskolák, különbségek – Egy differenciális tanításelmélet vázlata* [Students, schools and differences – An outline of a differential theory of instruction]. Budapest: Tankönyvkiadó.

Brand-Gruwel, S., Aarnoutse, C. A. J., & Van den Bos, K. P. (1998). Improving text comprhension strategies in reading and listening settings. *Learning and Instruction, 8*, 63–81.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*, 141–178.

Campione, J. C., Brown, A. L., & Connel, M. L. (1988). Metacognition: On the importance of understanding what you are doing. In R. I. Charles & E. A. Silver (Eds.), *Research agenda for mathematics education: The teaching of mathematical problem solving* (pp. 93–114). Hillsdale, NJ: Erlbaum.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, UK: Cambridge University Press.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* New York: Academic.

Cromley, J. G., & Azevedo, R. (2006). Testing and refining the Direct and Inferential Mediation (DIME) model of reading comprehension. *Journal of Educational Psychology, 99*, 311–325.

Cross, D. R., & Paris, S. G. (1988) Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology, 80*, 131–142.

Cs. Czachesz, E., & Vidákovich, T. (1994). Melyik módszer? Öt olvasástanítási módszer összehasonlító vizsgálata [Which method? A comparative study of five reading instruction methods]. *ÉKP Hírek, 6–7*, 6–12.

Csíkos, C. (2003). *How many buses are needed? Hungarian students' achievement on 'problematic' word problems.* Paper presented at the 10th European Conference for Research on Learning and Instruction, Padova, Italy.

Csíkos, C. (2005a). *A metacognition-based training program in grade 4 in the fields of mathematics and reading.* Paper presented at the 11th European Conference for Research on Learning and Instruction, Nicosia, Cyprus.

Csíkos, C. (2005b). Metakogníciókra alapozott fejlesztő kísérlet 4. osztályos tanulók körében a matematika és az olvasás területén [A metacognition-based training program in grade 4 mathematics and reading.]. *Magyar Pedagógia, 105*, 127–152.

Csíkos, C. (2008). *A large scale 3-year longitudinal study on Jacobs and Paris' IRA questionnaire among 3rd-5th grade students*. Paper presented at the 3rd Biennial Meeting of the EARLI Special Interest Group 16 (Metacognition), Ioannina, Greece.

Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation, and self-regulation. In A. Efklides, J. Kuhl, & R. M. Sorrentino (Eds.), *Trends and prospects in motivation research* (pp. 297–323). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.

Elley, W. B. (Ed.). (1994). *The IEA study of reading literacy: Achievement and instruction in thirty-two school system.* Oxford, UK: Pergamon.

Flavell, J. H. (1971). First discussant's comments: What is memory development the development of? *Human Development, 14*, 272–278.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34*, 906–911.

Garner, R. (1987). *Metacognition and reading comprehension.* Norwood, NJ: Ablex.

Gaskins, I. W. (1994). Classroom applications of cognitive science: Teaching poor readers how to learn, think, and problem solve. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 129–156). Cambridge, MA: The MIT Press.

Gaultney, J. F. (1995). The effect of prior knowledge and metacognition on the acquisition of a reading comprehension strategy. *Journal of Experimental Child Psychology, 59*, 142–163.

Hacker, D. J. (1998). Definitions and empirical foundations. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 1–23). Mahwah, NJ: Erlbaum.

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.). (1998). *Metacognition in educational theory and practice.* Mahwah, NJ: Erlbaum.

Hall, K., Bowman, H., & Myers, J. (1999). Tasks, texts and contexts: A study of reading and metacognition in English and Irish primary classrooms. *Educational Studies, 25*, 311–325.

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist, 22*, 255–278.

Keene, E. O., & Zimmermann, S. (1997). *Mosaic of thought: Teaching comprehension in a reader's workshop.* Portsmouth, NH: Heinemann.

Kentridge, R. W., & Heywood, C. A. (2000). Metacognition and awareness. *Consciousness and Cognition, 9*, 308–312.

Kluwe, R. H. (1987). Executive decisions and regulation of problem solving behavior. In F. E. Weinert, & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 31–64). Hillsdale, NJ: Erlbaum.

Koriat, A., & Levy-Sadot, R. (2000). Conscious and unconscious metacognition: A rejoinder. *Consciousness and Cognition, 9*, 193–202.

Meloth, M. S., & Deering, P. D. (1992): Effects of two cooperative conditions on peer-group discussions, reading comprehension, and metacognition. *Contemporary Educational Psychology, 17*, 175–193.

Meneghetti, C., Carretti, B., & De Beni, R. (2006). Components of reading comprehension and scholastic achievement. *Learning and Individual Differences, 16*, 291–301.

Molnár, E. K. (2006). Olvasási képesség és iskolai tudás [Reading ability and school-based learning]. In K. Józsa (Ed.), *Az olvasási képesség fejlődése és fejlesztése* (pp. 43–60). Budapest: Dinasztia Tankönyckiadó.

Myers II, M., & Paris, S. G. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology, 70*, 680–690.

Nagy, J. (2007). *Kompetencia alapú kritériumorientált pedagógia* [Competency-based and criterion-referenced education]. Szeged, Hungary: Mozaik Kiadó.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102–116.

Nelson, T. O., & Narens, L. (1992). Metamemory: A theoretical framework and new findings. In T. O. Nelson (Ed.), *Metacognition. Core readings* (pp. 117–129). Boston, MA: Allyn and Bacon.

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Application, interpretations, and limitations. *Contemporary Educational Psychology, 25*, 241–286.

Organisation for Economic Co-operation and Development. (2006). *Assessing scientific, reading and mathematical literacy. A framework for PISA 2006.* Paris: Organisation for Economic Co-operation and Development.

Palincsar, A. S., & Brown, A. L. (1985). Reciprocal teaching: Activities to promote read(ing) with your mind. In T. L. Harris & E. J. Cooper (Eds.), *Reading, thinking and concept development: Strategies for the classroom.* New York: The College Board.

Pap-Szigeti, R., Zentai, G., & Józsa, K. (2006). A szövegfeldolgozó képességfejlesztés módszerei [Methods for improving abilities through text processing]. In K. Józsa (Ed.), *Az olvasási képesség fejlődése és fejlesztése* (pp. 235–258). Budapest: Dinasztia Kiadó.

Perfetti, C. A. (1985). *Reading ability.* New York, NY: Oxford University Press.

Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. III, pp. 545–561). Mahwah, NJ: Erlbaum.

Pressley, M. (2002). Metacognition and self-regulated comprehension. In A. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 291–309). Newark, DE: IRA.

Pressley, M., & Gaskins, I. W. (2006). Metacognitively competent reading comprehension is constructively responsive reading: How can such reading be developed in students? *Metacognition and Learning, 1*, 99–113.

Schellings, G., Aarnoutse, C., & van Leeuwe, J. (2006). Third-grader's think-aloud protocols. Types of reading activities in reading an expository text. *Learning and Instruction, 16*, 549–568.
Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*, 32–71.
Steklács, J., & Csíkos, C. (2007). *Improving students' reading comprehension by means of developing reading strategies. Results from a Hungarian experiment.* Paper presented at the 12th European Conference for Research on Learning and Instruction, Budapest, Hungary.
Steklács, J., & Csíkos, C. (2009). Developing reading strategies among 4th grade students in Hungary. *The New Mexico Journal of Reading, 19(3)*, 15–19.
Thorndike, R. L. (1973). *Reading comprehension education in fifteen countries: An empirical study.* New York: Wiley.
van den Bos, K. P., Brand-Gruwel, S., & Aarnoutse, C. A. J. (1998). Text comprehension strategy instruction with poor readers. *Reading and Writing: An Interdisciplinary Journal, 10*, 471–498.
van der Stel, M., & Veenman, M. V. J. (2008). Relation between intellectual ability and metacognitive skillfulness as predictors of learning performance of young students performing tasks in different domains. *Learning and Individual Differences, 18*, 128–134.
Van Kraayenoord, C., & Schneider, W. (1999). Reading achievement, metacognition, reading self-concept and interest: A study of German students in Grade 3 and 4. *European Journal of Psychology of Education, 14*, 305–324.
Veenman, M. V. J., Elshout, J. J., & Meijer, J. (1997). The generality vs. domain-specificity of metacognitive skills in novice learning across domains. *Learning and Instruction, 7*, 187–209.
Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modelling of school arithmetic word problems. *Learning and Instruction, 7*, 339–359.
Walczyk, J. J. (1995). Testing a compensatory-encoding model. *Reading Research Quarterly, 30*, 396–408.
Whitebread, D., Coltman, P., Pasternak, D. P., Sangster, C., Grau, V., Bingham, S., Almeqdad, Q, & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning, 4*, 63–85.
Wood, K. D. (2001). *Literacy strategies across the subject areas.* Boston, MA: Allyn & Bacon.

Chapter 17
Metacognition and Spelling Performance in College Students

Ruth Vanderswalmen, Joke Vrijders, and Annemie Desoete

1 Introduction

Proficient spelling is crucial in convincing someone of your expertise (Harris, Graham, Brindle, & Sandmel, 2009). The volume of studies on spelling in younger children shows the importance of the topic (see, Defior, Jimenez-Fernandez, & Serrano, 2009; Landed, Thaler, & Reitsma, 2008; Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008; Verhoeven, Schreuder, & Baayen, 2006; Wakely, Hooper, de Kruif, & Swartz, 2006). Several of these studies show that being proficient at the lower levels of writing skills, such as spelling, helps to ease the demands on working memory when writing. When students allocate their working memory resources to figuring out how to spell a word, they may forget what ideas they were going to write next (Carlisle, 1994). Also Wakely et al. (2006) found that students who had more problems with spelling wrote a rather undeveloped story, that is, a story with sentences that described more than one event but with few details about the setting. They conjectured this may be due to a lack of automaticity in spelling, which undermines students' ability to produce ideas fluently and disrupts their composition of sentences and their monitoring of the writing process.

Students seem to have increasing difficulties with spelling nowadays compared to the past (Claes & Moeyaert, 2003). A study by Herbots (2005) revealed that one out of three university students could not write a short text without making some spelling errors. In the Netherlands, we also see ominous messages in the media: 68% of first-year students undergoing teacher training fail a test in their mother tongue (Grezel, 2007). Harris et al. (2009) described a similar case. Specifically, only 25% of the students in the United States were classified as competent writers. In addition, almost one in every five first-year college students in the United States requires a remedial writing class and more than a half of new college students are unable to write a paper relatively free of errors. Most importantly, spelling errors

R. Vanderswalmen (✉)
University College Arteveldehogeschool, Campus Kantienberg, Opleiding Logopedie, Voetweg 66, Gent B-9000, Belgium
e-mail: ruth.vanderswalmen@arteveldehs.be

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_17, © Springer Science+Business Media, LLC 2010

are not made in rarely used words; rather, basic errors in everyday words have become common in higher education.

What are the reasons for these weak written-language skills? There are many different reasons, but none of them, by itself alone, is sufficient to explain the phenomenon. Among the reasons advocated are, first of all, the priorities in language teaching. Nowadays, clusters such as grammar, spelling, and sentence composition receive less attention than in the past. Teachers' overemphasis on macro-level writing processes (i.e., planning, organization, and self-monitoring) and lack of emphasis on improving lower level skills necessary for writing (i.e., fluent handwriting, grammar, and spelling) are often reported (e.g., Hayes, 1996).

Another potential reason is the use of new communication technologies. According to Dutch teachers (Soenens, 2002) text messaging and instant messaging culture is the main 'culprit'. The impact of new communication technologies is not to be underestimated (Vlaamse Onderwijs Raad [VLOR], 2006). Due to the speed of communication, less attention is being paid to proper and appropriate language.

Third, students' spelling is not only insufficient in terms of prior knowledge and skills. In addition, students often lack the attitude and self-awareness of proficient spellers (Vrijders, Vanderswalmen, & Beeckman, 2007). The experience of teachers in higher education suggests that students cannot judge their own strengths and weaknesses correctly. For example, they make three verb errors in an e-mail but still say they almost never make spelling errors in verb spelling.

To sum up, although spelling receives a lot of research interest in the context of young students' emerging literacy skills, there is less research on older students' spelling skills and metacognitive awareness of their spelling behaviour. This chapter is focusing on the latter issue. In what follows, first we present a theory on spelling and a classification of spelling errors. The aim is to make explicit that spelling depends on phonological, morphological and lexical skills. Then, the facets of metacognition in relation to spelling are discussed. We claim that metacognitive experiences, metacognitive knowledge, and metacognitive skills are all involved in proficient spelling. Then, an empirical study is presented regarding the relations of the facets of metacognition with spelling performance and the implications for future research are discussed.

1.1 Spelling and Spelling Errors

Spelling depends on the appropriate translation of phonemes (sounds) into graphemes (letters) and on a proficient segmentation of graphemes (Steffler, Varnhagen, & Friesen, 1998). Transparent orthographic systems are characterised by high degree of consistency in the translation of phonemes into graphemes and are mainly governed by bi-univocal phoneme-to-grapheme correspondence (PGC) rules (Defior et al., 2009); in the bi-univocal rule there is one-to-one grapheme-to-phoneme correspondence. In contrast, opaque or deep orthographic systems, such as English, have graphemes with various corresponding phonemes and vice versa, with a large number of irregular, orthographically exceptional and inconsistent words (Verhoeven et al., 2006). French, Portuguese, and Danish are also orthographically deep languages (Verhoeven et al., 2006), whereas the Spanish orthographic code is characterised by high level of consistency (Defior et al., 2009). Spelling words that have regular phoneme-grapheme correspondences is influenced by phonological skills (Gentry, 1982; Henderson & Beers, 1980) in addition to orthographic knowledge (Templeton & Morris, 2000). Such words can be spelled by applying a phonological strategy because of the fully consistent relationships between phonemes and graphemes (e.g., <pen>, [pen] in Dutch). The same is true for consonant clusters, although correct segmentation is crucial in this case. For example, poor spellers often omit the consonant immediately following the vowel in consonant clusters (e.g., writing <stop> instead of [stomp] in Dutch; Van Bon & Uit De Haag, 1997). In addition, some words can be spelled via reasoning by analogy because of similar phonemes (e.g., [aai] in Dutch) or letter combinations (e.g., [cht] in Dutch). If children know how to spell <m*aai*en> and <l*uch*t>, then they can also spell <l*aai*end> and <z*ucht*en> through reasoning by analogy.

However, the orthographic depth hypothesis does not provide us with sufficient insights into the access to orthographic representations in the mental lexicon, because it is not fine-grained enough (Verhoeven et al., 2006). Learning to spell words without a regular phoneme-grapheme correspondence (i.e., morphological words) is more than merely memorizing letter sequences. Written Dutch also includes aspects of morphology that are not represented phonologically. For some words the environment of the phoneme is determinative for the manner of writing, and a rule-based approach is necessary (Keuning & Verhoeven, 2008). In order to arrive to a full understanding of the spelling processes, it is also necessary to take into account that spelling rules are not always directly governed by phonotactic rules. The reader must convert sounds to an underlying orthographic representation to which spelling adaption rules are applied, independent of the pronunciation (Verhoeven et al., 2006). In Dutch polysyllabic words there is the complicated grapheme-phoneme conversion rule, pertaining to vowel and consonant letter doubling. Long vowels in Dutch can be written in two ways, namely as two identical vowel letters as in <boom> or with a single vowel letter as in <b*o*men>. Dutch short vowels are represented by a single vowel letter (e.g., <bom>); in plural formation this consonant is geminated (e.g., <bo*mm*en>, i.e., "bombs") with a consonant geminate (e.g., [mm]) (see Verhoeven et al., 2006). The general rule is that the contrast between short and long vowels in open syllables is expressed by the alternation of single and double consonant letters.

Another morphological rule is also needed to write correctly (Sénéchal, Basque, & Leclaire, 2006). In Dutch word-ending devoicing is a systematic phonological process. For example, words like <bed> and <krab> are pronounced [bet] and [krap]. However, the orthography operates as though this devoicing did not take place. Writers have to use their morphological understanding of the relationship between <bed> and <bedden> and <krab> and <krabben> in order to spell accurately.

Knowledge of spelling rules appears to be critical in the ability to spell words without a regular phoneme-grapheme correspondence (Rittle-Johnson & Siegler, 1999).

However, the vowel reduction rule (needed to write<bomen>), the consonant doubling rule (needed to write<bommen>) and the word-ending devoicing rule (needed to write<bed>or<krab>) are not sufficient to spell lexical words without errors. Some words (so called "lexical words") can only be learnt by memorizing them because current spelling rules do not apply to them and analogical reasoning cannot offer a solution. In the case of the graphemes [au], [ou], and [ei], [ij], one just has to know which of the two alternatives is the correct one based on a visual imprint strategy. The same applies to the spelling of foreign words or loan words such as<mail>where PGC rules cannot be applied. In this case lexical knowledge is needed.

The development of a child's abilities underlying the spelling skills has been studied within several theoretical frameworks. It is often assumed that spelling skills and strategies are acquired during the learning process following a sequence of qualitatively distinct stages in which different sources of knowledge are used (Ehri, 1992; Henderson, 1992; Templeton & Bear, 1992). All stage theories presume a transition from relying on phonological properties of words to recognizing and representing orthographic and morphological regularities and rules (Keuning & Verhoeven, 2008). However, some researchers have suggested that variability of strategy use in spelling may be better described in terms of the general learning framework of "overlapping waves" as proposed by Siegler (2000).

1.2 Metacognition and Its Facets

Metacognition has been introduced to describe and explain how people gain control over their learning and thinking, particularly in the case of cognitive failures (Efklides & Sideridis, 2009; Flavell, 1976) and difficulties they meet when dealing with information processing and problem solving (Brown, 1980, 1987; Desoete & Veenman, 2006; Efklides, 2001; Flavell, 1976; Montague, 1998). The model of metacognition by Nelson and Narens (1990) has served as a theoretical framework for the conceptualisation of metacognition. Three principles underlie this model: (a) mental processes are posited to function at two levels, the cognitive (or object) level and the metacognitive level, (b) the metacognitive level represents a dynamic model of the cognitive level and (c) there are two dominant functions, namely control and monitoring, which are defined in terms of the direction of flow of information between the meta-level and the object-level. It is widely accepted that metacognition influences reading, writing, and text studying (Afflerbach, 1990; Nist, Simpson, & Olejnik, 1991; Otero, Campanario, & Hopkins, 1992; Pugalee, 2001; Van Kraayenoord & Schneider, 1999; Veenman & Beishuizen, 2004; Zhang, 2001). However, before looking at the relations of metacognition with spelling, a brief description of the facets of metacognition will be made in order to highlight the complexity of notion of metacognition and its relations with cognition.

Metacognition has been described as having three facets, namely metacognitive knowledge, metacognitive experiences and metacognitive skills (Efklides, 2001, 2008;

Flavell, 1979). *Metacognitive knowledge* has been described as the knowledge and deeper understanding of cognitive processes and products (Flavell, 1976). Children may know, for example, that they have to check their spelling after writing a text or email. According to Efklides (2008, p. 208) metacognitive knowledge is « declarative knowledge stored in the memory and comprises models of cognitive processes. It also encompasses information about people (including one's self), as well as information about tasks, strategies, and goals. Metacognitive task-knowledge involves task categories and their features, relations between tasks, as well as the ways they are processed. Metacognitive strategy-knowledge involves knowledge of multiple strategies as well as the conditions for their use (e.g., when, why and how a strategy should be used). Finally, metacognitive goal-knowledge involves knowledge of what sort of goals people pursue when confronted with specific tasks or situations.»

Another related conceptualization of metacognitive knowledge distinguishes declarative, procedural and conditional (or strategic) metacognitive knowledge. *Declarative metacognitive knowledge* is described as « what is known in a propositional manner » (Jacobs & Paris, 1987, p. 259) or the assertions about the world and the knowledge of the influencing factors (memory, attention and so on) of human thinking. *Procedural metacognitive knowledge* (also called "metacognitive strategies" or "metacognitive skills") can be described as « the awareness of processes of thinking » (Jacobs & Paris, 1987, p. 259), or « the knowledge of the methods for achieving goals and the knowledge of how skills work and how they are to be applied. Procedural knowledge is necessary to carry out procedures in order to apply declarative knowledge and reach goals » (Harris et al., 2009, p. 133). *Conditional or strategic metacognitive knowledge* is considered to be « the awareness of the conditions that influence learning such as why strategies are effective, when they should be applied and when they are appropriate » (Jacobs & Paris, 1987, p. 259). Conditional knowledge is critical to effective use of strategies (Harris et al., 2009). Novices have been found to possess poorer metacognitive skills than experts (Kruger & Dunning, 1999). Students doing poorly on tests predicted less accurately which questions they would get right than students doing well (Kruger & Dunning, 1999; Sinkavich, 1995)

Metacognitive experiences are « what the person is aware of and what she or he feels when coming across a task and processing the information related to it » (Efklides, 2008, p. 279). They take the form of metacognitive feelings, metacognitive judgments/estimates, and online task-specific knowledge. *Metacognitive feelings* are non-analytic representations of knowing states with an affective and cognitive character. The affective character of metacognitive experiences can be explained by two feedback loops. The first one is related to the outcome of cognitive processing and detects the discrepancy from the goal set. Error detection (as discrepancy from the goal) and feeling of difficulty (as lack of processing fluency) are associated with negative affect (Efklides, 2006). *Metacognitive judgments/estimates* include analytic and non-analytic processes, such as judgment of learning, estimate of effort expenditure, estimate of time needed or spent, but also estimate of solution correctness. When people are asked to make a judgment about their confidence there are two sources of information on which they rely, according to

Efklides (2008), namely their estimate of solution/response correctness (as discrepancy of the response to the goals) and their feeling of difficulty (as cue that the response might not be correct). Metacognitive experiences, in essence, make the person aware of his or her cognition and trigger control processes that serve the pursued goal of the self-regulation process (Efklides, 2008; Koriat, 2007). However, the person can feel highly confident, even if the outcome of cognitive processing is not correct, just because the solution was produced fluently, thus endangering appropriate control decisions. This is particularly true for persons who are not aware of their ignorance (Efklides, 2008; Kruger & Dunning, 1999).

Metacognitive skills refer to « the deliberate use of strategies (procedural knowledge) in order to control cognition » (Efklides, 2008, p. 280). According to Brown (1980), executive control (or "metacognitive skills") can be seen as the voluntary control people have over their own cognitive processes. There are four basic metacognitive skills identified in the literature: prediction, planning, monitoring, and evaluation (Desoete, 2007a, 2007b; Lucangeli & Cornoldi, 1997). In spelling, test prediction refers to student activities aimed at differentiating which words will require attention and possible further action (such as words with [ei] or [ij]). Planning involves analysing the demands of the spelling exercises, retrieving relevant domain-specific knowledge and skills (e.g., when to use capitals), and sequencing of problem-solving strategies. Monitoring is related to questions such as "am I following my plan?", "should I write a word on another piece of paper to check if the spelling on the test sheet is correct?" and so on. In evaluation there is self-judging of the answer and of the process of getting to this answer.

There are different methods of *assessing metacognition* (Desoete, 2008; Sperling, Howard, Miller, & Murphy, 2002; Tobias & Everson, 2000; Veenman, Van Hout-Wolters, & Afflerbach, 2006). Self-report questionnaires are frequently used to assess metacognitive knowledge and self-ratings are usual measures for metacognitive experiences (Efklides, 2008). The prospective measurement of metacognitive knowledge has to do with metacognitive judgments elicited before problem solving. Retrospective measures of metacognitive knowledge involve self-reports of strategies or metacognitive experiences after problem solving. Several studies underlined the importance of questionnaires and ratings (Busato, Prins, Elshout & Hamaker, 1998; Zimmerman & Martinez-Pons, 1990). However, Veenman et al. (2006) pointed out the limited explained variance towards learning outcomes by self-report questionnaires. Moreover, only moderate correlations were demonstrated between prospective and retrospective measurements of metacognitive knowledge (Veenman, 2003). Hence, in addition to the self-report measures, think-aloud protocols or systematic observation of behaviour can take place to measure metacognitive skills (Veenman & Elshout, 1999). These analyses were found to be very accurate, but time-consuming, techniques to assess metacognitive skills (Pressley, 2000). Recently, *multi-method* techniques are also being used. Often these techniques combine measurements of metacognitive experiences and/or knowledge (e.g., Dermitzaki & Efklides, 2003). For example, students are asked, before and after the processing of a task, to assess the difficulty they experience, the correctness of the solution (conceived or produced), the effort required, and to make subjective estimations about the use of problem-solving strategies. Finally, in calibration studies a comparison is made of whether the prediction before the tasks ("calibration" or comprehension paradigm) or the evaluation after a task ("performance calibration" or postdiction paradigm) corresponds with the actual performance on the task (Glenberg, Sanocki, Epstein, & Morris, 1987; Lin & Zabrucky, 1998; Schraw, Potenza, & Nebelsick-Gullet, 1993). Calibration studies are therefore most closely related to the assessment of metacognitive experiences and refer to the reliability of metacognitive experiences.

To conclude, several problems emerge in the assessment of metacognition (Artzt & Armour-Thomas, 1992). On the one hand, there seem to be various facets of metacognition (metacognitive knowledge, metacognitive experiences, and metacognitive skills) to be assessed with different techniques. On the other hand, from mathematical problem-solving research, we know that how we test influences what we find (Desoete, 2007a). The present study aimed to add some data into the debate on the value of questionnaires and ratings in combination with calibration measures to predict spelling skills during adolescence. Moreover, we aimed to investigate the relationship between spelling performance and spelling-related metacognitive knowledge, metacognitive skills, and metacognitive experiences of college students.

1.3 Spelling and Metacognition

Hacker, Keener, and Kircher (2009) argued that metacognitive monitoring and control are essential components of proficient writing and spelling. Actually, Hacker et al. (2009) defined writing as applied metacognition. In writing, *declarative metacognitive knowledge* can take many forms. First, there is the knowledge that the writer has about himself or herself as a writer, including what knowledge they are comfortable with and which components of spelling they have not yet mastered. In addition, there is metacognitive knowledge regarding the writing task, including strategies specific to a particular writing task. Also, declarative knowledge includes the writer's knowledge about their own affect related to writing, including their self-efficacy for writing in general and specific writing (with students overestimating or underestimating themselves), their motivation to write and how these and other affective factors may influence their writing (Harris et al., 2009). In addition, writing *procedural metacognitive knowledge* includes general and genre-specific strategies the writer is knowledgeable of as well of knowledge of how skills work and when they are needed and the knowledge of one's own optimal writing environment (Harris et al., 2009). Finally, *conditional metacognitive knowledge* includes evaluating the writing task and determining the skills and strategies needed, selecting among alternative strategies, identifying the environmental conditions that can be addressed to make writing conducive, identifying when and why to engage in different components of the writing process and so on (Harris et al., 2009).

Metacognitive experiences and metacognitive knowledge may be involved in what people are aware of when spelling such as awareness of similarly sounding but different diphthongs (‘ou’ or ‘au’ and ‘ij’ or ‘ei’) in spelling. However, Kruger and Dunning (1999) and Kruger (2002) showed that people who are unskilled in, for example, spelling suffer a dual burden. Not only do these people reach erroneous conclusions and make unfortunate choices in their spelling, but their incompetence also robs them of the metacognitive competence to realise it. For example, they found that participants scoring in the bottom quartile on a test of English grammar grossly overestimated their spelling performance and ability. Improving the spelling skills of participants and thus increasing their metacognitive competence helped them recognize the limitations of their ability to produce and recognize written documents that conform to grammar rules and facts. The skills that engender competence to write grammatical English are the very same skills necessary to evaluate competence in that domain. Because of their incompetence, individuals lack the ability to know how well one is performing, when one is likely to be accurate in judgment and when one is likely to be in error. The same skills that enable one to spell without errors are the skills necessary to recognise an error, and these are the same skills that determine if an error has been made. In short, the same knowledge that underlies spelling ability to write without errors is also the knowledge that underlies the ability to make correct estimates about one’s spelling.

1.4 The Present Study

Research comparing different types of measures of older students’ metacognition related to spelling is relatively limited; namely, few studies combine measures on metacognitive knowledge (MK), metacognitive skills (MS), and metacognitive experiences (ME). In the present study we aimed to contribute to the body of knowledge concerning the relationship between the different facets of metacognition and spelling in higher education.

1.4.1 Research Questions: Hypotheses

There were three research questions:

1. At the performance level, what type of spelling errors do college students make? Is there a variability in the errors, that is, do they make basic errors (e.g., <misdrij*f*en> instead of <misdrij*v*en> for “crimes”) as well as rule-related errors (e.g., <*k*erstmis> instead of <Kerstmis> for Christmas) and memory-related errors (e.g., <*copie*> instead of <kopie> for “copy”)? Or do they only make errors in the higher stages of spelling acquisition (only memory-related and non-spelling-related errors)? Following the stage theories, such as that of Ehri (1992), it was hypothesized that no basic errors or rule-related errors would occur but

only memory-related errors or non-spelling-related errors, because the transition from relying on phonological properties of words to recognizing and representing orthographic and morphological regularities and rules has already taken place in their earlier school years (Hypothesis 1).

2. What is the relationship between spelling performance and MK, ME, and MS? It was hypothesized that incompetent spellers will have poor MK and MS, and less accurate ME (Hypothesis 2a). In addition, it was predicted an "above-average effect", or the tendency of the average person to believe he or she is above average (Hypothesis 2b), as found by Kruger and Dunning (1999).
3. Which type of metacognitive measures can most adequately predict proficient spelling? It was hypothesized that measures of MK, MS, and ME would equally well predict spelling, because there is no available evidence to suggest that some facet of metacognition would be more accurate in predicting spelling than the others (Hypothesis 3).

2 Method

2.1 Participants

A total of 2,095 first year bachelor students participated in the study (594 boys and 1,501 girls). At the time of testing their mean age was 18.82 years ($SD = 1.80$). The professional and academic bachelor students were registered in colleges and universities in Ghent, Brussels and Leuven. Several fields of study were selected in order to make the sample representative. These fields were grouped to three major study fields. Specifically, the study field *Education* was represented by the bachelor of primary education and the bachelor of secondary education. The study field *Business and Languages* was represented by the bachelor of business management and the bachelor of translation studies. The study field *Health Care* was represented by the bachelor of audiology, occupational therapy, speech therapy, podiatry, and the bachelor of nursing. Students taking the bachelor programme in *Social Work* were also tested. Participants were informed about the research and consented to participate.

2.2 Instruments

2.2.1 Dictation Test

To measure spelling performance of participants a Dictation test was developed. The instrument met the following three criteria: (a) The instrument should test spelling skills rather than spelling knowledge; that is, to test whether students use rules in practice (during dictation of sentences) so that the test is not limited to word recognition. (b) The instrument should reveal the type of errors students make. Hence, the sentences in the Dictation test contained several phonological, morphological, and lexical target words. The words were of low, medium and high frequency. (c) Finally, the instrument should also address spontaneous writing, that is, use of complex sentences; however spontaneous writing is not included in the data presented in the present chapter. The result was a Dictation test consisting of 12 paragraphs. Each paragraph comprised three coherent sentences.

Classification of Spelling Errors

Performance on the Dictation test was scored by counting the number of spelling errors. Also, the errors were classified in four main categories based on the classification by Kleijnen (1992) and the AT-GSN[1] dictation (Gauderis, Heirman & Vandenhoof, 2004). In this way the spelling errors were both quantitatively and qualitatively analysed. The analyses of spelling errors provided a more differentiated picture of spelling performance. Examples of spelling errors are shown in Table 17.1.

Table 17.1 Examples of spelling errors

	Correct spelling in Dutch	English translation
Basic errors		
een *aanzienelijk aantal	een aanzienlijk aantal	a substantial number
een *effective	een infectie	an infection
een *aambod	een aanbod	an offer
*ecpletici	Epileptici	Epileptics
*misdrijfen	Misdrijven	Crimes
Rule-related errors		
*kerstmis	Kerstmis handycap	Christmas
een *anti-sociale houding	een antisociale houding	antisocial behaviour
een *muziek groepje	een muziekgroepje	a music group, a band
de musici werden *geïnspireert	de musici werden geïnspireerd	the musicians were inspired
*Elke's promoter	Elkes promoter	Elke's supervisor, the supervisor of Elke
Memory-related errors		
een *copie	een kopie	a copy
een *handicap	een handicap	a handicap
Hij wordt door de ziekteverzekering *gewijgerd	Hij wordt door de ziekteverzekering geweigerd	They refused him health insurance.
Non-spelling-related errors		
enige Oost-Vlamingen	enkele Oost-Vlamingen	some East-Flemish people
een spot	een preventiespot	a prevention advertisement

*Underlined letters refer to the mistakes students make.

[1]AT-GSN stands for "Algemene Toets Gevorderde Spelling van het Nederlands" (Ghesquière, 1998).

The first three categories reflected the three strategies spellers use, that is, the phonological, the morphological, and the lexical or mnemonic strategy. Category 1 was labelled "Basic Errors". It included errors in words that could be spelt by the *phonological* strategy. This kind of errors is often made by dyslexics or novice spellers. Category 2 was labelled "Rule-Related Errors" and regarded errors in *morphological* words that could be explained by spelling rules. Verb spelling in Dutch is rule-based, as is the spelling of capitals, of open and closed syllables and the spelling of hyphenated and spaced words. Category 3, called "Memory-Related Errors", involved memory of similar (e.g., <looplank> instead of <loop-plank>) and lexical words. Rules are not sufficient to explain the orthography of this kind of words. In this category three types of errors were included: (a) Errors in loan words (e.g., <computer>, <fitness>); (b) Errors in similarly sounding diphthongs [ei/ij] or [ou/au] (e.g., <lijden> means "to suffer", whereas <leiden> means "to lead"); (c) Errors in adopted words, which in the past quite often had two accepted spellings, a traditional and a progressive one (e.g., <apotheek> and <apoteek>, <chronisch> and <kronisch>, <productie> and <produktie>); since 2007, however, one of them was chosen as the preferred one. Category 4, called "Non-Spelling-Related Errors", involved errors in the Dictation test that are not related to spelling. When a word was added or forgotten it was included in this category. This was also the case when a word was replaced by another word that was meaningful in the context.

The psychometric properties of the classification scheme of the dictation errors were tested on a sample of 2,089 Dutch-speaking students in Flanders (Vrijders et al., 2007). The internal consistency for this test was very satisfactory (Cronbach's $\alpha = 0.89$).

2.2.2 Metacognition Questionnaires

Two questionnaires were created for the present study, namely a prospective and a retrospective metacognition questionnaire.

The Prospective Metacognition Questionnaire

The PMQ assessed students' MK of the self as speller and students' use of MS in spelling, namely checking of spelling.

The MK of the self as speller was measured as follows. Participants were required to rate their own *spelling skills*, as compared to peers, on a 7-point scale ranging from 1 (very bad) to 7 (very good).

They were also required to report the kind of *spelling difficulties* they had by selecting one of the spelling categories, such as verb spelling, English verbs, use of apostrophe and dieresis, use of capital letters, memory-related words (e.g., [c/k] or [ij/ei]), and writing words with/without hyphenation (e.g., "semi-" or "semi…").

Their responses were on a 3-point scale ranging from 1 (many difficulties) to 3 (not many difficulties).

The use of MS was assessed with one item by asking participants how often they read through their own texts, letters, and e-mails to check for any spelling errors. Responses were on a 5-point rating scale, varying from 1 (never) to 5 (always).

The PMQ was tested in previous studies in order to determine its reliability. Test-retest correlation of 0.81 ($p < 0.01$) was found.

The Retrospective Metacognition Questionnaire

The RMQ assessed metacognitive experiences, namely feeling of confidence (FOC; metacognitive feeling) and estimate of the number of spelling errors (EOSE; metacognitive judgment). Also, a score showing the correspondence between the ratings of FOC and actual performance was calculated as well as a calibration index using the EOSE.

To assess the *feeling of confidence* (FOC) participants were asked to look at ten words of the Dictation test. They were asked to rate how sure they were for the spelling of each word on a 4-point rating scale, ranging from 1 (I am absolutely sure it is incorrect) to 4 (I am absolutely sure it is correct).

Participants might be sure that their spelling was correct whereas they had spelled the word incorrectly or vice versa. To assess the *correspondence between FOC and actual spelling performance* the ratings of FOC that fully corresponded to the actual spelling performance (e.g., the response "I am absolutely sure I wrote the word correctly" and correct answer and the response "I am absolutely sure I did not write the word correctly" and incorrect answer) received 2 points; the response "I am sure I wrote (did not write) the word correctly" and corresponding spelling performance received 1 point, while the response "I am absolutely sure I wrote (did not write) the word correctly" and not corresponding spelling performance received a 0 point. Cronbach's alpha for the scores was 0.87.

To assess the *estimate of the number of spelling errors* (EOSE), participants were asked to estimate the number of errors they had made (e.g., six errors) in three randomly selected paragraphs of the Dictation test (paragraphs 10, 11, and 12).

To assess the students' *calibration index* between the actual performance score and the estimated score of their spelling performance (e.g., "If I lose 0.5 point for each error, I think I will score 7/10 on this paragraph for the six errors I have made") the score participants attributed to their performance (e.g., 7 out of 10) was subtracted from their actual performance score (e.g., 8 out of 10 for four errors they made).

The PMQ and RMQ were tested in a pilot study in order to determine their reliability for measuring individual differences in spelling and metacognition. Gutmann's split-half and Spearman-Brown's coefficients were 0.70 and 0.72, respectively. Furthermore, all variables were normally distributed and test-retest correlations of 0.85 ($p < 0.001$) were found.

2.3 Procedure

Participants took the Dictation test during the first semester of the academic year. The test was dictated in the following way. First, a paragraph was read aloud twice. Then students had to write down on a sheet of paper the paragraph that was dictated in sentence parts. After dictating all 12 paragraphs the complete dictation test was read aloud once more to give the students the opportunity to check for mistakes. The PMQ was completed before the Dictation test. The RMQ was completed after the Dictation test. All sessions were carried out collectively in classrooms, after assuring good testing conditions.

3 Results

3.1 Spelling Performance

On all paragraphs of the Dictation test students made an average of 24 spelling errors ($SD = 13$) in 410 words. Concerning the Basic Errors category, 637 students (30.5%) made no errors at all, 571 (27%) made one error and 348 (16.7%) made two errors ($M = 1.95$, $SD = 3.25$), that is, there was a downward trend with the increase in the number of errors. This trend did not occur with the Rule-Related Errors category. In this case there was a normal distribution in relation to the number of errors ($M = 15.74$, $SD = 6.79$), that is, the number of students rose in direct proportion to the number of errors until a peak was reached with 150 students (7.2%) who made 13 errors; after that peak, there was a decrease in the number of students who made such errors. In the case of the Memory-Related Errors category, the errors were less than in the case of Rule-Related Errors category ($M = 1.63$, $SD = 1.70$). Finally, concerning the Non-Spelling-Related Errors category, students made a relatively large number of non-spelling-related errors ($M = 4.82$, $SD = 5.47$).

Our main focus, however, was on the spelling of relatively "incompetent" participants, which we defined, in line with Kruger and Dunning (1999), as those whose test score fell in the bottom quartile ($n = 520$); their mean errors were 41.29 ($SD = 13.59$), whereas college students in the 3rd quartile made 25.11 errors ($SD = 2.78$), students in the 2nd quartile made 18.65 ($SD = 2.02$) errors, and students in the top quartile made 12.01 ($SD = 2.94$). It is worth noting that incompetent spellers made all kinds of errors, but mainly rule-related errors. Specifically, they made a mean number of 4.61 basic errors ($SD = 5.37$), 23.83 rule-related errors ($SD = 6.63$), 2.46 memory-related errors ($SD = 2.78$), and 10.14 non-spelling-related errors ($SD = 8.16$).

The very competent spellers (in the top quartile) also made mainly rule-related errors. Specifically, they made a mean number of 0.43 basic error ($SD = 0.65$), 9.00 rule-related errors ($SD = 2.57$), 1.00 memory-related error ($SD = 0.76$), and 1.58 non-spelling-related errors ($SD = 1.32$).

Table 17.2 Means (and *SD*) of spelling errors in the quartiles as a function of error category along with the respective *F* values

	Bottom quartile	3rd Quartile	2nd Quartile	Top quartile
Category 1	4.62[d] (537)	1.76[c] (1.38)	1.01[b] (0.97)	0.43[a] (0.65)
Category 2	23.83[d] (6.63)	16.84[c] (3.25)	13.33[b] (2.39)	9.00[a] (2.57)
Category 3	2.46[d] (2.78)	1.66[b] (1.57)	1.41[b] (0.92)	1.00[a] (0.76)
Category 4	10.13[d] (8.16)	4.73[c] (2.47)	2.87[b] (1.79)	1.58[a] (1.32)

Note: Number sharing the same index (a, b, c, d) did not significantly differ between them

The MANOVA with group (bottom quartile, 3rd quartile, 2nd quartile, top quartile) as independent variable and the four types of spelling errors as dependent variable was significant, Wilks's lambda$=$0.27, $F(12, 5511.39)=292.47$, $p<0.001$, partial $\eta^2=35$. There were differences between groups for basic errors, $F(3, 2086)=225.19$, $p<0.001$, partial $\eta^2=0.25$, for rule-related errors, $F(3, 2086)=1231.28$, $p<0.001$, partial $\eta^2=0.64$, for memory-related errors, $F(3, 2086)=75.61$, $p<0.001$, partial $\eta^2=0.09$, and for non-spelling-related errors, $F(3, 2086)=381.66$, $p<0.001$, partial $\eta^2=0.35$. For a summary of the mean number of errors (*M*) and the *SD* per error category, see Table 17.2.

3.2 Metacognitive Knowledge and Metacognitive Skills

The Prospective Metacognition Questionnaire (PMQ) was used to assess student's MK of the self as speller and student's use of MS in spelling, namely checking of spelling.

Overall, in our sample students rated themselves as above medium spellers ($M=4.30$, $SD=0.95$). Approximately 40.3% of the students in the sample considered themselves almost as good as their peers as far as their spelling skills were concerned (score 4), whereas 34.4% thought they were slightly better compared to their peers (score 5). In addition 6.9% believed that they were better spellers than their peers (score 6), and 0.6% thought they were much better than their peers (score 7). Only 12.2% of the students rated themselves as doing rather worse than their peers (score 3), 3.4% rated themselves as worse than their peers (score 2), and 0.6% admitted performing much worse than their peers when it came to spelling (score 1).

The PMQ also included a rating of the difficulties students had with spelling. Students reported difficulties with verb spelling ($M=2.30$, $SD=0.72$), English verbs ($M=2.09$, $SD=1.28$), the use of apostrophe and dieresis ($M=2.18$, $SD=0.63$), the use of capital letters ($M=2.64$, $SD=0.55$), memory-related words ($M=2.53$, $SD=0.64$), and writing words with/without hyphenation ($M=1.95$, $SD=0.59$). These are all rule- and memory-related errors, and this finding suggests that the difficulties reported correspond to the kind of errors most often made in the Dictation test.

In addition the ANOVA with group (bottom quartile, 3rd quartile, 2nd quartile, top quartile) as independent variable and MK of the self as speller as dependent variable was significant, $F(3, 2074) = 130.19$, $p < 0.001$, partial $\eta^2 = 0.16$. Post hoc analyses revealed that all groups significantly differed from each other. Participants in the bottom quartile rated themselves as less competent ($M = 3.76$, $SD = 1.02$) than students in the 3rd quartile ($M = 4.18$, $SD = 0.86$) and students in the 2nd ($M = 4.46$, $SD = 0.81$), or top quartile ($M = 4.79$, $SD = 0.79$).

The PMQ also included an assessment of MS. Participants had to rate how often they read through their own tests, letters, and e-mails to check for spelling errors. The mean number of checking for spelling errors was 2.85 ($SD = 0.96$). Approximately 42.6% of students in the sample stated that they usually checked the material they were writing themselves, while 22.3% claimed that they always checked it. Finally, 2.1% of the students admitted that they never and 9.7% that they very seldom checked their spelling. The other 23.3% of the students rated that they sometimes checked the material they were writing themselves.

The ANOVA with group (bottom quartile, 3rd quartile, 2nd quartile, top quartile) as independent variable and use of MS as dependent variable was significant, $F(3, 2067) = 25.36$, $p < 0.001$, partial $\eta^2 = 0.04$. Post hoc analyses revealed that students in the bottom quartile checked their texts less ($M = 2.51$, $SD = 0.99$) than peers in the 2nd quartile ($M = 2.77$, $SD = 0.95$) and peers in the top quartile ($M = 3.01$, SD = 0.89). Students in the 3rd quartile ($M = 2.63$, $SD = 1.01$) differed from peers in the top quartile. Students in the 2nd quartile differed from students in the bottom and top quartile.

3.3 Metacognitive Experiences

In response to a ten-word list the students were asked to report retrospectively, after the Dictation test, their FOC; correspondence of FOC with actual spelling performance was further investigated. Also, based on the three paragraphs (i.e., paragraphs 10, 11, and 12) of the Dictation test students were asked to report their EOSE; a calibration index between actual and estimated performance scores was also calculated.

3.3.1 Feeling of Confidence

There were three words that were written incorrectly by a high number of students; specifically, <gecanceld>, <lijdt>, and <antisociale> (see Table 17.3). These words were most frequently misspelled without the students realising it. A total score of FOC was firstly computed for all ten words together for each student. The mean total FOC score for the whole sample was $M = 20.84$ ($SD = 3.64$). The ANOVA with the sum score as dependent variable and the group (bottom quartile, 3rd quartile, 2nd quartile and top quartile) as independent variable was significant, $F(3, 2036) = 122.18$, $p < 0.001$, partial $\eta^2 = 0.15$. Post hoc analyses revealed that students in the bottom quartile had a significantly lower FOC ($M = 18.95$. $SD = 3.46$) compared to students in the 3rd quartile ($M = 20.13$, $SD = 3.48$), or to students in the 2nd quartile

Table 17.3 Descriptives of the various measures of feeling of confidence (FOC) as a function of the ten words

	FOC *M* (*SD*)	Correct spelling (%)	++ (%)	+ (%)	– (%)	–– (%)	0 Point (%)	1 Point (%)	2 Points (%)
Firma's	1.72 (0.71)	93.2	42.5	43.3	13.4	0.7	33.3	45.6	21.1
Vind	1.45 (0.69)	91.7	65.1	25.4	8.2	1.3	12.0	24.6	63.5
Lijdt	2.19 (0.94)	12.5	26.5	39.0	24.1	10.4	62.4	25.0	12.6
Ondervraagd	1.73 (0.76)	84.2	44.4	39.6	14.1	1.8	18.8	38.2	43.0
Georganiseerd	1.61 (0.67)	94.9	48.9	41.4	9.0	0.7	10.9	41.1	49.0
Geleide	2.10 (0.78)	73.1	23.2	45.8	28.5	2.6	33.3	45.6	21.1
Gecanceld	2.72 (0.75)	29.4	13.7	48.3	33.5	4.5	56.2	37.7	6.1
Antisociale	2.05 (0.70)	27.2	21.4	53.5	24.1	1.1	63.8	30.5	5.8
Hondenweer	2.79 (0.70)	73.1	14.4	51.3	32.4	1.8	40.6	46.4	12.9
Oost-Vlaanderen	3.20 (0.73)	86.6	37.4	45.2	16.5	0.9	18.6	44.5	36.9

Note: ++ absolutely sure, + sure, – not sure, –– absolutely not sure, *0 Point* no correspondence between FOC and actual spelling performance, *1 Point* partial correspondence between FOC and actual spelling performance, *2 Points* correspondence between FOC and actual spelling performance

($M = 21.57$, $SD = 3.40$), and to high proficient spellers in the top quartile ($M = 22.72$, $SD = 3.05$). Students in the 3rd quartile had lower FOC compared to students in the 2nd or top quartile, while students in the top quartile were more confident than all other students.

3.3.2 Correspondence of Feeling of Confidence with Actual Spelling Performance

There was a significant correlation between FOC and the number of spelling errors, $r = -0.38$, $p < 0.001$. The correspondence was also significant for FOC and basic spelling errors, $r = -0.20$, $p < 0.001$, FOC and rule-related errors, $r = -0.37$, $p < 0.001$, FOC and memory-related errors, $r = -0.13$, $p < 0.001$, and for FOC and non-spelling-related errors, $r = -0.27$, $p < 0.001$.

3.3.3 Estimate of Number of Spelling Errors

Concerning the EOSE in the three paragraphs of the Dictation test, it was observed that the students usually gave a higher estimate of errors than they actually had made in the three paragraphs. Over 60% of the students thought that in each paragraph, they were making two or fewer errors while, on average, they made one error in paragraph 10 and 11, and two errors in paragraph 12. For the overall results see Table 17.4.

To look for differences in EOSE between the groups of students as regards their spelling performance (quartiles), a MANOVA was conducted with the EOSE scores in the three paragraphs as dependent variables and group (bottom quartile, 3rd quartile, 2nd quartile, top quartile) as independent variable. The multivariate effect was significant, Wilks's lambda $= 0.85$, $F(9, 5042.85) = 37.69$, $p < 0.001$, partial $\eta^2 = 0.05$. Students in the bottom quartile estimated that they made more errors compared to students in the other quartiles on paragraph 10, $F(3, 2074) = 55.07$, $p < 0.001$, partial $\eta^2 = 0.07$, on paragraph 11, $F(3, 2074) = 108.19$, $p < 0.001$, partial $\eta^2 = 0.14$, and on paragraph 12, $F(3, 2074) = 61.77$, $p < 0.001$, partial $\eta^2 = 0.08$. Specifically, the students in the bottom quartile estimated that they made more mistakes in paragraph 10 ($M = 1.87$, $SD = 1.59$), in paragraph 11 ($M = 2.92$, $SD = 1.65$), and in paragraph 12 ($M = 3.03$, $SD = 2.17$) than students in the top quartile, whereas students in the top quartile estimated that they had made few

Table 17.4 Means (and *SD*) of the estimate of number of spelling errors (EOSE) and of calibration

	EOSE	Estimated score	Actual score	Calibration
Paragraph 10	1.35 (1.26)	9.33 (0.63)	9.51 (0.41)	0.18 (0.65)
Paragraph 11	2.12 (1.54)	8.94 (0.77)	9.31 (0.63)	0.37 (0.83)
Paragraph 12	2.25 (1.71)	8.88 (0.85)	9.01 (0.60)	0.13 (0.92)

Note: Range from 1 to 10.

mistakes in paragraph 10 ($M = 0.94$, $SD = 1.01$), in paragraph 11 ($M = 1.40$, $SD = 1.19$), and in paragraph 12 ($M = 1.75$, $SD = 1.75$).

3.3.4 Calibration Index

To calculate the calibration index and to see if the calibration discrepancy was larger in spellers within the bottom quartile (Kruger & Dunning, 1999), we took the difference between the actual performance score and the performance score estimated by the student (see Fig. 17.1) for each of the three paragraphs of the Dictation test.

For paragraph 10, 38.40% of the students had a calibration index of 0, that is, perfect calibration, whereas for paragraph 11 and for paragraph 12, 27% and 25.10%, respectively had perfect calibration. To compare proficient spellers with below average spellers, a MANOVA was conducted on the calibration indices. The MANOVA with the calibration indices in the three paragraphs as dependent variables and group (bottom quartile, 3rd quartile, 2nd quartile, top quartile) as independent variable showed a significant multivariate effect, Wilks's lambda = 0.98, $F(9, 5042.85) = 4.04$, $p < 0.001$, partial $\eta^2 = 0.01$. However, students in the bottom quartile did not differ significantly in calibration from the other groups on paragraph 10, $F(3, 2074) = 0.39$, *ns*, or on paragraph 12, $F(3, 2074) = 1.96$, *ns*. They only differed significantly on paragraph 11, $F(3, 2074) = 8.21$, $p < 0.001$, partial $\eta^2 = 0.01$. Specifically, the students in the bottom quartile were better calibrated in paragraph 11 ($M = 0.22$, $SD = 1.00$) than the other students who tended to underestimate their spelling performance even more. They differed from students in the 3rd quartile ($M = 0.42$, $SD = 0.89$) and students in the 2nd quartile ($M = 0.41$, $SD = 0.73$) and from students in the top quartile ($M = 0.45$, $SD = 0.61$). Students in the bottom quartile estimated 2.92 errors ($SD = 1.65$). Thus, their estimated spelling score was $10 - 2.92/2 = 8.54$, whereas their actual spelling score was 8.76 out of 10 (SD = 0.73). Post hoc analyses revealed that students in the 3rd quartile (actual score 9.24, $SD = 0.51$; estimated score 8.82, $SD = 0.79$), 2nd quartile (actual score 9.49, $SD = 0.42$; estimated score 9.09, $SD = 0.64$), or top quartile (actual score 9.75, $SD = 0.32$; estimated score 9.30/10, $SD = 0.60$) did not differ from each other but they did differ from students in the bottom quartile. These data reveal that incompetent spellers underestimated their spelling skills less compared to peers with better spelling skills.

3.3.5 Relations Between MK, MS, and ME

To investigate the relations between MK, MS and ME, Pearson correlations were computed on the respective scores (see Table 17.5). Table 17.5 also shows the correlations between the facets of metacognition and the actual spelling performance.

Most metacognitive measures were significantly intercorrelated. Low, but significant, and positive correlations between MK of the self as speller and FOC ratings were found. Moreover, there was a low, but significant, and positive correlation

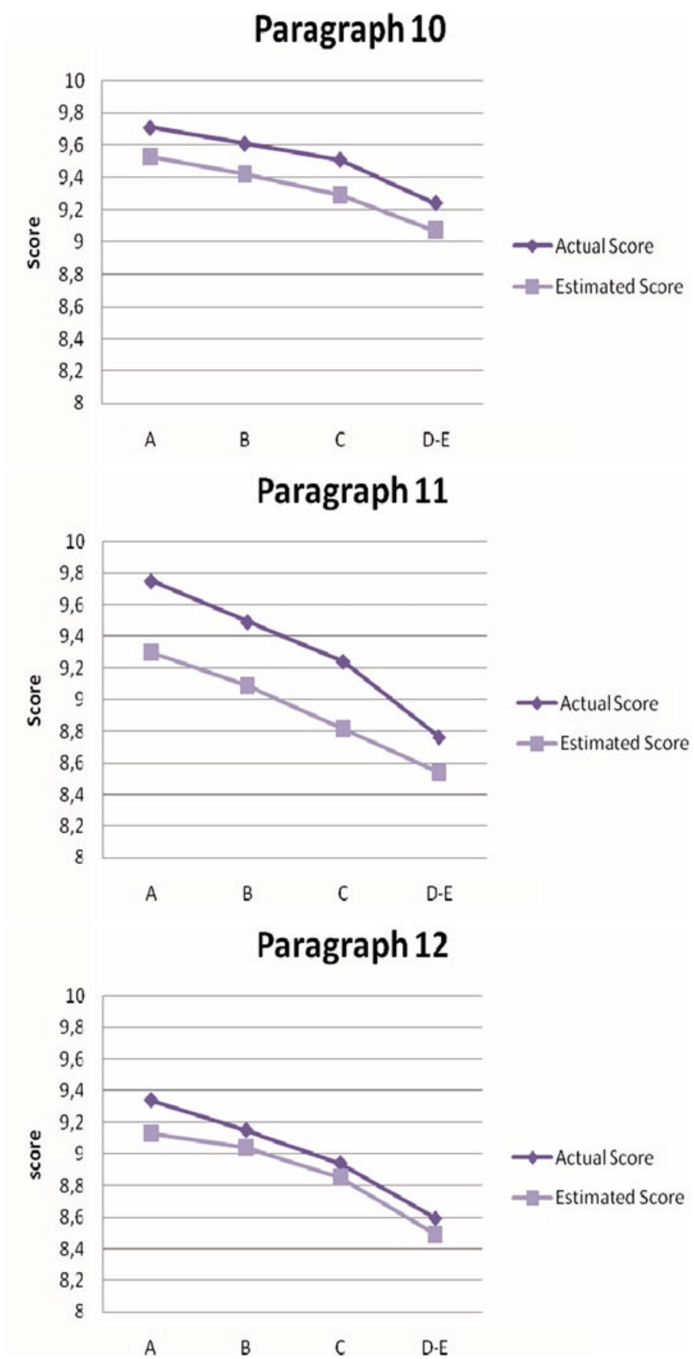


Fig. 17.1 Estimated performance score versus actual performance score in the three paragraphs of the Dictation test

Table 17.5 Relations between metacognitive knowledge (MK), metacognitive skills (MS), metacognitive experiences (ME) measures, and actual performance

	MK	MS	FOC total	Calibration paragraph 10	Calibration paragraph 11	Calibration paragraph 12	PERF
Prospective metacognition questionnaire							
MK of the self as speller	–	0.197*	0.449*	–0.195*	–0.167*	–0.189*	–0.389*
MS (checking for errors)	–	–	0.174*	–0.043*	–0.029	–0.013	–0.148*
Retrospective metacognition questionnaire							
FOC total (for the 10 words)	0.449*	0.174*	–	–0.262*	–0.230*	–0.209*	–0.373*
Correspondence of FOC with actual performance	–0.437*	–0.129*	–0.491*	0.642*	0.595*	0.644*	0.394*
EOSE for paragraph 10	–0.326*	–0.102*	–0.412*	0.798*			0.326*
EOSE for paragraph 11	–0.413*	–0.135*	–0.479*		0.690*		0.375*
EOSE for paragraph 12	–0.382*	–0.095*	–0.376*	.		0.776*	0.318*

Note: *PMQ* prospective metacognition questionnaire, *RMQ* retrospective metacognition questionnaire, *FOC* feeling of confidence, *EOSE* estimate of the number of spelling errors, *PERF* actual performance

*$p < 0.001$

between the MS (i.e., checking for spelling errors) rating assessed prospectively and the FOC rating assessed retrospectively. There was also a high and significant positive correlation between the EOSE rating and the calibration index, which is understandable since the calibration index includes the EOSE. Moreover, there were moderate and negative correlations of MK, MS, and FOC with EOSE in the three paragraphs. The correlations between MK of the self as speller and MS with the calibration index were negative. As regards actual performance, the number of errors actually made were negatively correlated with MK of the self as speller and MS, although the latter correlation was low, and with FOC. The relation with EOSE was positive and moderate.

3.4 Can Metacognition Predict Proficient Spelling?

3.4.1 Can Prospective Metacognitive Measures Predict Spelling Performance?

A regression analysis was performed on spelling performance as dependent variable with MK of the self as speller and MS entered simultaneously as predictor variables. The MK of the self as speller and MS predicted 16% of the variance of spelling performance, and MK was a stronger predictor, $\beta = -0.375$, $t = -18.229$, $p < 0.001$, than MS, $\beta = -0.073$, $t = -3.531$, $p < 0.001$. The negative sign suggests that the higher the MK and MS, the less the errors made.

3.4.2 Can Retrospective Metacognitive Measures Predict Spelling Performance?

A regression analysis was performed on spelling performance as dependent variable with the retrospectively assessed word-specific FOC scores entered simultaneously as predictor variables (see Table 17.6). This treatment was dictated by the fact that the various words represented different categories of spelling errors and word-specific FOC was assumed to represent a more accurate predictor than an undifferentiated overall FOC score. The FOC ratings predicted 23.8% of the variance of spelling performance. Of the various predictors, FOC ratings on the words <Oost-Vlamingen>, <hondenweer>, <geleide>, <ondervraagd>, <vind>, and <georganiseerd> were significant. These words are all words that belong to the Rule-Related Errors category.

In addition, a regression analysis was conducted on spelling performance as dependent variable with the EOSE scores in the three paragraphs entered simultaneously as predictor variables. The R^2 was 0.157 and $F(3, 2074) = 130.16$, $p < 0.001$. All three predictors were significant. Specifically, for paragraph 10, $\beta = 0.110$, $t = 3.938$, $p < 0.001$; for paragraph 11, $\beta = 0.232$, $t = 7.874$, $p < 0.001$; and for paragraph 12, $\beta = 0.112$, $t = 4.228$, $p < 0.001$.

Table 17.6 Prediction of spelling performance based on feeling of confidence (FOC)

FOC on	β	*t*	*p*
Oost-Vlaanderen	−0.132	−6.611	<0.001
Gecanceld	0.035	1.803	0.072
Hondenweer	0.056	2.889	0.004
Lijdt	0.012	0.596	0.551
Antisociale	−0.050	−2.593	0.010
Geleide	−0.120	−5.885	<0.001
Firma’s	0.013	0.657	0.511
Vind	−0.144	−6.849	<0.001
Ondervraagd	−0.211	−9.281	<0.001
Georganiseerd	−0.163	−7.281	<0.001

$R^2 = 0.21$, $F(10, 2025) = 64.72$, $p < 0.001$

Finally, a regression analysis was conducted on spelling performance as dependent variable with the three calibration indices in the three paragraphs as predictor variables. The R^2 was 0.028 and $F(3, 2074) = 20.71$, $p < 0.001$. However, only the calibration index for paragraph 10, $\beta = 0.116$, $t = 4.604$, $p < 0.001$, and for paragraph 11, $\beta = -0.178$, $t = -6.971$, $p < 0.001$, were significant predictors of spelling performance. What is worth noting is that the calibration index for paragraph 10 positively predicted spelling performance, whereas for paragraph 11 negatively. This reflects the more accurate calibration that was detected in paragraph 10 and the less accurate in paragraph 11. In the latter case students tended to underestimate their performance. In paragraph 12 there was a very accurate calibration which probably did not leave score variability to sufficiently predict performance.

4 Discussion

Following the stage theories, such as that of Ehri (1992), it was hypothesized that no basic errors or rule-related errors would occur in college students but that only memory-related errors or non-spelling-related errors, because the transition from relying on phonological properties of words to recognizing and representing orthographic and morphological regularities and rules has already taken place in their earlier school years (Hypothesis 1). The findings of the present study do not confirm the proposed stage hypothesis. Since both weak and proficient spellers made several types of errors, and since they made especially rule-related errors a stage paradigm is not tenable.

Moreover, the present study revealed that quite a large number of college students made spelling errors. Three words were misspelled with striking frequency, namely <gecanceld>, <lijdt>, and <antisociale>. Half the students were unable to assess themselves correctly (correspondence score) when it came to the spelling of these words. For the spelling of the word <hondenweer>, the correspondence between FOC rating and actual performance appeared completely wrong in two-fifths of the cases. Writing of words like <firma’s>, <vind>, and <georganiseerd> was estimated as “definitely correct” by two-fifths, two-thirds and two-fifths of cases, respectively, although these words were spelled correctly by more then four out of five students. A potential explanation for the good spelling performance with these words is that they are frequently recurring words in the Dutch language. The two past participles are regular weak verbs which have a clear conjugation rule, namely the “t’kofschip” rule (mnemonic for voiceless consonants of Dutch; [ge] + stem + [d], [ge] + stem + [t] when the stem ends in a consonant contained in the mnemonic “t’kofschip”). The conjugated verb <vind> drops the final [t] because of the [je] after the finite form. This is a rule that is already taught in primary school. This also applies to the plural of nouns that end in a consonant preceded by one grapheme, such as <firma’s>.

As to Hypothesis 2a (incompetent spellers have poor MK and MS and less accurate ME), in line with Hacker et al. (2009) the present study revealed that students who spell well and, therefore, make few errors (i.e., in the top quartile) also appear to perceive themselves as competent spellers, that is, their MK of the self as speller represents their competence; they also assess themselves as using more often MS and have higher FOC after the Dictation test than students of the other quartiles. Concerning spelling performance, participants in the bottom quartile rated themselves as less competent spellers compared to students in the other quartiles. Therefore, in line with Harris et al. (2009) and Zimmerman and Reisemberg (1997), spelling performance was related to MK of the self as speller and use of MS. Moreover, in line with Efklides (2002), ME, such as FOC and EOSE, were also found to be related to spelling performance. Students in the bottom quartile rated themselves lower compared to all other groups and they had a lower feeling of confidence than students in the other quartiles.

To investigate (Hypothesis 2b) if the calibration discrepancy was bigger in spellers within the bottom quartile (Kruger & Dunning, 1999) compared to spellers in the top quartile, the difference between the actual score on spelling and the spelling score estimated by the students was computed. Students differed significantly only on paragraph 11 of the Dictation test. However, in contrast with the Kruger and Dunning (1999) data, the students in the bottom quartile in this study did not overestimate themselves more than proficient spellers. Moreover, there was only a very weak prediction of the variance in spelling performance by calibration indices. It can be concluded that, in line with Desoete (2008) and Desoete and Roeyers (2006), the way in which calibration is assessed and, especially, the facet of metacognition that is involved in the computation of the calibration index (FOC vs. EOSE), are important. Calibration might be a time-saving assessment technique. It is not, however, a good way to predict spelling performance in college students.

Moreover, Hypothesis 3 stated that measures of MK, MS, and ME would equally well predict spelling. The present study revealed that all three facets of metacognition were significantly correlated with performance. Specifically, ratings of MK of the self as speller and use of MS predicted about one sixth of the variance of spelling performance. Also FOC ratings predicted about one fourth of the same variance. Finally, EOSE predicted about one sixth of the same variance and in all three paragraphs of the Dictation test EOSE predicted proficient spelling. However, calibration indices only predicted about one thirtieth of the spelling performance. Perhaps, for FOC ratings the choice of the words on which FOC was reported was important. Higher FOC rating on <Oost-Vlamingen>, <geleide>, <ondervraagd>, and <georganiseerd>, in particular, predicted making few spelling errors, whereas the opposite was true for the FOC rating on <hondenweer> that positively predicted spelling performance.

The present study had a number of limitations. Since we opted for a large group of students, we could not incorporate other kinds of measures of metacognition such as think-aloud protocols or online recording (see also Veenman, 2003) into our study. Follow-up research using those techniques to assess metacognitive skills is certainly to be recommended. Furthermore, we studied only a limited number of aspects of the three facets of metacognition, namely MK of the self as speller, use of MS such as checking for errors, FOC, and EOSE, because these aspects have been shown by clinical experience to be frequently disturbed in poor spellers. Of course, follow-up research is necessary into other MS (such as prediction, planning and monitoring skills) and into other aspects of the broader metacognition related to spelling. It was certainly not the intention to deny the importance of these aspects, but merely to make a start with research into an instrument that could assess (screen) the metacognition of spellers, in order to be able subsequently to research into those who underperform in terms of spelling and/or metacognition. In addition, the difference between our data and the data of Kruger and Dunning (1999) might be caused by our calibration instrument and by the rather limited number of errors that were taken into account for the calibration measure in this study. Additional research is needed to investigate whether another instrument might lead to other conclusions.

Nevertheless, based on these studies, we can conclude that metacognitive knowledge, skills and experiences are successful in predicting part of the variance in spelling performance. Certainly in the case of students with problems, it may be advisable to examine these metacognitive facets. There is evidence that metacognitive knowledge and skills can give valuable information on the spelling skills of college students. We suggest that researchers who are interested in students' skills should use multiple-method designs, including ratings, questionnaires and think aloud protocols.

References

Afflerbach, P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly, 25*, 31–46.

Artzt, A. F., & Armour-Thomas, E. (1992). Development of a cognitive-metacognitive framework for protocol analysis of mathematical problem solving in small groups. *Cognition and Instruction, 9*, 137–175.

Brown, A. L. (1980). Metacognitive development and reading. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 453–481). Hillsdale, NJ: Erlbaum.

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. Reiner & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65–116). Hillsdale, NJ: Erlbaum.

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (1998). Learning styles: A cross-sectional and longitudinal study in higher education. *British Journal of Educational Psychology, 68*, 427–441.

Carlisle, J. F. (1994). Morphological awareness, spelling, and story writing: Possible relationships for elementary-age children with and without learning disabilities. In N. C. Jordan & J. Goldsmith-Philips (Eds.), *Learning disabilities: New directions for assessment and intervention* (pp. 123–145). Boston: Allyn & Bacon.

Claes, B., & Moeyaert, S. (2003). Opstelling en aanzet tot normering van een instaptoets spelling voor beginnende logopediestudenten [Setting up of and encouragement to standardise an entry test in spelling for new logopaedics students]. *Tijdschrift voor Logopedie & Audiologie, 33*(4), 157–166.

Defior, S., Jimenez-Fernandez, G., & Serrano, F. (2009). Complexity and lexical effects on the acquisition of Spanish spelling. *Learning and Instruction, 19*, 55–65.

Dermitzaki, I., & Efklides, A. (2003). Goal orientations and their effect on self-concept and metacognition in adolescence. *Psychology: The Journal of the Hellenic Psychological Society, 10*, 214–227.

Desoete, A. (2007a). Evaluating and improving the mathematics teaching-learning process through metacognition. *Electronic Journal of Research in Educational Psychology, 5*(3), 705–730.

Desoete, A. (2007b). Students with mathematical disabilities in Belgium: From definition, classification and assessment to STICORDI device. In T. E. Scruggs & M. A. Mastropieri (Eds.), *Advances in learning and behavioural disabilities: Vol. 20. International perspectives* (pp. 181–222). Oxford, England: Elsevier.

Desoete, A. (2008). Multi-method assessment of metacognitive skills in elementary school children: How you test is what you get. *Metacognition and Learning, 3*, 189–206.

Desoete, A., & Roeyers, H. (2006). Metacognitive macroevaluations in mathematical problem solving. *Learning and Instruction, 16*, 12–25.

Desoete, A., & Veenman, M. (Eds.). (2006). *Metacognition in mathematics education.* New York: NOVA.

Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation, and self-regulation. In A. Efklides, J. Kuhl, & R. M. Sorrentino (Eds.), *Trends and prospects in motivation research* (pp. 297–323). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2002). The systemic nature of metacognitive experiences: Feelings, judgments, and their interrelations. In M. Izaute, P. Chambres, & P.-J. Marescaux (Eds.), *Metacognition: Process, function, and use* (pp. 19–34). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*, 277–287.

Efklides, A., & Sideridis, G. D. (2009). Assessing cognitive failures. *European Journal of Psychological Assessment, 25*, 69–72.

Ehri, L. C. (1992). Review and commentary: Stages of spelling development. In S. Templeton, & D. R. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy: A memorial festschrift for Edmund H. Henderson* (pp. 307–332). Hillsdale, NJ: Erlbaum.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–235). Hillsdale, NJ: Erlbaum.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34*, 906–911.

Gauderis, K., Heirman, E., & Vandenhoof, T. (2004). *Normering van AT-GSN, TTR en CDR-5 bij leerlingen van het zesde jaar Algemeen Secundair Onderwijs* [Standardisation of AT-GSN, TTR and CDR-5 among sixth-year students in General Secondary Education]. Unpublished doctoral thesis, University College Arteveldehogeschool, Ghent, Belgium.

Gentry, J. R. (1982). An analysis of developmental spelling in GNYS AT WRK. *The Reading Teacher, 36*, 192–200.

Ghesquière, P. (1998). *Algemene toets gevorderde spelling van het Nederlands (AT-GSN), verantwoording en handleiding. Rapport van een specialisatiejaar: onderzoek AT-GSN-dictee* [General Test in Dutch Advanced Spelling (AT-GSN), justification and manual. Report of a specialisation year: study on AT-GSN dictation]. Unpublished manuscript, University of Leuven, Belgium.

Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General, 116*(2), 119–136.

Grezel, J. E. (2007). Die spelfouten maken toch niets uit? Zorgen over taalvaardigheid van studenten [So those spelling errors don't matter? Concerns about students' language skills]. *Onze Taal, 76*, 148–151.

Hacker, D. J., Keener, M. C., & Kircher, J. C. (2009). Writing is applied metacognition. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 154–172). New York: Routledge.

Harris, K. R., Graham, S., Brindle, M., & Sandmel, K. (2009). Metacognition and children's writing. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 131–153). New York: Routledge.

Hayes, J. (1996). A new framework for understanding cognition and affect in writing. In M. Levy & S. Ransdell (Eds.), *The science of writing: theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Erlbaum.

Henderson, E. H. (1992). The interface of lexical competence and knowledge of written words. In S. Templeton & D. R. Bears (Eds.), *Development of orthographic knowledge and the foundations of literacy: A memorial festschrift for Edmund H. Henderson* (pp. 1–30). Hillsdale, NJ: Erlbaum.

Henderson, E. H., & Beers, J. W. (1980). *Developmental and cognitive aspects of learning to spell: A reflection of word knowledge.* Newark, DE: International Reading Association.

Herbots, K. (2005, October, 27). Een op de drie studenten kan niet zonder fouten schrijven [One in every three students cannot write without mistakes]. *De Morgen.* Retrieved July 6, 2009, from the World Wide Web: http://www.standaard.be/Artikel/Detail.aspx?artikelId=dmf27102005_002

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist, 22* (3&4), 255–278.

Keuning, J., & Verhoeven, L. (2008). Spelling development throughout the elementary grades: The Dutch case. *Learning and Individual Differences, 18*, 459–470.

Kleijnen, R. (1992). *Hardnekkige spellingfouten. Een taalkundige analyse* [Persistent spelling mistakes. A linguistic analysis]. Lisse, The Netherlands: Swets & Zeitlinger.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289–325). Cambridge, UK: Cambridge University Press.

Kruger, J. (2002). Unskilled and unaware – but why? A reply to Krueger and Mueller. *Journal of Personality and Social psychology, 82*, 189–192.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1121–1134.

Landed, K., Thaler, V., & Reitsma, P. (2008). Spelling pronunciations: Transforming irregularity into regularity. *Learning and Instruction, 18*, 295–308.

Lin, L.-M., & Zabrucky, K. M. (1998). Calibration of comprehension: research and implications for education and instruction. *Contemporary Educational Psychology, 23*, 345–391.

Lucangeli, D., & Cornoldi, C. (1997). Mathematics and metacognition: What is the nature of the relationship? *Mathematical Cognition, 3*, 121–139.

Montague, M. (1998). Research on metacognition in special education. In T. E. Scruggs & M. A. Mastropieri (Eds.), *Advances in learning and behavioural disabilities* (pp. 151–183). Greenwich, CT: JAI.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation, 23*, 125–141.

Nist, S. L., Simpson, M. L., & Olejnik, S. (1991). The relation between self-selected study processes and test performance. *American Educational Research Journal, 28*, 849–874.

Otero, J., Campanario, J. M., & Hopkins, K. D. (1992). The relationship between academic achievement and metacognitive comprehension monitoring ability of Spanish secondary school students. *Educational and Psychological Measurement, 52*, 419–430.

Pressley, M. (2000). Development of grounded theories of complex cognitive processing: Exhaustive within- and between study analyses of thinking-aloud data. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 262–296). Lincoln, NE: Buros Institute of Mental Measurements.

Pugalee, D. K. (2001). Writing, mathematics, and metacognition: Looking for connections through students' work in mathematical problem solving. *School Science and Mathematics, 101*, 236–245.

Rittle-Johnson, B., & Siegler, R. S. (1999). Learning to spell: Variability, choice and change in children's strategy use. *Child Development, 70*, 332–348.

Savolainen, H., Ahonen, T., Aro, M., Tolvanen, A., & Holopainen, L. (2008). Reading comprehension, word reading and spelling as predictors of school achievement and choice of secondary education, *Learning and Instruction, 18*, 201.

Schraw, G., Potenza, M. T., & Nebelsick-Gullet, L. (1993). Constraints on the calibration of performance. *Contemporary Educational Psychology, 18*, 455–463.

Sénéchal, M., Basque, M. T., & Leclaire, T. (2006). Morphological knowledge as revealed in children's spelling accuracy and reports of spelling strategies. *Journal of Experimental Child Psychology, 95*, 231–254.

Siegler, R. S. (2000). The rebirth of children's learning. *Child Development, 71*, 26–35.

Sinkavich, F. J. (1995). Performance and metamemory: Do students know what they don't know? *Instructional Psychology, 22*, 77–87.

Soenens, D. (2002, May 8). Een nul voor Nederlands [They failed Dutch]. *Knack*.

Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology, 27*, 51–79.

Steffler, D. J., Varnhagen, C. K., & Friesen, C. K. (1998). There are more to children's spelling than the errors they make: Strategic and automatic processes for one-syllable words. *Journal of Educational Psychology, 30*, 492–505.

Templeton, S., & Bear, D. R. (Eds.). (1992). *Development of orthographic knowledge and the foundations of literacy: A memorial festschrift for Edmund H. Henderson.* Hillsdale, NJ: Erlbaum.

Templeton, S., & Morris, D. (2000). Spelling. In M. L. Kamil, P. B. Mosenthal, D. P. Pearson, & R. Barr (Eds.), *Handbook of reading research* (pp. 525–543). Hillsdale, NJ: Erlbaum.

Tobias, S., & Everson, H. (2000). Assessing metacognitive knowledge monitoring. In G. Schraw & J. C. Impara (Eds.), Issues in the measurement of metacognition (pp. 147–222). Lincoln, NE: Buros Institute of Mental Measurements.

Van Bon, W. H. J., & Uit De Haag, I. J. C. A. F. (1997). Difficulties with consonants in the spelling and segmentation of CCVCC pseudowords: Differentiation among Dutch first graders. *Reading and Writing: An Interdisciplinary Journal, 9*, 363–386.

Van Kraayenoord, C. E., & Schneider, W. E. (1999). Reading achievement, metacognition, reading self-concept and interest: A study of German students in grade 3 and 4. *European Journal of Psychology of Education, 14*, 305–324.

Veenman, M. V. J. (2003, August). The assessment of metacognitive skills: What can be learned from multi-method designs? Paper presented at the 10th Conference of the European Association for Research on Learning and Instruction, Padova, Italy.

Veenman, M. V. J., & Beishuizen, J. J. (2004). Intellectual and metacognitive skills of novices while studying texts under conditions of text difficulty and time constraint. *Learning and Instruction, 14*, 619–638.

Veenman, M. V. J., & Elshout, J. J. (1999). Changes in the relation between cognitive and metacognitive skills during the acquisition of expertise. *European Journal of Psychology of Education, XIV*, 509–523.

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning. Conceptual and methodological considerations. *Metacognition and Learning, 1*, 3–14.

Verhoeven, L., Schreuder, R., & Baayen, R. H. (2006). Learnability of graphotactic rules in visual word identification. *Learning and Instruction, 16*, 538–548.

Vlaamse Onderwijs Raad. (2006). De lat hoog voor talen in iedere school. Goed voor de sterken, sterk voor de zwakken [High standards for languages in every school. Good for the best, tough for the weaker ones]. Retrieved December 7, 2006, from the World Wide Web: http://www.ond.vlaanderen.be

Vrijders, J., Vanderswalmen, R., & Beeckman, A. (2007). Tien vragen over de spelling van studenten Hoger Onderwijs [Ten questions about the spelling of higher education students]. *Logopedie, 20*(4), 77–88.

Wakely, M. B., Hooper, S. R., de Kruif, R. E. L., & Swartz, C. (2006). Subtypes of written expression in elementary school children: A linguistic-based model. *Developmental Neuropsychology, 29*, 125–159.

Zhang, L. J. (2001). Awareness in reading: EFL students' metacognitive knowledge of reading strategies in an acquisition-poor environment. *Language Awareness, 10*, 268–288.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*, 51–59.

Zimmerman, B., & Reisemberg, R. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology, 22,* 73–101.

Chapter 18
Computer Use in a Primary School: A Case-Study of Self-Regulated Learning

Lena Swalander and Anne-Mari Folkesson

1 Introduction

Learning strategies adopted in the first school years might influence students' learning habits through the following school years; thus, it is important to investigate learning environments in primary school. Classroom studies have found that use of computer increases motivation (Alexandersson, Linderoth, & Lindö, 2000; Enochsson, 2004; Folkesson, 2004; Rosas et al., 2003). Moreover, computer use can help students to become self-regulated learners as shown empirically by Watts and Lloyd (2004) and in research overviews by Azevedo (2005) and Schraw (2007). On the other hand, computer use may create superficial learning habits and this has been found to cause a worry and an uncertainty on the part of the teachers (Alexandersson et al., 2000; Watts & Lloyd, 2004). To be able to explore children's learning habits in a computer-supported learning environment, with reference to its opportunities for self-regulated learning (SRL), more studies on young children are needed. According to Dermizaki, Leondari, and Goudas (2009), there is also a need for more comprehensive accounts of SRL in context since the early identification of student's strengths and weaknesses during problem solving might be facilitated. Therefore, the main objective for the present study was to make an in depth exploration of the features of SRL in a computer-supported learning environment in primary school.

The role of computers in primary education has mostly been investigated with focus on technical issues or on attitudes towards computer use. Only few studies have examined the educational value of computer use (Tondeur, van Braak, & Valke, 2007). Tondeur et al.'s (2007) review was a starting point for the development

L. Swalander (✉)
Department of Human Sciences, Kalmar University, 391 82 Kalmar, Sweden
and
Department of Behavioural Sciences and Learning, Linköping University, 581 83, Linköping, Sweden
e-mail: lena.swalander@lnu.se

A.-M. Folkesson (✉)
Department of Human Sciences, Kalmar University, 391 82, Kalmar, Sweden
e-mail: anne-mari.folkesson@hik.se

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_18,

of an instrument aiming at the construction of a typology of computer use in primary school. These authors rejected a one-dimensional approach with limited foci, such as software applications, and teacher or pupil attitudes. Instead, they developed a more comprehensive scale, based on items in teacher surveys, which resulted in three factors: “the use of computers as an information tool”, “the use of computers as a learning tool” and “learning basic skills”. Nevertheless, Tondeur et al.’s (2007) study also takes its starting point in the computer per se and not in the learning environment, as other researchers had done (Alexandersson et al., 2000; Chen & Chang, 2006; Waite, 2004). Thus, although Tondeur et al. (2007) have provided a valuable instrument to investigate the different types of computer use in primary school, they did not discuss their results with respect to different learning strategies. This is important since, as mentioned above, computer-supported learning has been found to enhance surface learning (Watts & Lloyd, 2004). In a study of different ways of using computers, Wittwer and Senkweil (2008) found that access to computers was not linked to achievement. However, for a small group of students who used computers in a self-determined way, a positive effect on achievement was found and the suggested rationale was that this kind of use engaged them in problem-solving activities (Wittwer & Senkbeil, 2008).

Therefore, understanding how young students use computers in regular classroom and how this is related to SRL is an issue that merits continued attention. This was the aim of the present study. In what follows, first, we discuss issues related to computer-supported collaborative learning and to SRL; then we focus on components of SRL, namely metacognition and motivation, as well as to learning environments with computer use that promote SRL. Finally, we present an observational study of computer use in the classroom of students aged 7–9 years as a case study of a learning environment that enhances self- and co-regulation through computer use.

1.1 Computer-Supported Collaborative Learning

A collaborative learning context is defined as one in which the participants engage in a “coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem” (Roschelle & Teasly, 1995, p. 70). It has also been noted that cognitive co-regulation involves continuous attempts by the partakers to coordinate their activities and language (Salonen, Vauras, & Efklides, 2005). Such self- and co-regulatory skills have been found to be enhanced when children are working together on computers (Denis & Hubert, 2001; Kusunoki, Sugimoto, & Hashizume, 2002). This falls well in line with the results from Zhang, Scardamalia, Lamon, Messina, and Reeve (2007), who found that collaborative work was beneficial for the advancement of community knowledge and that young children can take responsibility for the improvement of their own knowledge.

However, Salovaara and Järvelä (2003), who studied computer-supported collaborative learning (CSCL), found that students often reported surface-level strategies and a fact-oriented approach. Nevertheless, in their database, there was also evidence of collaborative work and discussion and use of deep-level cognitive strategies. Although the Salovaara and Järvelä (2003) study was limited to only one school subject, using one specific computerised forum, their findings, indicating high levels of collaboration and discussion, have been found in other studies employing different designs as well (Alexandersson et al., 2000; Sullivan & Pratt, 1996; Watts & Lloyd, 2004). According to Scardamalia, Bereiter, McLean, Swallow, and Woodruff (1989) a unique advantage of educational software is the possibility to initiate cross-curricular links and to integrate knowledge from several disciplines. They argue that it is unlikely that the computers per se can foster cooperative learning but they could play a role in a classroom culture where cooperative learning is encouraged (Scardamalia et al., 1989).

In an experimental study of CSCL in primary school, Prinsen, Volman, Terwel, and van den Eeden (2009) showed an effect of such a learning environment both on the degree of and on the quality of participation in CSCL. However, there were individual differences in CSCL of students with respect to learner characteristics; that is, students with minority background benefited less in terms of degree of participation and boys benefited less in terms of quality. Furthermore, the review by Winters, Greene, and Costich (2008) showed that learning strategies in the context of computer-based learning environments tended to be more effective for academically successful students. They also showed that planning and monitoring were more prevalent among students with high prior knowledge than among those with lower prior knowledge (Winters et al., 2008). These individual differences may have implications for the development of SRL.

1.2 Self-Regulated Learning (SRL)

The notion of SRL has been defined as the degree to which students are "metacognitively, motivationally and behaviourally active participants in their own learning process" (Zimmerman, 1986, p. 308). The definitions of the notions of self-regulation, SRL, and metacognition, and their theoretical and empirical boundaries have been discussed by several researchers in the field, who have found that the boundaries between these notions are not always clear and that they sometimes are used interchangeably (Dinsmore, Alexander, & Loughlin, 2008; Lajoie, 2008). In the present study, Zimmerman's (1986) definition of SRL was adopted. It is pointed out that SRL should not be understood as a mental ability or an academic performance skill; rather "self-regulation refers [instead] to the self-directive *process* through which learners transform their mental abilities into task-related academic skills" (Zimmerman, 2001, p. 1).

In their review of self-regulation in the classroom, Boekaerts and Corno (2005) analysed various conceptualisations of self-regulation that exist in educational psychology and the instruments constructed to investigate capabilities to self-regulate. They stated that educational psychology researchers have narrowed the notion of SRL by persistently focusing their attention on learning and achievement goals.

SRL includes a compound set of functions situated at the crossroads of several fields of research in cognitive psychology, such as problem solving, decision making, metacognition, conceptual change, motivation, and volition (Boekaerts & Corno, 2005). Although the various SRL theories emphasise somewhat different features "all theorists assume that students who self-regulate their learning are engaged actively and constructively in a process of meaning generation and that they adapt their thoughts, feelings, and actions as needed to affect their learning and motivation" (Boekaerts & Corno, 2005, p. 201). This assumption is in accordance with Greene and Azevedo's (2007) theoretical review, where they state that "although there are important differences between various theoretical definitions, self-regulated learners are generally characterised as active, efficiently managing their own learning through monitoring and strategy use" (Greene & Azevedo, 2007, p. 334–335). Self-regulated students are also assumed to be able to select their use of metacognitive and motivational strategies as well as to create a beneficial learning environment. They can also independently choose both the form and the amount of their needed instruction (Zimmerman, 2001).

1.2.1 Metacognition in SRL

According to Zimmerman (1986), an important aspect of SRL is metacognition. The definition of metacognition used by Efklides (2006) follows the original definition by Flavell (1979) as cognition about cognition. However, there are many definitions and overlaps between metacognition, self-regulation and self-regulated learning (Dinsmore et al., 2008). In a more recent theoretical article, Efklides (2008) extended the original Flavell definition as well as the definitions by Nelson (1996) and Nelson and Narens (1994). According to Efklides (2008), metacognition can be divided into three facets, namely metacognitive knowledge (MK), metacognitive experiences (ME), and metacognitive skills (MS). Specifically, MK involves declarative knowledge stored in memory, information regarding persons, tasks, strategies, and goals; ME involve a person's awareness and feelings about a task as well as the awareness and feelings when processing relevant information, which can be seen as the interface between a person and a task. Moreover, Efklides (2006, 2008) stressed that the nature of metacognitive feelings is both affective and cognitive. Finally, MS involves the intentional strategy use for controlling cognition and these skills encompass strategies for orientation, planning, regulation, monitoring, and evaluation.

Most SRL research emphasize MS, that is, control processes. However, Salonen et al. (2005) argued that being aware of one's own and fellow students' cognition through ME is necessary for metacommunication and regulatory control processes. They reported that co-regulation, in certain learning situations, could create obstacles for learning, such as misunderstandings about the others' cognition. Salovaara and Järvelä (2003), on the other hand, suggested that a lack of developed inquiry and collaborative learning culture could explain the procedural learning strategies found among students rather than metacognition in its various facets.

Therefore, it is not clear what exactly is the reason for students' lack of SRL or co-regulation. To complicate things further, Patrick, Ryan, and Kaplan (2007) showed that when students experience academic support from their peers they were more prone to self-regulation as well as engagement in task-related interactions, that is, they were more motivated. This implies that motivation is as important for SRL as metacognition.

1.2.2 Motivation in SRL

As argued above, motivation is a significant component of SRL, which in its simplest definition can be defined as being moved to do something (Ryan & Deci, 2000). Goal-orientation theories (Dweck, 1986, 2000) suggest that, as students' interest in learning for the sake of improving their knowledge or skills increases, so does their valuing and reported use of cognitive learning strategies and self-regulation. In a study of the interrelations between cognitive self-regulation and motivation, Dermizaki et al. (2009) found that there was a reciprocal relation between motivation and cognitive self-regulation. Moreover, people vary both in their level of motivation and in their orientation of motivation (Ryan & Deci, 2000), and this can be connected to what goals they choose to adopt. The primary attitudes and goals that set off action can be considered as the basis for the orientation of motivation. Deci (1992) proposed that individuals are intrinsically motivated when they undertake an activity because of their interest in it. Therefore, intrinsically motivated behaviours are manifested when an individual is doing an activity because he or she is interested in it or finds it enjoyable (Ryan & Deci, 2000; Ryan, Kuhl, & Deci, 1997). The other type of motivation is extrinsic motivation, which "is a construct that pertains whenever an activity is done in order to attain some separable outcome" (Ryan & Deci, 2000, p. 60).

Although intrinsic motivation is clearly an important type of motivation, most activities are not, strictly speaking, intrinsically motivated. This is especially the case after early childhood, as the freedom to be intrinsically motivated becomes increasingly curtailed by social demands and by roles that require individuals to assume responsibility for non-intrinsically interesting tasks. In school, it appears that intrinsic motivation becomes weaker with each advancing grade (Ryan & Deci, 2000).

Pintrich and Schrauben (1992) hypothesized that students who are interested in their courses and judge them to be important will employ cognitive and metacognitive strategies more often than students who are not as interested in their courses or who do not feel that their courses are important to them. Pintrich (2003) developed a motivational science perspective, which summarises aspects of student motivation in search of motivational generalisations. He described five different conceptualizations of what characterizes motivated students: (a) adaptive self-efficacy and competence beliefs; (b) adaptive attributions and control beliefs; (c) higher levels of interest and intrinsic motivation; (d) higher levels of value; (e) goals (Pintrich, 2003).

1.2.3 Learning Environment and SRL

One important issue regarding studies of learning is that learning environments differ with respect to self-regulation. According to the Swedish National Assessment of the quality of reading and writing processes (Skolverket, 1998a), learning environments promoting SRL were very rare in primary school. Similar results were also reported by Tondeur et al. (2007) who, almost 10 years after the Skolverket (1998a) study, found that students writing their own texts on the computer was still not prevalent in the early grades of primary school. Nevertheless, in a case study of primary school, where computers were frequently used for extended production of children's own texts (Folkesson, 2004), a positive effect on reading comprehension was found (Folkesson & Swalander, 2007). However, the specific learning conditions (i.e., SRL) in the classrooms studied was not in focus of the Folkesson's studies; therefore SRL could not be separated from the extended writing on computers, which was the main explanatory factor for children's reading achievement. Hence, there is a need for an empirical description and analysis of the features of the computer-supported learning environment that enhanced students' reading performance. Indeed, for improving the conceptual understanding of self-regulation in computer-supported learning environments, Schraw (2007) suggested the formulation of explicit process models that explain the effect of computer-based scaffolding on SRL. The study reported below is an attempt to describe the processes involved in students' SRL in a computer-supported learning environment, although no a priori theoretical model was formulated.

To summarise, both self-regulation and co-regulation are closely linked to motivation and metacognition. Also, SRL has been related to computer use. However, most research has been conducted in the higher school grades, whereas research in the lower grades of primary school is sparse. As for research on computer use it has often focused on different computer software and experiments of learning in different subjects and there are few studies exploring computer-supported classrooms in a naturalistic setting. Thus, in order to develop our understanding of young school children's handling of SRL in a computer-supported learning environment, the present study was focused on children in the age of 7–9 years in regular classroom work, in which computers were extensively used.

1.3 The Present Study

The aim of the present study was to describe and analyze the learning environment of a computer project in primary school by exploring how SRL was manifested in an educational setting in which computer use was an integrated tool in everyday classroom work. Special focus was on how three aspects of SRL, namely metacognition, motivation, and behaviour, were manifested. Following Zimmerman (1986), the assumption guiding the classroom observations was that the integration of computer use in everyday classroom work will provide many opportunities for SRL as well as for co-regulation. This is so because computer use creates more opportunities for new activities that allow both self- and co-regulation for carrying out computer-supported learning.

2 Method

2.1 Design

This is an exploratory case study of a 3-year computer project conducted in a Swedish primary school. The project was aimed at developing computers as a pedagogical tool in a broad sense and not related to specific subjects or software. During 3 years time the learning environment of two teachers and their students were investigated using multiple methods. The project started when the students attended Year 2 and these students were followed through Year 3. In the third year of the project the teachers started over again with a new group of students in Year 1, hence "Year 1" is the last year of the project. Consequently, the students in Years 2 and 3 formed one cohort (Cohort A) and the students in Year 1 another cohort of students (Cohort B). The case study was conducted in a naturalistic setting (described below), where different kinds of data were collected, and a comprehensive picture of the learning environment could be obtained. In Table 18.1 data collection is presented in chronological order.

In sum the data consisted of 36 h of participant observation documented by field notes and tape-recorded conversations, including informal conversation with the children and the teachers. Furthermore, there were 37 h of un-structured teacher interviews regarding the teaching process with special focus on the perceived differences when using the computers as compared with using paper-and-pencil. Additionally, the data include 5 h of semi-structured group interviews with the children (see Appendix A) and, finally, different written documents from the children (see Table 18.1). Using observational data with small children has at least three strengths: (a) they reveal what the children really do rather than what they say they do; (b) behaviour can be linked to the context; (c) small children do not have the verbal ability to inform about their actions (Winne & Perry, 2000). However, one needs to be cautious when interpreting children's behaviour since behaviour can have many explanations (Winne, 2004). The rationale for using interviews and documents as complements to the observations is that information "the researchers are unable to see for themselves is obtained by interviewing people who did see them or by finding documents recording them" (Stake, 2005, p. 453). Many researchers in the field have also underlined the importance of authentic and naturalistic contexts, which can provide new valuable information about SRL-processes (Whitebread et al., 2009; Zimmerman, 2008).

Finally, focusing on two particular classes as a case study has the advantage that it focuses on the discovery, interpretation, and the shedding of light on the

Table 18.1 Overview of data collection Year 2, 3, and 1

Fall; Year 2	Spring; Year 2	Fall; Year 3	Spring; Year 3	Fall; Year 1	Spring; Year 1
Teacher interview – Planning (2 h) – September	Document – Children's stories about the theft of the computers – January	Participant observation of a day in school (6 h) – November	Participant observation of a day in school (12 h) – March and May	Participant observation of a day in school (6 h) –October	Participant observation of a day in school (6 h) – February
Document – Children's evaluation of their different tasks, ranking of tasks and rationale – December	Participant observation of a day in school (6 h) – March	Document – Children's evaluation of their different tasks, ranking of tasks and rationale – December	Teacher interview (5 h) – June	Teacher interview (5 h) – December	Teacher interview (10 h) – April and May
	Group interviews with all children (5 h) – April	Teacher interview (5 h) – December		Document – Children's evaluation of their different tasks, ranking of tasks and rationale – December	Document – Children's evaluation of their different tasks, ranking of tasks and rationale – December
	Document – E-mails from the children to the researcher – March to June				
	Teacher interview (10 h) – June				

interplay between significant factors where it is impossible to differentiate between cause and effect (cf. Merriam, 1998). A case study focuses on making holistic descriptions and explanations and is mainly appropriate in situations where the separation between a variable and its context is unattainable. Since the activities in a naturalistic setting "are expected to be influenced by context, [so] contexts need to be described, even if evidence of influence is not found" (Stake, 2005, p. 452). Hence, if the aim of a case study is not to generalise in a positivistic sense, data gathering is conducted to "seek what is ordinary in happenings, in settings" (Stake, 2005, p. 453).

2.2 Participants

Cohort A consisted of 39 children (age ranged from 7 to 9 years), who were followed during their second and third school year, and Cohort B of 29 children (age ranged from 6 to 7 years) who were followed during their first school attended year; all students attended a compulsory school in a small Swedish town. The distribution of gender was 22 girls and 17 boys in Cohort A and 15 girls and 14 boys in Cohort B. Pupils in this school were of low and medium socio-economic status. Of the total sample, 54 children had Swedish as their mother tongue and 14 of the pupils had a different mother tongue than Swedish.

In this school, two teachers had voluntarily developed a computer project. The teachers in the project were two women of about 40 years of age who had a long experience as primary school teachers. Their pedagogical ideas can be summarised as an inspiration of a "little of the best" and that they wanted the children to take on more responsibility and to act more independently. However, they did not lean against any single pedagogical model. The teachers had always been interested in new ideas and constantly asked themselves how their work could be developed.

2.3 Setting and Procedure

The teachers mainly used open-ended software, with which the children could express themselves in their own language and, specifically, avoided computer software aimed at drill and practice. They also used multimedia software for the integration of text, picture and sound. The children were not allowed to reproduce texts but had to write everything in their own words. The children used computers in their early reading and writing development and the computers were used as a pedagogical tool in their everyday classroom work. Hence, computer use was not limited to specific software or specific subjects, as often has been the case in previous research. The teachers had very limited computer-use skills and very limited technical knowledge, but from the beginning of the project they received technical support every week. Thus, their learning developed together with that of their students.

The two teachers had classrooms close to each other, with a library and a computer room on the same floor. Altogether they had 20 computers, of which ten were located in the computer room and five in each of the two classrooms. They also had digital cameras, a scanner, a server, and Internet connection. The children were free to work in any room on the whole floor, including the library and the corridors, and to collaborate with peers of their own choice. This environment has many features in common with a Swedish pre-school environment and is rather unlike a traditional primary school classroom where the children sit at their desk listening to the teacher and work individually with readymade materials. In the classrooms, the desks were sometimes put together in small groups and sometimes placed in a U-shape. The five computers in each classroom were placed in a row on one side of the room. There were only short periods of "chalk-and-talk", mostly of practical and social character, in the morning, before and after lunch, and at the end of the day. The teachers introduced thematic work to the whole class, and now and then the pupils presented their work to their peers. The schedule did not indicate different subjects since reading, writing and mathematics often were integrated into thematic works. However, in mathematics ready-made workbooks were also used. Thus, it was not possible to distinguish different approaches to computer use linked to different school subjects.

Teaching was normally carried out as conversations related to the pupils' own production. During the whole day, both pupils and teachers initiated talk like an everyday conversation among equals, that is, there were no traditional teacher-directed dialogues, where the pupils were to provide an answer, evaluated as wrong or right by the teacher, as often is the case in traditional school environments. The work was characterized by the children's own production, authentic writing, discussion about the productions between the children, thematic work, participation, and involvement (Folkesson, 2004). After the school day, the teachers made sure that there was a progression in their learning by checking the children's files on the server. As for the quality of learning the children in Year 2 were tested on reading comprehension and compared to a national sample, where the results showed that these children performed significantly better (Folkesson & Swalander, 2007).

During the day, the children could choose from tasks listed on the white board. They worked at their own pace and there were no rules that regulated the children's collaboration with other children. The teachers moved around in the classrooms and in the computer room helping and supporting the children with comments on language and spelling as well as with technical help when needed. A typical task was to produce own stories where text, pictures and sound were integrated on the computer. These stories had various contents linked to the actual thematic work with the common concept of being a digital story combining the text and the pictures into a slide show with digitally recorded sound. Everything children produced was sorted into folders on a server with one folder for each child, labelled with his or her name. The children had their own passwords, and only the teachers had access to all folders.

2.4 Data Analysis

All data were transcribed and content analysed (Wilkinson, 2003) in two steps. The first step was deductive, that is, the content of the transcriptions was first categorised into the three theory-driven main themes revealed from SRL definitions, namely metacognition, motivation, and behaviour (Zimmerman, 1986). The second step was an inductive step where the authors studied the data transcriptions independently in order to find examples of SRL (Zimmerman, 1986) with the intention of creating sub-themes to the main themes. In line with Whitebread et al. (2009), the analysis was a "blend of a priori categories [of behaviour] deriving from previous research literature and new categories emerging from a 'grounded' analysis of data" (p. 72). The findings were compared and discussed by the authors until there was full agreement so as to find typical events that could be related to the critical features of SRL. The sub-themes were generated from the data to reflect the width and depth in the material in order to reveal SRL activities in this learning environment. In what follows, the sub-themes are presented with cited examples from data transcriptions together with reflective comments on the findings.

3 Results

The aim of the present study was to describe and analyze the learning environment in a primary school computer project by exploring how SRL was manifested in everyday classroom work. The focus was on three main aspects of SRL, namely metacognition, motivation and behaviour. The results were categorised with the three aspects as main themes followed by their respective sub-themes.

3.1 Metacognition

The observed data regarding metacognition indicated that the children were aware of different learning processes and that they were able to verbalise this awareness. Five different sub-themes were found: (a) awareness of the processes in the learning environment, (b) self-monitoring the working process, (c) planning, (d) awareness of the relations between different learning activities in the learning process, and (e) no benefit from computer use.

3.1.1 Awareness of the Processes in the Learning Environment

The first excerpt illustrates the awareness of *active learning processes* as verbalised by two boys when asked to describe what they do when they learn.

Boy 1: We talk to each other.
Boy 2: You can write…and then you can watch.

(Group interview, spring semester, Year 2)

These boys seem to understand that learning and gaining knowledge is about being an active participant. Such understanding is also the case in the following excerpts, where the children also show an awareness of the role of *individual differences in learning*.

Interviewer: Do you know what you do when you learn?
Girl 1: You learn by listening…with your heart. In a way…you listen in your own way. You can understand in your own way with your heart.
Girl 2: You learn everything by yourself I think.

(Group interview, spring semester, Year 2)

Both these girls seem to understand that they have a personal role when listening and processing information. Another example shows children having *strategies for self-regulation and self-evaluation*.

Interviewer: How can you be sure that you know something?
Boy 1: I test many times.
Boy 2: You can try many times and if you succeed nearly every time you know it rather good.
Girl 1: You see if you know and then if you don't know you can try again.

(Group interview, spring semester, Year 2)

The children in the excerpt above seem to have strategies to evaluate and decide whether or not they have learnt something. Since they are aware of strategies for learning, they are not dependent on the teacher telling them. In the next excerpt the children evaluate material and make decisions, and also exemplify strategy use, namely quick reading to get a notion about a text.

> The boys are discussing some news they found on the Internet. They click here and there, they scroll up and down, read some headlines and stop to think. Finally they choose a text about robbers and policemen. They say that they will read aloud in front of the class.

(Participant observation, fall semester, Year 3)

These boys read small pieces of text, they think, and finally they choose. The awareness that learning can be about *active searching for knowledge* is shown in the next excerpt.

> I think you have computers to learn a lot of things…and…you get to know more if you have done research about something. You can always search the Internet and things like that.

(Group interview, spring semester, Year 2)

These boys seem to link the search for knowledge to the use of computers. The computers are also prominent when it comes to self-monitoring one's work, as shown in the next sub-theme.

3.1.2 Self-Monitoring the Working Process

The data showed that the children were able to self-monitor their own working process when using the computers across a variety of situations, that is, complex procedures and sequential procedures; the data also showed some ambivalence to this. First a boy shows his capability of monitoring the *complex procedures* when creating a digital story.

> A boy has created a digital story in KidPix about a boy who is going fishing. He is asked whether he makes the pictures or the text first. He answers that he makes the picture first and creates the story at the same time. He then writes down the words and finally audio tapes the text.

(Participant observation, spring semester, Year 2)

This child was obviously familiar with the procedures he used when creating the story, and he was able to self-monitor the order of putting words, pictures, and sounds together. The next excerpt is about a boy who explains the *sequential procedures* and at the same time he is self-monitoring his work when using the table functions for a school task.

> A boy is working at the computer and the observer asks him what he is doing. He answers: "and than you ca…go to Table, insert table…and then you can…, how many columns you want, two then, you need twooo (he is talking and clicking at the same time)…then the width should be…and here you can write something."

(Participant observation, spring semester, Year 2)

The following excerpt also shows the same familiarity with the software, but this time illustrating how the children created a digital story.

> We are making a slide show in KidPix. First you make 5, 6 pictures, then you put the pictures in PowerPoint and talk in the sound in microphones.

(Excerpt from an e-mail three children wrote to the observer, spring semester, Year 2)

These children show an easiness to self-monitor several functions when using the computer for making a digital story. The fact that it was easy to monitor the work on the computer could also cause some *ambivalence* for some children.

> "On the computer you can cheat by erasing" she says giggling and kind of embarrassed. She adds that it is not really cheating. You just change and it becomes nicer.

(Participant observation, spring semester, Year 2)

This reflection shows that the girl seemed to like the possibilities to make things nicer and the embarrassment and giggling can be seen as affective components that underlie her thoughts about cheating.

3.1.3 Planning

The children were expected to take a great deal of responsibility for planning their work. One of the main reasons for the teachers to start this project was to teach the children to take responsibility for their learning. The following excerpt is an example of the plans that the teachers wrote on the whiteboard. In this plan the numbered assignments were compulsory but the children could choose in which order to fulfil them. The assignments without numbers were voluntary and could be worked on when the children wanted to.

> On the whiteboard in both rooms is written:
>
> 1. Math pp. 58–61; 2. Diary week # 12; 3. "Research"; 4. Theatre picture; 5. Manuscript + text (slide show); 6. Slide show KidPix to PowerPoint + sound, Review, Handwriting, Reading, Practice book, Math game, Lexia, Extra slide show PowerPoint.

(Participant observation, spring semester, Year 2)

The teachers had done this planning but everyday the children had to make choices and explicitly tell the teacher that they were *having own plans for the day*. The following excerpt is an example of this.

> After the break the teacher asks the children to tell her their plans for the rest of the morning lessons. Each child tells her what he or she is planning to do. All children are expected to have plans of their own.

(Participant observation, spring semester, Year 2)

This way of working implicitly requires children to self-regulate their learning processes since the teachers do not tell every child what he or she is supposed to do at a specific time.

3.1.4 Awareness of the Relations Between Different Activities in the Learning Process

The data showed that the children were aware of how different learning activities were interrelated, that is, the letters of the alphabet as the basis for reading, the close link between reading and spelling, and reflection about layout and readability. The first excerpt illustrates the children's awareness of *the letters of the alphabet as a basis for reading.*

> Two children took an autonomous initiative and started to work on a letter task. They are working in several steps all aimed at recognition, writing and finally reading. One of the children says 'It's really fun. It is good to know the letters so that I can read. Fun to write the letters. I have learnt a lot.'

(Participant observation, fall semester, Year 1)

The next excerpt shows another basic relation, namely *the close link between reading and spelling*. When reading a computer written text aloud, it is easy to detect mistakes in writing. In this example a boy finds that putting spaces in the wrong place makes the text impossible to understand.

> A boy is reading his text aloud to another boy. He then notices a mistake, that he had written 'in te' [inte = not, in Swedish] with a space. With the help of another boy he changes to 'inte'.

(Participant observation, spring semester, Year 1)

Here we see that the boy understands that it is not only the letters that matter to make up the word but also the spaces in between the letters. From the more experienced readers (Year 3) the teachers commented that the children showed a more developed understanding and *reflection about layout and readability*, as shown in the excerpt below.

> In WordArt we showed the children all the possibilities. Now the children are able to decide what is suitable, the effect of the layout. In the beginning everything was very flashy when they, for example, were asked to make a poster. Now the children can reflect on what is suitable and if it is easy to read.

(Teacher interview, spring semester, Year 3)

3.1.5 No Benefit from Computer Use

In the material there were also three situations when the use of computers did not help the children: (a) problems with technological messages, (b) problems with the spelling function, and (c) turning down help from the computer. The first excerpt exemplifies a girl who had *problems with technological messages* and clearly needed teacher support.

Teacher: What does it say?
Girl: Something appeared.
Teacher: Not connected to the server. It's probably…in a way…. something's awkward…OK! And then you may search here and see if you find something.
Girl: Mmmm.

(Participant observation, fall semester, Year 3)

This girl encounters a problem when a warning sign appears on the screen. She does not understand what it says and needs the teacher to help her solve the problem. Problems in reading ability are also present in the next example when a boy is writing a text on the computer and have *problems with the spelling function*.

> A boy had misspelled the word 'nestan' [nästan = almost, in Swedish] and the computer indicated it as misspelled and gave several suggestions for other words. The boy quickly chose 'Märsta' [a Swedish town] and happily went on writing since there were no indications of misspelling. After a lot of guessing, when trying to answer questions from the observer, the teacher came and solved the problem.

(Participant observation, spring semester, Year 2)

The example above shows a child who does not understand how to use the spelling function in a real situation. He only knows that it exists so by himself he cannot benefit from this function and needs the teacher to help him use it properly. The two examples above show children whose reading ability is not developed enough to understand the different suggestions the word processor gives. Another kind of problem with the word processor is that some children do not like the spelling function and they were even *turning down help from the computer*.

Girl: Then the computer shows if it's wrong or something and you're not supposed to know that. You're supposed to learn if it's wrong. In that case it is better that the teacher tells you so.
Interviewer: Is it better to hear it from the teacher than to see on the computer?
Girl: Yes, then you know when you've made a mistake.
Interviewer: Do you think it's a bit like cheating (i.e. using the spelling function)?
Three girls: YES!

(Group interview, spring semester, Year 2)

These girls think it is wrong to self-regulate their learning and they want to be led by the teacher; therefore, they do not seem to appreciate the benefit from all functions of the computer.

3.2 Motivation

In line with previous research, it was found that the computers were highly motivating for many different reasons. The data in this study were categorised into five sub-themes: (a) positive affect; (b) negative affect; (c) opportunities for making choices; (d) work eagerness; and (e) teacher demands.

3.2.1 Positive Affect

From the observations it was obvious that the children really enjoyed working on the computers. This was evident in the pleasure they had making things happen on the screen, the potential for many different activities with the computer, and feelings of emptiness without the computers. It seemed as if they especially took *pleasure in making things happen on the screen*.

> A girl finds working on computers fun and she likes to "scribble" with KidPix. She said: 'It is fun to fiddle with the stamps [readymade pictures from ClipArt].'

(Participant observation, spring semester, Year 2)

Another example of this enjoyment of watching things happen on the screen is shown below.

> A boy who is making a slide show about himself appears to find it really funny to see images of his friends come flying in from different angles on the computer screen. He continues with a second theme in his story about himself after having a supporting conversation with a peer.

(Participant observation, spring semester, Year 2)

Since the children appeared to enjoy how easy it was to create something and then make changes they were stimulated to test new ideas. On a computer everything is changeable and this might be inspiring. In writing and drawing with a pencil you can also erase and change things but with the computer it is easier since less effort is needed because there is no manual work. It is also immediately perceptible. It is, however, important to notice that in these classes they often used crayons, which are not possible to erase, but on the computer you can use colours and still be able to erase.

Another source of positive affect was *the potential of many different activities with the computer* and that just thinking about or seeing the computer gives rise to a lot of ideas of what the children would like to do.

Interviewer: What do you think about when you see a computer?
Boy: Ehhh, I think it's fun and ehhh, you can use it for different things. You can….play games and search the Internet and yes…Something that's very good, is that you can use it for so many things.
Girl: When I see a computer I think that I want to play… and write a lot of things…because I always feel like writing when I see a computer…if you are bored… it's fun to work on computers I think.

(Group interview, spring semester, Year 2)

When seeing all the different positive opportunities the children talk about in relation to the computer, it is understandable that they really experienced *feelings of emptiness without the computers* when they were stolen for a short period of time as shown below.

> The classroom was all empty…very boring in school…really boring not to be able to use the computers. I hope it will be as before again. I think it's been quieter in the classroom. It's hard to write everything by hand. The worst is that PowerPoint is lost; it is the most fun program.

(A written story about when the computers were stolen, spring semester, Year 2)

Interviewer: What did you do when the computers were stolen?
Boy 1: I started to cry, I did and my best friend he tried to comfort me.
Boy 2: It was sad… I almost fainted.

(Group interview, spring semester, Year 2)

These examples give support to the idea of an overall positive affect most children had because of their access to computers in the classroom. Nonetheless, it is noteworthy that for a few children there was also negative affect experienced.

3.2.2 Negative Affect

Negative affect was associated with frustration due to technical problems and physical fatigue. The first example is about children reporting *frustration due to technical problems* and, therefore, the computers were not experienced in a positive way.

Interviewer: What do you do when you have technical problems?
Girl 1: Then you have to tell a teacher and then you think 'oh no not again', it happens to me rather often. And then when the teacher stands there talking and stuff, then you get tired…you even get a headache, when the teacher is checking if it has frozen or if it is something else or if a picture has disappeared or the whole file...I think it's very irritating if something like that happens.

(Group interview, spring semester, Year 2)

Sometimes the children did not complain but they still reported frustration that could be interpreted as negative affect.

> It makes trouble sometimes. Difficult to log on. The CD-ROMs get sticky and the computer freezes.

(Written evaluation, spring semester, Year 1)

> I think that it is not so easy to write on a computer, because I don't find the keys.

(Group interview, spring semester, Year 2)

Another perceived problem for some children was thus that they preferred handwriting and in addition, the child in the excerpt below reported *physical fatigue*.

> I think it's bad when the computer is broken, sometimes it's nuts. I think it was more fun without the computers. Because sometimes the computers are irritating, yes irritating… your eyes hurt in a way.

(Group interview, spring semester, Year 2)

Thus, the negative affect was caused by quite ordinary technical problems even for experienced users as well as problems with being a beginner at using computers.

3.2.3 Opportunities for Making Choices

There were many opportunities for making choices that often led to formulation of new tasks as well as to development of present tasks. In this learning environment the children had opportunities to perform other tasks than those created by the teacher. This implies that the children had the opportunities to use their intrinsic motivation for the *formulation of new tasks*.

> One boy asks the teacher if it was OK if he made a new slide show. The teacher was surprised that he remembered how to do it, since he had not used that programme for a long time. This time he wants to do a story about himself, going on vacation, his best friends and so on. He collects images of his best friends on the server.

(Participant observation, spring semester, Year 2)

The new tasks that the children invented were of various kinds. The boy in the excerpt above used knowledge from one situation and transferred it to a new self-chosen activity. However, the children did not always find completely new tasks, rather they generated a *development of the present tasks* as they got along.

> One boy who is making a digital story about himself…he is working in a non-linear and associative way. His ideas seem to come from his imagination as well as from photos. Now he is writing about London since he is going there soon. To illustrate he is looking for pictures in ClipArt. There he finds other pictures that give him new ideas to write about.

(Participant observation, fall semester, Year 2)

The many opportunities to make choices seem to stimulate the children in their working process and this might make them eager to work.

3.2.4 Work Eagerness

This was shown by a strong focus on task and also by a kind of playfulness at the computer. It was striking how the computer-supported learning environment encouraged the children to *focus on task*. This was observed on many occasions.

> When it is time for recess, the children are not eager to leave their work. A child seems very focused on creating her digital story and when the bell rings the observer asked if she is leaving for recess; "I don't know" she answers without any sign of interest in whether or not there is recess time now.

(Participant observation, spring semester, Year 2)

When creating something on the computer the children seem to be very focused and not eager to leave their work even when the bell rings for recess. One could even say that the children showed *playfulness at the computer.*

> A boy is working on his 'race car track', an assignment for illustrating a math problem; 'oops the whole car disappeared' he says to himself, starting all over again calmly and without frustration.

(Participant observation, fall semester, Year 1)

In the example above the computer could be related to the work eagerness but there were also examples of this attitude in the learning environment as a whole.

Interviewer: Is there something that is especially good that you learn in school?'
Girl: Yes, it's like when you have finished (a task) then you get new tasks to do.

(Group interview, spring semester, Year 2)

As shown above, the girl seems eager to work but there is also evidence of children having other thoughts about a demanding learning environment.

3.2.5 Teacher Demands

The teachers encouraged *hard work* from the children and this was sometimes experienced as positive extrinsic motivation and sometimes as negative.

> It took a lot of work to finish the Moominbook. But I was very pleased with it.

(Written evaluation, spring semester, Year 1)

Interviewer: Is there something you don't like in school?
Girl 1: Research is not very fun. It's tiresome.
Girl 2: Also research. It's so hard to write and stuff, and you have to search and write and find a lot of pictures and stuff.

(Group interview, spring semester, Year 2)

In sum, the data illustrating the theme motivation are mostly showing how the children are positive and encouraged to work when using computers. There are, however, indications of that this way of working does not suit all children.

3.3 Behaviour

The third main theme concerns the behavioural aspects of SRL. The most important finding was the high degree of *helpfulness* among the children and this constitutes the first sub-theme. The second sub-theme was *involvement*, that is, the children got very engaged in each other's work.

3.3.1 Helpfulness

The children were helping each other, both spontaneously and on the teacher's request. The helpfulness was manifested for different reasons, for example when help was needed because the teacher did not know how to solve a problem.

> The teacher is helping a girl with a digital story – something went wrong and neither the teacher nor the girl could solve the problem. A girl says, "I think S knows" and fetches her. S solves the problem as a natural thing and without "unnecessary" talking. Then, she goes straight back to her own work.

(Participant observation, spring semester, Year 2)

Help was also needed when the teacher did not have time to help a child at once due to being busy helping another child.

> A girl asks the teacher something. The teacher is currently helping someone else and cannot come to immediate assistance. The teacher then asks a boy to help the girl and he does it immediately without hesitating despite that he has to leave his friend and their collaborative work for a while.

(Participant observation, spring semester, Year 2)

The excerpt above and the next one are two examples of how the children seem to think it is natural and no struggle to help each other. Rather it is the norm as opposed to traditional classrooms where children are expected to sit quiet and where helping each other is regarded as a disturbance since the children then talk too much. What's more, helping or receiving help has also traditionally been regarded as cheating.

Interviewer: What are you doing now?
Boy: I'm helping.
Interviewer: Is he helping you?
Girl: He is my assistant.

(Participant observation, spring semester, Year 3)

The helpfulness was expressed in a variety of situations when the children were giving technical support, procedural instructions, and theoretical instructions or explanations. The first excerpt is an example of *technical support* where a modification of pictures was needed.

> A girl is creating a digital story about Ariel, the little mermaid, by using ready-made Clip Art objects. She wants to make a story about two friends; therefore she needs to invert one of the Ariels. Two boys acknowledge that she is having technical problems and they give suggestions of different solutions and together they solve the problem. The boys leave and the girl continues with her story.

(Participant observation, spring semester, Year 2)

The next excerpt from an interview confirms that the children were aware of the value of helping each other.

Interviewer: What do think you learn from the computers?
Girl: I think I learn to save stuff so…so that I can help others.

(Group interview, spring semester, Year 2)

The next example illustrates how the children from the project classes also helped older children who sometimes visited these classes.

> A boy is helping a fourth grader who is looking for football pictures to copy them into his own document. The boy shows him which keys to press, how to save, and cut and 'save as'.

(Participant observation, fall semester, Year 3)

Since the children could not read the messages in English, it was sometimes evident that the children was not really sure themselves. Nevertheless, they still managed to succeed in helping without asking the teacher, probably by deduction from earlier experiences.

> A girl is helping a boy. She shows him how to open KidPix. A warning sign appears, it is in English. 'I suppose I press OK. Yes it worked'.

(Participant observation, fall semester, Year 1)

Helpfulness was not only expressed as technical advice but also on more complicated matters as when the children were giving *procedural instructions* to one another.

> A girl is writing a text to a digital story. Together with a peer she is watching a picture and at the same time she explains to two boys how to proceed with a math game.

(Participant observation, spring semester, Year 2)

The last aspect of helpfulness differs from the previous ones in the sense that the behavioural outcome was not always successful, namely when trying to give *theoretical instructions or explanations*, which is shown in the next example.

> A boy is struggling with a math problem in his textbook. Several other children appear trying to help and explain subtraction to him. He does not understand.

(Participant observation, spring semester, Year 2)

As shown above the children used to help each other and they tried to do so in all kinds of different situations when they saw a peer in distress, whether they knew how to do it or not. It seems that when it comes to theoretical instruction the children are not always reliable helpers. Nevertheless, the children sometimes are successful in instructing each other even in mathematics.

Interviewer: What do you learn from your friends?
Girl: When there was a difficult page in the math book I learnt from T how to manage. There were a lot of car numbers and you were supposed to find the lowest number and then next and so on. How to put it together into sentences… She taught me how to do that page. Then she's taught me to double things. Because I didn't really understand what to do on the page where you should double numbers and stuff.

(Group interview, spring semester, Year 2)

Besides helping, there was a high degree of active involvement, which constitutes the second sub-theme of SRL behaviour.

3.3.2 Involvement

The observations revealed children's engagement in the work of their peers as well as an active participation in collaborative projects. A typical incident of the children's way of spontaneously showing *engagement in the work of their peers* comes from a situation where a girl was making a digital recording on the computer for her digital story and her peers got involved in the organization of her work.

> In the classroom, five children are working at four computers. One is working with teacher support, one is writing letters, and two others are helping each other to log in. The fifth child is about to audiotape her story. She asks everyone to be silent. The other children are considerately listening and giving her good advice. The recording girl lost her reading fluency at the end of the story and wants to do it all over again. "We can listen to it first", the other children say. Then everyone listens to the recording and comments when she makes a reading error. The teacher and the child discuss whether or not to record it again. The recording girl asks somebody to close the door. A girl rushes willingly to her assistance.

(Participant observation, spring semester, Year 2)

The next excerpt is an example of the children's *active participation in collaborative projects*, where the children had been organized by their teacher to work together in different groups.

> Seven children are making a story together. All children give suggestions for the story, they make drawings and they make clay sculptures. Then they organise the material to make a coherent story. The teachers photograph the material and insert it in a PowerPoint presentation. The manuscript is organised so that each picture got its own text. The children take turn reading the text and digitally record it. The children organise this work themselves and they all take part in the discussion. They create and solve the problem together and the children are particular about that every child's work is represented in the slide show.

(Teacher interview, spring semester, Year 1)

According to the teachers, the children get more interested in each other's work when working on computers than when using paper-and-pencil, as outlined in the excerpt below.

> When working with a story in PowerPoint the children are curious about other children's work and they never ask the teachers what to write. When working in PowerPoint there is need for less teacher support than when they use paper-and-pencil, the children have no trouble in making up a story and are more interested in each other's work.

(Teacher interview, spring semester, Year 2)

As this excerpts show, the self-regulated learning environment was characterised by a high level of interaction both between child and child and between child and teacher. By helping each other the children were active in their own learning process as well as in that of their peers.

In sum, the data has shed light on young students' computer use in an SRL environment, where the many different features have been structured into themes and sub-themes. A summary of these findings is presented in Table 18.2 below.

Table 18.2 Overview of the themes and sub-themes

Metacognition	Motivation	Behaviour
Awareness of the processes in the learning environment	Positive affect	Helpfulness
Active learning processes	Pleasure in making things happen on the screen	Technical support
Individual differences in learning	The many offers of different activities by the computer	Procedural instructions
Strategies for regulation and evaluation	Feelings of emptiness without the computers	Theoretical instructions or explanations
Active searching for knowledge	Negative affect	Involvement
Self-monitoring the working process	Frustration due to technical problems	Engagement in the work of their peers
Complex procedures	Physical fatigue	Active participation in collaborative projects
Sequential procedures	Opportunities for making choices	
Ambivalence	Formulation of new tasks	
Planning	Development of the present tasks	
Having own plans for the day	Work eagerness	
Awareness of the relations between different learning activities in the learning process	Focus on task	

(continued)

Table 18.2 (continued)

Metacognition	Motivation	Behaviour
The letters of the alphabet as the basis for reading	Playfulness at the computer	
The close link between reading and spelling	Teacher demands	
Reflection about layout and readability	Hard work	
No benefit from computer use		
Problems with technological messages		
Problems with the spelling function		
Turning down help from the computer		

4 Discussion

The aim of the present study was to describe and analyze the learning environment in a primary school computer project by exploring how SRL was manifested in an educational setting using computers as an integrated tool in everyday classroom work. The focus was on metacognition, motivation and behaviour, three important aspects of SRL as linked to computer use. As suggested by Zimmerman (2001), the children in this learning environment showed metacognitive awareness and motivation as well as active behavioural involvement in their own learning environment. Metacognitive skills like planning and monitoring, were found when the children were talking to each other about what they do when they learn or when planning (Salonen et al., 2005; Sullivan & Pratt, 1996) as well as when self-monitoring and evaluating their work. This was, for example, shown when the children created digital stories and constructed tables, where they constructed and reconstructed until they were satisfied with the result. Thus, it is probably beneficial for the development of children's metacognitive experiences and metacognitive knowledge that teachers try to encourage the children in primary school to talk to each other when learning and to trust their willingness to take on responsibility for learning, as previously shown by Zhang et al. (2007).

It also seems important that the teachers explicitly instructed the children to produce their own texts and not to re-produce texts from books or web pages. Thus, the children got opportunities to develop their own thoughts, which in turn likely gave rise to many different ideas to discuss. Such co-regulation (Salonen et al., 2005) has been experienced as problematic by teachers who are afraid that undirected learning would lead to "bad learning habits" (Watts & Lloyd, 2004). The two teachers in the present study were, however, not afraid of "bad learning habits" due to letting the children work autonomously, which probably could be explained by the teachers' openness to new teaching methods and rather constructivist epistemological beliefs that previously have been related to the way computers are used in school (Hermans, Tondeur, van Braak, & Valke, 2008).

The children in the present study were however not undirected in the sense that they could do whatever they wanted to. Instead, the self-regulation mainly regarded their liberty to carry out the assignments in different modes, such as using pictures, sounds and texts. For some assignments, the children could also choose if they wanted to use the computers or paper-and-pencil. They could also work in different sequences, in their own pace, as well as in different peer-relation settings. Furthermore, they were free to invent new and developed tasks when the compulsory ones were completed. Nonetheless, some learning problems occurred, like when the children were helping each other with theoretical instructions and explanations. This could give rise to misunderstandings (Salonen et al., 2005) and it is not sure that the teachers had time to notice that when it happened. However, in mathematics the children used workbooks, where the teachers could control their knowledge.

As for metacognitive strategies, the easiness to make changes to one's product when working with a computer seems to give an opportunity to perceive and handle the school work as a process with molar steps rather than tedious procedural ones. The latter may obscure the overall organization of the work at the expense of planning. Before the introduction of computers, the children wrote their stories in notebooks and these paper-and-pencil products were hard to rearrange especially when using coloured crayons and watercolours. As opposed to traditional schoolwork in primary school, different presentation programmes make it easier to rearrange the pictures and the text in ways that better suits the story logically. The use of these programmes might cause metacognitive experiences associated with positive affect (Efklides, 2006, 2008), since the children seem to find it both easy and enjoyable to make reflections in order to make the story connected logically.

Nevertheless, it is important to point out that not all children benefited from computer use. The reasons found in this study were that some of the children had low reading ability, thus, they were not helped by the spelling function and could not understand the warning signs on the screen. Another reason was that some children found it very hard to search for new material to use for their assignments, which also might be linked to low reading ability. The fact that not all children benefit from a non-traditional learning environment has previously been found by Prinsen et al. (2009). Furthemore, Wittwer and Senkweil (2008) showed that only students who used computers in a self-determined way improved their math achievement, whereas Winters et al. (2008) found a relation between SRL-processes and academic success. Further investigations are needed in order to find out who will not benefit and how to handle that situation.

When using computer, there is no need to create the story serially, from a beginning to an end, since the computers allow non-linear modes of working and different alternatives to be tested, as in the case of the boy who made the story about himself (see also Folkesson, 2004; Somekh, 1991; Tondeur et al., 2007). One aspect of computer-supported learning concerns the many opportunities for making choices. This can be related to the development of metacognitive strategies for orientation and evaluation. The children spent a lot of time on layout issues and evaluated different alternatives. This was in accordance with the results from a similar project in England where the teachers stressed the value of the opportunities for the children to experiment with colours and shapes (Somekh, 1991). Since it is so easy to make changes on the computer the children do not need to feel a negative affect even if they are not pleased with their first attempts as suggested by Efklides (2008). Easiness gives the children the energy required to look closely, test, and evaluate in order to choose the best alternative (Somekh, 1991). However, cognitive ability or prior knowledge is important for metacognitive skills. For example, when reading ability is low the children cannot regulate their work on the computer and they need assistance.

Another prerequisite for metacognitive skills to be used seems to be the allowing attitude of help-seeking and help-giving and the encouragement from the teachers. The ways in which the computers are used have previously been shown to depend on teacher attitudes (Alexandersson et al., 2000; Chen & Chang, 2006; Hermans et al., 2008; Tondeur et al., 2007) and experience (Chen & Chang, 2006; Watts & Lloyd, 2004), that is, the level of teacher- or self-regulation. These findings are supported in the present study, where the teachers expressed explicit strategies aiming at the children's production of texts written in their own words. Hence, teachers were encouraging a non-linear and self-regulated way of working.

Another important finding of the present study was the prevalence of motivational evidence found in the data. As previously shown (Alexandersson et al., 2000; Enochsson, 2004; Rosas et al., 2003; Watts & Lloyd, 2004) motivation is closely linked to computer use. A reason for this is that children will engage more in activities they find interesting and important, as suggested by Pintrich and Schrauben (1992), and computer use in school might be perceived as important, because it is a "real thing" used by grown-ups in their professions. Hence, even for small school children, computer use could be very important for their motivation to engage in schoolwork, as shown by Enochsson's (2004) observation of older children's Internet searching, which motivated younger children in learning to read. Moreover, as discussed by Dermizaki et al. (2009), motivation and volition give feedback to appropriate self-concept. Thus, computers might enhance also domain-specific self-concept since they enhance motivation by being perceived as "fun" to work at, but this needs to be studied further in a more controlled design.

The children seem able to be much focused on task when using the computers (Sullivan & Pratt, 1996) and there seems to be a "fun factor" involved, as pointed out by several researchers (Alexandersson et al., 2000; Enochsson, 2004; Folkesson, 2004; Skolverket, 1998b). When the children in the present study were asked why they liked to work on the computers, they often simply answered that it was "fun", without being able to verbalise the reasons for it. One explanation might be the opportunities of making choices, as mentioned above. Another important explanation might be that the computer is a medium where the children can experiment with pictures, sounds, and motions, that is, more senses are involved at the same time than with traditional paper-and-pencil work. Alexandersson et al. (2000) found that the children showed more vivid body language and movement and also that they were even singing while working at the computer as compared to working with paper-and-pencil. Their findings are in line with the present study, when, for example, a little girl made up the story about the two pictures of Ariel as if she was playing with dolls. From these examples it seems that the computer gives rise to positive feelings, which are assumed to enhance motivation (Efklides, 2006, 2008). Moreover, when working on computers there is an immediate response from the computer itself and in this self-regulated environment the children also were involved in each other's work, increasing motivation further, which is in line with the findings of Enochsson (2004).

By using the computers, the children showed clear evidence of being intrinsically motivated. The assignments were given from an outer source, that is, the teacher, but it seems that, as the children found working on the computer so enjoyable, their extrinsic motivation in getting the assignments done was turned into intrinsic motivation (Deci, 1992) when they got the opportunity to solve the problem on the computer. The computer, thus, seems to be a medium that can turn extrinsic motivation into intrinsic. As pointed out by Ryan and Deci (2000), most activities that older children and adults do are not intrinsically motivated. Yet, in early childhood there could still be possibilities for children to be intrinsically motivated if the school environment endorses SRL. The relation between SRL and computer use seems to be reciprocal where one supports the other and vice versa, especially for the motivational and behavioural aspects of SRL. For example, the children's enjoyment at the computers and the opportunity to make choices together with the freedom to collaborate with peers seemed to enhance a positive attitude to work. This reciprocity can be linked to the concept of reciprocal determination (Bandura, 1997), in the sense that the children influence the learning environment and the learning environment influences the children.

A striking observation was that the children so willingly left their own work to help a peer, especially with technical problems, and this could be interpreted as an aspect of the prevalence of co-regulation in this environment. This is contrary to what Colnerud (1999) reported, as she found a prevalent helplessness pattern in relation to computer use in primary school. The helpfulness found in the present study could be related to the concept of co-regulation, as in this classroom the children were allowed and even encouraged to help each other when the teacher did not know how to solve a problem or did not have time. These teachers even used the computers strategically as opposed to the teacher-regulated classrooms studied by Colnerud (1999), where the teachers were afraid of not being in control. Thus, the helpfulness found in an early study (Riis, 1991) was confirmed here. The overall impression from the results is in line with Azevedo (2005) and Schraw (2007), that is, the use of computers can serve as a pedagogical tool regarding self- and co-regulation. An explanation from the teachers in the present study is that the computers were perceived as motivating as well as helpful for the students wanting to develop new ideas.

A specific issue of student differences was the attitude among some of the children who thought that it was wrong to self-regulate learning with computer support. They wanted the teachers to tell them what was right and wrong and for these children the computers caused a negative affect. However they did not abandon their work, probably since there was an extrinsic motivation (Ryan & Deci, 2000) in terms of teacher demand.

The contribution of the present study is theoretical, methodological, and educational. Theoretically, the contribution is that the empirical findings have shed light on the theoretical concepts of SRL with respect to primary school children and their abilities as well as their difficulties. For improving the conceptual understanding of self-regulation and computer environments, Schraw (2007) suggests a development of explicit process models that explain the effect of scaffolding on SRL. Since SRL in primary school is not very frequent in previous research the themes and sub-themes found in this study might be of interest for future research, for example, making comparisons between computerised and traditional school settings regarding SRL.

Methodologically, the contribution is the use of a naturalistic method where the interplay between the significant factors of computer use in a SRL-promoting environment is shown. In this study multiple data collected from an authentic context have been used and this enables more valid conclusion than a single method study would (Zimmerman, 2008). However, a problem with the present study is that the data were collected in a broader framework and were not specifically focused on the features of SRL. Thus, it would be important to collect new data with this specific focus to elaborate the present findings.

The educational implication is that computer use could provide opportunities for SRL, which may be implemented strategically by teachers already in primary school. Situations where teacher help is especially needed are also revealed. Thus, the results are promising and encouraging for teachers who want to use computers for a non-traditional educational situation.

Acknowledgements We would like to thank the anonymous reviewer and especially Professor Anastasia Efklides for their insightful and helpful comments on this chapter. We would also like to thank Professor Stefan Samuelsson, and Assistant Professors Anna-Lena Eriksson-Gustafsson and Ulla-Britt Persson, Linköping University, for their valuable comments on previous versions of this paper.

The order of the authors is arbitrary, equal responsibility is assumed.

5 Appendix A: Guide for the Group Interviews

1. What do you learn in school?
2. Who teaches you that?
3. How do you learn that?
4. What is good about school?
5. What is bad about school?

The children were asked to freely answer the following questions in a narrative way without further prompts from the interviewer.

1. What do you think of when you see a computer?
2. What is it used for?

3. When is the computer good to use?
4. When is the computer bad to use?
5. What did you do when the computers were stolen?
6. What is the difference between now when you have the computers, and before, when they were stolen.

References

Alexandersson, M., Linderoth, J., & Lindö, R. (2000). *Dra den dit och lägg den där! En studie om barns möten med datorn i skolan* [Move it over there and put it there! A study about children's encounters with the computer in school] (IPD-rapporter No. 2000:15, Institutionen för Pedagogik och Didaktik). Göteborg, Sweden: University of Göteborg.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist, 40*, 199–209.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review, 54*, 199–231.

Chen, J-Q., & Chang, C. (2006). Using computers in early childhood classrooms: Teachers' attitudes, skills and practices. *Journal of Early Childhood Research, 4*, 169–188.

Colnerud, G. (1999). *Beyond the rhetorical curtains. Moral messages hidden in current trends*. Paper presented at the 9th International Study Association on Teachers and Teaching (ISATT) Conference, Dublin, Ireland.

Deci, E. L. (1992). On the nature and functions of motivation theories. *Psychological Science, 3*, 167–171.

Denis, B., & Hubert, S. (2001). Collaborative learning in an educational robotics environment. *Computers in Human Behaviour, 17*, 465–480.

Dermizaki, I., Leondari, A., & Goudas, M. (2009). Relations between young students' strategic behaviours, domain-specific self-concept, and performance in a problem-solving situation. *Learning and Instruction, 19*, 144–157.

Dinsmore, D. L., Alexander, P. A., Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychological Review, 20*, 391–409.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040–1048.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Lillington, NC: Psychology Press.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*, 277–287.

Enochsson, A. (2004). Internet seeking and motivation for learning to read. *Learning for Innovation in Technology Education, TERC2004, 1*, 178–188.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34*, 906–911.

Folkesson, A.-M. (2004). *Datorn i det dialogiska klassrummet* [The computer in the dialogical classroom]. Lund, Sweden: Studentlitteratur.

Folkesson, A.-M., & Swalander, L. (2007). Self-regulated learning through writing on computers: Consequences for reading comprehension. *Computers in Human Behavior, 23*, 2488–2508.

Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research, 77*, 334–372.

Hermans, R., Tondeur, J., van Braak, J., & Valke, M. (2008). The impact of primary school teachers' educational beliefs on the classroom use of computers. *Computers and Education, 51*, 1499–1509.

Kusunoki, F., Sugimoto, M., & Hashizume, H. (2002). A group learning support system enhancing the externalization of thinking. *System and Computers in Japan, 33*, 93–100.

Lajoie, S. P. (2008). Metacognition, self regulation, and self-regulated learning: A rose by any other name? *Educational Psychology Review, 20*, 469–475.

Merriam, S. B. (1998). *Case study research in education: A qualitative approach*. San Francisco, CA: Jossey-Bass.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102–116.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT.

Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology, 99*, 83–98.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*, 667–686.

Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (pp. 149–183). Hillsdale, NJ: Erlbaum.

Prinsen, F. R., Volman, M. L. L., Terwel, J., & van den Eeden, P. (2009). Effects on participation of an experimental CSCL-programme to support elaboration: Do all students benefit? *Computers & Education, 52*, 113–125.

Rosas, R., Nussbaum, M., Cumsille, P., Marianov, V., Correa, M., Flores, P., Grau, V., Lagos, F., Lopez, X., Rodriguez, P., & Salinas, M. (2003). Beyond Nintendo: Design and assessment of educational video games for first and second grade students. *Computers & Education, 40*, 71–94.

Riis, U. (1991). *Skolan, och datorn. Satsningen Datorn som pedagogiskt hjälpmedel 1988–1991* [The school and the computer. The realization of the computer as a pedagogical tool 1988–1991] (Tema T Rapport 24). Linköping, Sweden: Linköpings Universitet.

Roschelle, J., & Teasly, S.D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 69–97). Berlin, Germany: Springer.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54– 67.

Ryan, R. M., Kuhl, J., & Deci, E. L. (1997). Nature and autonomy: An organizational view of social and neurobiological aspects of self-regulation in behavior and development. *Development and Psychopathology, 9*, 701–728.

Salonen, P., Vauras, M., & Efklides, A. (2005). Social interaction – What can it tell us about metacognition and coregulation in learning? *European Psychologist, 10*, 199–208.

Salovaara, H., & Järvelä, S. (2003). Students' strategic actions in computer-supported collaborative learning. *Learning Environment Research, 6*, 267–285.

Scardamalia, M., Bereiter, C., McLean, R. S., Swallow, J., & Woodruff, E. (1989). Computer supported intentional learning environments. *Journal of Educational Computing Research, 5*, 51–68.

Schraw, G. (2007). The use of computer-based environment for understanding and improving self-regulated learning. *Metacognition and Learning, 2*, 169–176.

Skolverket. (1998a). *Skolverkets rapport 160, Nationella kvalitetsgranskningar. 1998* [National Agency for Education report 160, National Assessment of Quality of Schools. 1998]. Stockholm: Skolverket.

Skolverket. (1998b). *Verktyg som förändrar. En rapport om 48 skolors arbete med IT i undervisningen* [Tools that make a difference. A report from 48 schools on using IT in education]. Stockholm: Skolverket.

Somekh, B. (1991). Pupil autonomy in learning with microcomputers. *Cambridge Journal of Education, 21*, 47–65.

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.

Sullivan, N., & Pratt, E. (1996). A comparative study of two ESL writing environments: A computer-assisted classroom and a traditional oral classroom. *System, 29*, 491–501.

Tondeur, J., van Braak, J., & Valke, M. (2007). Towards a typology of computer use in primary education. *Journal of Computer Assisted Learning, 23*, 197–206.

Waite, S. (2004) Tools for the job: A report of two surveys of information and communications technology training and use for literacy in primary schools in the West of England. *Journal of Computer Assisted Learning, 20*, 11–20.

Watts, M., & Lloyd, C., (2004). The use of innovative ICT in the active pursuit of literacy. *Journal of Computer Assisted Learning, 20*, 50–58.

Whitebread, D., Coltman, P., Pino Pasternak, D., Sangster, C., Grau, V., Bingham, S., Almeqdad, Q., & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning. *Metacognition and Learning, 4*, 63–85.

Wilkinson, S. (2003). Focus group. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods* (pp. 184–204). London: Sage.

Winne, P. H. (2004). Comments on motivation in real-life, dynamic, and interactive learning environments. *European Psychologist, 9*, 257–263.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In P. Pintrich, M. Boekaerts, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 532–565). Orlando, FL: Academic.

Winters, F. I., Greene, J. A., & Costich, C. M. (2008). Self-regulation of learning within computer-based learning environments: A critical analysis. *Educational Psychological Review, 20*, 429–444.

Wittwer, J., & Senkweil, M. (2008). Is students' computer use at home related to their mathematical performance at school? *Computers and Education, 50*, 1558–1571.

Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of 9- and 10-year-olds. *Education Technology Research Development, 55*, 117–145.

Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology, 11*, 307–313.

Zimmerman, B. J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed., pp. 1–37). Mahwah, NJ: Erlbaum.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*, 166–183.

Chapter 19
University Teachers Engaged in Critical Self-Regulation: How May They Influence Their Students?

Kathryn Bartimote-Aufflick, Angela Brew, and Mary Ainley

1 Introduction

In this chapter we present a model of the reflective processes that underpin continuing professional development of university teachers. The teachers under consideration here are employed in academic positions that involve research, teaching, and service. At the outset it is acknowledged that university teachers are not a homogeneous group and that individuals' teaching experience and expertise vary widely; as do teaching situations – from small tutorial groups to classes of hundreds or thousands.

Professional development is about intentional engagement in change, that is, transformation or conversion. To change in this way is to take on, to expand, and to let go, of knowledge and/or habits. In this work we pursue professional learning as "intentional conceptual change" (Sinatra & Pintrich, 2003) initiated and consciously controlled by the learner. Associated with this type of thinking are constructs such as metacognition and self-regulation.

In addressing how university teachers may continue to gain discipline and pedagogical knowledge we propose a four-phase model of *critical self-regulation* (CSR) that draws on the student learning literatures of metacognition and self-regulated learning (SRL), and critical reflection from adult education. To Zimmerman's (2004) three-phase model of SRL, we add a prior stage that includes teachers' reflection on the basic premises of their instruction and consideration of higher-order instructional goals. At the appraisal end of the process, the evaluation phase of SRL is extended to incorporate critical (or premise) reflection. We argue that critical reflection (as described by Mezirow (1990, 1991)) provides a qualitatively different and a deeper reflection than the reflection referred to in existing metacognition and SRL models.

K. Bartimote-Aufflick (✉)
Institute for Teaching and Learning, University of Sydney, Carslaw F07, Sydney, NSW, 2006, Australia
e-mail: k.aufflick@usyd.edu.au

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_19,

Following the presentation of the model of CSR, situations and tools for developing CSR are considered, including how teachers may learn individually and together. We focus on learning that arises because of the perceived need by the teacher to address some learning or teaching dilemma. The formulation of a problem or question in response to such a dilemma, and subsequent learning task, corresponds to the personal concerns and intent of the teacher. Learning through inquiry is emphasised as a means of dealing with the problem, once formulated. We discuss the need for a mentor or 'other' to support teachers' inquiry learning.

2 Theoretical Background

Metacognition may be conceptualised as two components, namely *knowledge* about cognition (strategies available and suitable for the task) and *regulation* of cognition (Baker & Cerro, 2000). Paris (2002) states that "Flavell (1979) described metacognition in terms of person, task, and strategy knowledge whereas Brown (1978) interpreted metacognition in terms of processes such as planning, monitoring, and regulating" (p. 106). Thus, at an early stage, the term metacognition encompassed both knowledge and control processes.

Pintrich (2002) described three types of metacognitive knowledge following Flavell's (1979) model. (a) *Strategic knowledge* is knowledge of general strategies for learning (rehearsal, elaboration, and organisation), metacognitive strategies (planning, monitoring, and regulation), and problem solving and thinking (e.g., means-end analysis, working backwards, induction, deduction). (b) *Task knowledge* includes knowledge of the difficulty of a task, and some conditional knowledge about when and why to use particular strategies. This conditional knowledge is said to include situational knowledge, such as social aspects, conventions, and cultural norms. (c) *Self knowledge* includes knowledge of one's strengths and weaknesses, and (against the norm for cognitive models) motivational beliefs (self-efficacy, goal orientation, interest, and task value).

Emerging from research into metacognition has been the construct of SRL, and an examination of factors outside the usual boundaries of metacognition such as motivation, affect, and attribution (Baker & Cerro, 2000). In addition, the traditional cognitive boundaries of metacognition have recently been widened to include consideration of the role of affect in metacognitive experiences and the accuracy of metacognitive knowledge (e.g., Efklides, 2006).

According to Pintrich (2000) SRL "is an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behaviour, guided and constrained by their goals and the contextual features in the environment" (p. 453). Zimmerman (2000, 2004) describes three cyclical phases of self-regulation: forethought (goal setting), control (including monitoring), and self-reflection. More recently, Zimmerman (2008) has further emphasised the proactive qualities required by the learner to display initiative, perseverance, and adaptive skill.

As useful as the student learning constructs of metacognition and SRL are to the context of university teacher learning, they do not fully cater for the complex thinking and transformation required in this work, particularly at the initial and closing stages of the cycle. Kreber (2004) asserts that SRL, as currently characterised in the literature, may achieve effective learning for teachers without achieving meaningful outcomes. That is, the learning task or goal achieved successfully by employing SRL might not have addressed the "right" problem. Therefore, a new phase before the SRL forethought phase is needed to cater for the difficulty involved in assessing the assumptions that influence the problem (or learning task) formulation process for a university teacher. Further we argue that the self-reflection phase, as defined by Zimmerman (2000, 2004), needs to be expanded so that not only is the relative success of the outcome in relation to the goal assessed, but also the utility of the outcome in terms of the how well the formulated problem and subsequent task or goal addressed the underlying issue of concern. Indeed, Mezirow (1991) following the lead of Dewey conceptualised learning as solving problems or dealing with "problematic situations" (Ozmon & Craver, 1990), and examining the relevance or value (*premise*) of the problem/question itself.

We propose that the complex thinking required for this new pre-phase and the expanded self-reflection phase may be facilitated through critical reflection (also known as premise reflection). Critical reflection provides a way by which one may undertake transformative learning, that is, learning in which one learns about and challenges the assumptions underlying perspectives and habits, and acts on these insights (Mezirow, 1990). It follows an a posteriori epistemology in the sense that knowing is through one's experience.

Critical reflection also relies on constructivist principles in that learning is viewed as the making of knowledge and meaning in an active purposeful way, and by acknowledging a learner's existing knowledge and seeking to build on, elaborate, or correct this knowledge. Commonly in this type of reflection cognitive approaches are used to make changes to one's thinking if one's critique of a presupposition leads to an untenable position.

3 Model of Critical Self-Regulation

The model of CSR aims to describe potential learning of adults, such as university academics, as in our opinion premise reflection requires a developmentally mature learner. We propose that a model of CSR may be useful for developing all three types of university teacher knowledge: (a) *discipline knowledge*, which is knowing the ways in which concepts and principles are organised to incorporate facts, and the ways in which validity or invalidity are established in one's discipline – equivalent to Shulman's (1986) content knowledge; (b) *pedagogical knowledge*, which is defined here as general knowledge of how students learn and of teaching strategies – combining Kreber and Cranton's (2000) instructional and pedagogical domains; and (c) *applied pedagogical knowledge*, which is knowledge of ways of representing and formulating the discipline that make it comprehensible to others – comparable to Shulman's (1986) pedagogical knowledge – and of a suite of teaching and learning activities that help students grow in their own understanding of the discipline.

In this chapter we focus on CSR as a way of developing pedagogical and applied pedagogical knowledge primarily. However we do realise that the CSR model may also be useful in thinking about the development of an academic's disciplinary knowledge, as research may be thought of as inquiry learning that deals with scientific and societal dilemmas. Further, conscious and intentional reflective processes can make discipline-based knowledge and skills explicit to teachers and therefore potentially available for sharing with students.

There are common threads amongst the three conceptualisations of learning discussed so far, namely metacognition, SRL, and critical reflection. One thread is a focus on conscious cognition. We note that Efklides (2008) describes a model of metacognition that incorporates both unconscious and conscious levels of cognition and regulation – the model of CSR operates at what Efklides calls the "personal awareness level" (p. 283). Another intersection point is reflection. In the model of expert learning developed by Ertmer and Newby (1996), metacognition is facilitated by reflection, and acts as a vehicle to activate and coordinate metacognitive knowledge and self-regulation. However, reflection takes different forms in the various phases of SRL. In forethought it is reflection that serves in analysing the task, in performance it is reflection to create self-awareness during the completion of the task, and in the self-evaluation phase it is reflection to consider and evaluate achievement.

In CSR the type of reflection possible within Zimmerman's (2000, 2004) SRL model is extended to include premise reflection. This extension is in both directions – at the beginning before the forethought phase, and at the end after the self-reflection phase. Specifically, CSR facilitates a deep questioning of one's perceptions of a teaching and learning problematic situation or dilemma and thus leads to a comprehensive consideration of the nature of the task, the enacting of self-regulatory processes in the completion of it, and a thorough analysis of the importance and utility of the resultant outcome. Figure 19.1 depicts CSR as a cycle.

3.1 Prior Phase

The prior phase in university teacher learning involves consideration of *higher-order goals*, for oneself and one's students. Higher-order goals include the pursuit of the ought (should be) self (Higgins, 1987), the ideal (would like to be) self (Higgins, 1987), or the possible self (Powers, 1973). More specific goals may be derived from higher-order goals, or may form as a reaction to one's environment. Goal setting and the definition of the nature of the task by the learner are very important as they determine the strategies employed in completing the task, the product of learning, and one's evaluation of it.

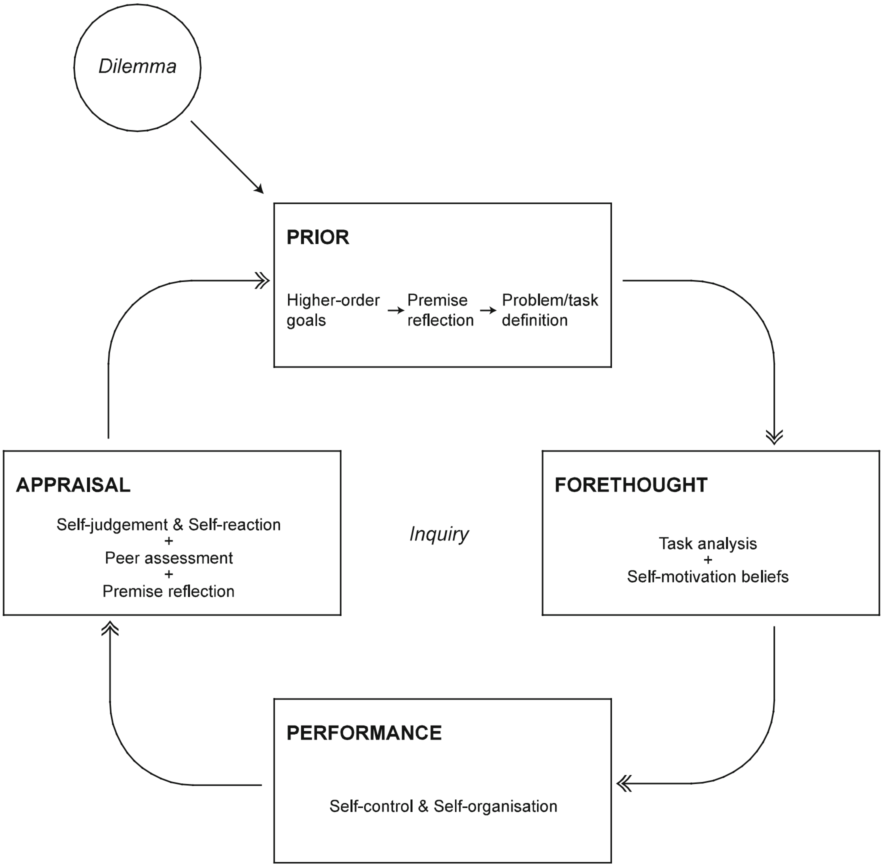


Fig. 19.1 Model of critical self-regulation

If a dilemma arises that creates disharmony between a teacher's perception of a situation and their higher-order goals for that situation, then a rationalisation for this vexation needs to take place. Reflection at this stage may result in one of two outcomes for the teacher involved – continue as things are, or change (McAlpine, Weston, Berthiaume, Fairbank-Roch, & Owen, 2004). If change is the choice, then we propose that premise reflection ensues.

Premise reflection is the reviewing and challenging of one's current views, knowledge, and beliefs – about oneself as a teacher and learner, and about one's students. From this altered meaning perspective, one composes problems differently than if this step had been overlooked. Once a situation has been problematised, the learning task becomes apparent, that is, to attempt to solve or address this newly formulated problem.

3.2 Forethought Phase

This phase begins with a description of the problem, which is referred to as *content reflection* (Mezirow, 1991) in the critical reflection model, and *task analysis* in the SRL model. McAlpine et al. (2004) refer to this phase as reflection-for-action. This leads to the usual SRL elements of *goal setting*, *strategic planning*, and also a consideration of *self-motivation beliefs*. Self-motivation beliefs incorporate motivation constructs that are well known: self-efficacy, which is belief about one's capability to learn or perform to attain goals (Bandura, 1997); outcome expectations, which are the expected consequences of achieving a particular goal – be they benefits or liabilities; interest; and goal orientation, which is the overall purpose for learning, namely mastery or comparing one's performance with others'.

In the CSR model we endorse an inquiry or research-led (Brew, 2006; Zamorski, 2002) approach to tackling the problem or question formulated in response to the dilemma experienced by the teacher. This approach influences the nature of the planning (hypothesis setting, choosing evidence collection methods, and analysis method choice).

3.3 Performance Phase

This phase proceeds as in the SRL model. In this phase it is the *awareness* of various types of knowledge (e.g., about oneself, learning strategies, regulatory strategies, assessment methods) made apparent via reflection, that is essential in informing understanding about teaching and learning. This phase is inclusive of Schön's (1995) reflection-in-action, and Mezirow's (1991) process reflection (the strategies and procedures of problem solving).

Self-control and *self-observation* are given by Zimmerman (2000, 2004) as two major types of performance or volitional control. Volition is that which "controls intentions and impulses so that action occurs" (Corno, 2001, p. 194). Self-control is promoted through self-instruction (verbalizing how to proceed during task execution), imagery (forming of mental pictures), attention focusing (e.g., structuring one's environment to eliminate distractions, or avoiding rumination about past mistakes), and task strategies (reducing a task to its essential parts, and organising those parts).

Zimmerman (2000) describes the features of effective self-observation as timely and highly informative self-feedback, accurate observations, and an emphasis on recording positive accomplishments rather than negative behaviours. However, more recently, Zimmerman (2004) instead refers to metacognitive self-monitoring and self-recording as the processes of self-observation – the former being a mental tracking of one's performance and the latter keeping physical records of performance.

3.4 Appraisal Phase

Appraisal has been used to name the final phase rather than Zimmerman's "self-reflection" as we felt using the word reflection here may cause confusion given the emphasis on reflection (and its various types) throughout each of the CSR model phases.

From the final phase of the SRL model we include in the CSR appraisal phase *self-judgement* and *self-reaction*. Self-judgment involves evaluating one's performance and determining and attributing causes for these performance outcomes. In self-evaluating a person compares self-monitored information with one of three types of criteria: mastery, previous performance (or self-criteria), or normative (social comparisons). Zimmerman (2004) asserts that proactive self-regulators self-evaluate using either self-improvement or mastery criteria as these provide feedback regarding the effectiveness of their learning processes. Attributing the cause of an outcome to an uncontrollable cause (e.g., fixed ability) leads to negative self-reactions and discourages further efforts to improve (Weiner, 1979), whereas causal attribution of outcomes to personally controllable sources, such as use of strategies, sustains learners' motivation and encourages further adaptation of strategies (Weiner, 2000; Zimmerman & Kitsantas, 1997, 1999).

At this stage one must also deal with self-reactions (as per the SRL model). Two key forms of self-reactions that Zimmerman (2004) outlines are self-satisfaction regarding one's performance which also depends on the intrinsic value placed on the task, and adaptive or defensive inferences. Adaptive inferences consider how to alter one's self-regulatory approach in subsequent efforts. Defensive inferences involve protecting oneself from future dissatisfaction leading, for example, to helplessness, procrastination, task avoidance, cognitive disengagement, apathy, or self-handicapping. The self-judgment and self-reactions elements of this final phase may be considered analogous to Schön's (1995) reflection-on-action. Boud, Keogh, and Walker (1985) describe this form of reflection as "reconstructing experience" in that a learner describes their experience, works through attitudes and emotions, and orders and makes sense of new ideas and information attained.

Boud (2000) asserts that lifelong learners are also "lifelong assessors" (p. 152). The development of the ability to monitor one's own learning and standards is enhanced by the development of devices and methods for self-monitoring and judging progression towards goals. Another fundamental in the success of lifelong learning is how well an individual can engage peer and mentor assessment of work as s/he progresses. Therefore, in addition to consideration of the outcome of learning by the university teacher individually, the appraisal phase of CSR involves the review of outcomes, in relation to the problem and task definition, by another person. We refer to this as *peer assessment*.

Following on from this is the possibility for *premise reflection* in that one may mull over whether or not the "right" problem or question was posed in the first place (in the prior phase). This premise reflection may potentially initiate another completion of the CSR cycle.

4 Situations and Tools for Developing Critical Self-Regulation

Moving away from theory to practice, we now consider findings from research on metacognition and SRL (areas conceptually related to CSR) that may give useful indications for conditions under which university teacher CSR may be fostered. Teacher CSR in its turn should promote student SRL. In what follows we start with situations that promote student SRL and then go on to contexts that are related to teacher learning and may support teacher CSR and, indirectly, student SRL.

4.1 Insights from Student Learning Research

Following a review of the research literature on student SRL, we highlight five conditions that may be relevant to the fostering of student SRL: (a) interaction with ‘an other’ in metacognition and SRL where the other may be a teacher or a fellow student; (b) explicit instruction of skills and knowledge; (c) embedded instruction; (d) inquiry work including project work; and (e) self and peer critique.

4.1.1 The Importance of Others in Learning

Students may not be mindful of, or not be able to explain their own strategic behaviour (Paris & Cunningham, 1996). A challenge in teaching SRL is to help students to *perceive* opportunities for the use of strategies (Randi & Corno, 2000), and also design situations that require their use. Randi and Corno (2000) champion the use of *cognitive apprenticeship* as an approach to instruction. In this approach, skills are made initially explicit through modelling, then the students acquire these skills themselves through scaffolded exercises, until finally support fades almost completely and students are left to utilise these skills independently (cf. a similar suggestion by Schunk & Ertmer, 2000). The idea of the zone of proximal development (ZPD; Vygotsky, 1978; see also Salomon, Globerson, & Guterman, 1989) is useful here in that the teacher guides the student to a level of development they could not have achieved alone. The work by Brown (1994) on reciprocal teaching highlights the possibility for that guidance to come from a student peer.

4.1.2 Explicit Instruction

Cornford (2002) states that “…it is an unwarranted assumption that learners automatically know how best to learn. The most sensible approach is not to assume the automatic development of learning skills but to teach them quite explicitly” (p. 361). However he notes that many aspects of self-knowledge would in fact be impossible to teach directly but that “good teachers… can organize learning situations so that individuals are forced to consider their own personal strengths and weaknesses, reflect on these, and learn from these experiences" (p. 360).

Boekaerts and Corno (2005) note that direct instruction of academic strategies are important for students of all ability levels, and also mention the rise of socio-culturally-based interventions such as cognitive apprenticeships through modelling, reciprocal teaching, computer-mediated learning environments, and collaborative learning.

Zohar and Ben David (2008) describe a continuum along which explicit teaching may fall – from transmission to guided discovery to discovery learning – and make the point that instruction should be designed to promote knowledge construction by students (i.e., methods away from the transmission and rote learning end of the spectrum). Zohar and Ben David (2008) recommend explicit instruction of metacognitive knowledge that is mediated by verbal discussion, has opportunities for practise across time, and includes individual student-teacher interactions. Key recommendations from Zohar's (2006, p. 339) work are that students need to "use a thinking strategy rather than addressing it in an abstract way", have "the same thinking skills addressed over and over again through different parts of the curriculum", and then practise generalising (thinking metacognitively) across these experiences to learn to recognise general thinking skills.

4.1.3 Inquiry Learning

Randi and Corno (2000) report on several strategy instruction initiatives. One of these is student- or learner-centred science classrooms (Blumenfeld et al., 1991) where students complete projects distinguished by two components: (a) they have a question or problem that organises and drives related activities, and (b) the activities result in the creation of artifacts or products that address the driving question. In these projects, goals are inherent in the project design, and strategy instruction helps students solve the problem and create an artifact representing the solution. There is also self-monitoring and feedback during the projects and the created products are shared and critiqued. Also in a science education setting, Dean and Kuhn (2007) endorse discovery learning and the development of inquiry skill, and time on task. They demonstrate that inquiry learning is preferable to direct instruction via worked examples.

4.1.4 Embedded Instruction

Many echo the call of Pintrich (2002) for the teaching of metacognitive knowledge to be embedded within the usual content-driven lessons, and for it to become part of the discourse of the classroom. Further, the modelling of strategies, accompanied by an explanation for them, is endorsed. Randi and Corno (2000) are also very firm in their assertion that self-regulation should be taught within the context of the regular curriculum. That is, as embedded instruction rather than in generic strategy training.

The integration of metacognitive and affective skills into a business economics program reported by Masui and De Corte (2005) is an example of embedded instruction for university students.

4.1.5 Self- and Peer-Critique

Paris and Paris (2001) advocate giving students opportunities to practise self-appraisal and self-management. In helping students to develop the ability to monitor their own learning and standards, we should be working with students to develop devices or methods for self-monitoring and judging progression towards goals (Boud, 1995, 2000). Boud (1995) states that peer-assessment and self-assessment are avenues through which self-regulation skills, such as, goal setting, monitoring/adjusting, and reflecting/evaluating, may be developed.

4.2 Insights from Teacher Learning Research

University teachers can take advantage of a range of learning avenues, such as discussing teaching approaches with experienced tutors in their teaching team and peers at academic conferences; observing capable students and fellow colleagues and speaking with them about their learning strategies; attending professional development workshops; completing formal postgraduate study in education; reading peer-reviewed publications on relevant topics; or engaging in the scholarship of teaching and learning (see Hutchings & Shulman, 1999; Kreber, 2005a). They may also learn through writing a review paper; keeping a reflective journal on learning about teaching; or seeking guidance from mentors (Kreber, 2005b; Kreber & Cranton, 2000; Randi, 2004).

Typical academic staff development activities at Australian universities include: completion of workshops (stand-alone and in series) and certificate courses; participation in committees and working parties; facilitation of workshops and courses; collaboration on teaching development projects; collaboration in educational research; and dissemination of development and research work. In addition, university teachers may also enhance their understanding of teaching practice through interpreting and summarising student, peer and quality assurance agency evaluations or through mentoring (Bartimote-Aufflick & Smith, 2008).

Although the value of all of these academic development activities is recognised, here, we wish to identify and pursue those settings which are most likely to support CSR in particular. Therefore in this section we have only summarised principles and key findings from teacher learning research in the related areas of critical reflection, metacognition and SRL.

As teaching SRL arises in the coming section as a method for learning SRL oneself, we also provide some insight into teacher characteristics that may help or hinder the teaching of SRL.

4.2.1 Embedded Instruction and Inquiry Learning

As for the student learning literature, we see in the teacher learning literature that embedded instruction and inquiry learning are important in shaping and developing metacognitive knowledge and regulatory skills in teachers. We discuss each in turn.

Within a university teacher learning space, whether this is in the form of a course, a project, or a less formal mentoring setting, the equivalent of embedded instruction is important. That is, that learning about SRL needs to be centred around authentic and useful activities for the teacher. This may include discussion of teaching dilemmas defined as conceptual conflicts to be resolved (Tillema & Kremer-Hayon, 2002), discussion of cases of student learning (Lin, Schwartz, & Hatano, 2005; Zohar, 2006), talking through critical (and/or recurrent) events (Lin et al., 2005), or attempting to teach SRL in one's own classes (Randi, 2004; Zohar, 2006). Critically, each of these activities requires some examination of their effectiveness within the CSR model.

Inquiry learning or collaborative projects that investigate and seek to influence student SRL seem an excellent way of providing self-directed learning opportunities for teachers (Van Eekelen, Boshuizen, & Vermunt, 2005). Such projects put control (Kreber, 2000) into the hands of the learning teacher.

Taking a collaborative innovation approach to research, Randi and Corno (2000) worked with and mentored teachers involved in projects investigating the learning of their students in a flexible way to help them improve their own SRL about their teaching. Another example of this approach is provided by Butler, Novak Lauscher, Jarvis-Selinger, and Beckingham (2004), who engaged teachers in a collaborative inquiry for professional development by gathering a small number of school teachers ($n=10$), primarily in learning assistant roles outside the classroom, to introduce SRL to students. Across the 2-year project, the authors found that teachers increasingly displayed aspects of SRL in their teaching. That is, they did actively reflect on their teaching, they did gain new conceptual understanding, and shifted their practice adaptively via a cyclical approach. For example, they set goals for practice, enacted changes, evaluated outcomes, and adapted approaches.

4.2.2 Teaching SRL as a Way of Teachers Learning SRL

Intertwined in the evidence (above), that embedded instruction and inquiry learning professional development initiatives can promote teacher SRL, is an indication that teachers can learn SRL personally through the process of *teaching* SRL to their students.

For students to be taught metacognitive or self-regulatory strategies and skills, teachers need to be willing and able to teach them. It is a definite possibility that they themselves have never been formally instructed in cognitive and metacognitive strategy use, and may feel uncertain about their ability to teach these skills (Cornford, 2002). As Zohar puts it "…teachers cannot teach effectively what they do not know" (2006, p. 335). It may be that the university teacher needs an 'other' to initiate this learning, provide a space for it, and lend their leadership and expertise.

4.2.3 The Need for 'The Other' in University Teachers' Learning

The importance of discourse (dialogue devoted to assessing contested beliefs) in premise reflection is highlighted by Marsick and Mezirow (2002). Samuels and Betts (2007) suggest that a challenge is often needed in order to prompt deeper reflection – of the deconstructing and reconstructing type they endorse. This type of reflection is analytical rather than merely descriptive, and in Samuels and Betts' (2007) study this analysis was evident through participants revisiting previously written reflections, taking more responsibility for the reflective process themselves, and displaying metacognitive awareness. A cue to deeper reflection may take the form of dialogue, of modelling, or of written prompts. In the appraisal phase of the CSR model the engagement of mentors and peers (Boud et al., 1985) is highlighted as an important dimension in judgement. However we see possibility for interaction with others to be of benefit throughout each of the phases of CSR. Efklides (2008) highlights the important role that others play in individual metacognition (at the meta-meta or social level)-we can draw on the cognition and emotions of others (and their regulation) through interaction.

One could argue that every learner needs a mentor or a learning companion (Cranton & Wright, 2008). For students it is the teacher, but for academic staff in their continued learning about teaching who is that learning companion? Butler et al. (2004) mention that participants in a collaborative development of teaching setting found strength and inspiration in the presence of a teaching mentor or leader. If we turn to a research analogy (McAlpine, Weston, & Beauchamp, 2002), perhaps we need to think about different roles as one grows in teaching (and filling different roles in varying contexts and times). A university teacher may progress in their expertise and subsequent role in the teaching community over time – from mentee to peer, then peer to mentor, and mentor to leader.

The mentor or leader in learning about CSR and related constructs may be a colleague whose specific role is to guide and support the growth of university teacher expertise (equivalently an academic developer in the UK and Australia, a faculty developer in North America, or an educational developer in Europe), or a senior colleague with an understanding of educational psychology in practice, or the two working together to combine their capabilities. For the academic developer Weston and McAlpine (2001) state three (staged) roles: providing a venue for people to learn, reframing activities suggested by academic colleagues, and working collegially with academic colleagues on a project to improve teaching or student learning.

4.2.4 Learning (and Using) SRL Across Multiple Domains

Along with developing pedagogical knowledge regarding SRL, and knowledge of how to teach SRL within the context of one's discipline (see also Weston & McAlpine, 2001; Zohar, 2006), some researchers promote the usefulness of teachers improving their SRL within their own discipline or content area (Kramarski, 2008).

This resonates with our earlier supposition that the CSR model may be useful in developing all three types of university teacher knowledge – discipline, pedagogical, and applied pedagogical.

Complementary to teaching the specifics of SRL, Zohar (2006) found that being able to articulate cognitive activities, question students, and facilitate metacognitive discussions are important skills to possess when teaching SRL. Here we see the potential for expert learners, that is, university academics successful in learning via research and inquiry in their own discipline, to begin to make their own learning strategies and skills explicit to themselves via reflection, and then articulate these to their students. Of further interest is the possibility for existing skills and strategies used in gaining discipline knowledge to also be employed in learning about teaching – this involves the much-debated idea of transfer.

4.2.5 Teachers' Approaches to Teaching and Their Own Learning

Important, also, is the nature of one's beliefs about learning and instruction (Zohar, 2006). One's approach to teaching, for example, a traditional transmission-of-knowledge approach to instruction versus a constructivist approach, will affect how possible it is for a teacher to teach thinking skills which by nature require students to actively learn (Zohar, 2004). This echoes the work in higher education on student-centred and teacher-centred approaches to teaching (see Prosser & Trigwell, 2006; Trigwell, Prosser, & Ginns, 2005; Trigwell, Prosser, & Waterhouse, 1999). Postareff and Lindblom-Ylänne (2008) rather than considering approaches to teaching to be dichotomous, have described the teacher- (or content-) centred approach as "less complete" (p. 120) than the student-centred approach.

Besides teachers' approaches to teaching, there are teacher characteristics and beliefs that may be related to their teaching and CSR. Gordon, Dembo, and Hocevar (2007) investigated the relationship between pre-service school teachers' achievement goal orientations (mastery versus performance) and their classroom control ideology. They borrow from Hoy's (2001) descriptions of custodial and humanistic control ideologies. Custodial control involves behaviour control and order maintenance, whereas humanistic control emphasises cooperative interaction and experience. Gordon et al. (2007) found that teachers with better SRL were more likely to take a mastery approach to their own learning, and a more humanistic classroom control ideology. Their conclusion was that teachers tend to teach in a similar way to the way they learn. Perhaps we can reason from this as university teachers are introduced to SRL they should also be introduced to goal orientations and the particular affordances of a mastery orientation (for themselves and their students).

In a study involving student teachers Malmberg (2008) found that mastery goal orientation was positively related to reflective thinking. This leads us to place further emphasis on the importance of goal orientation (that is, having a mastery orientation) for the advancement of the CSR phases and in particular the influence goal orientation plays in the forethought phase when goals and plans are being set.

5 Conclusions and Further Questions

In the present chapter we have introduced a model for university teacher learning called critical self-regulation. This model involves four phases: the prior, forethought, performance, and appraisal phases. It builds on the research into metacognition and SRL but extends this for the adult learner to incorporate premise reflection.

To make some recommendations as to how CSR may be developed by university teachers, we have reviewed the student learning and teacher learning literature around the development of SRL, as we anticipate parallels. From this review we conclude that the development of CSR may be encouraged when teachers are involved in meaningful and authentic tasks that are of direct relevance to their teaching work (equivalent to embedded instruction).

We suggest that participation in inquiry projects alongside peers and mentor(s) will be a key vehicle for professional development of CSR. The ‘other’ in teacher learning is important. A mentor or leader in learning about teaching and student learning can offer a cognitive apprenticeship, facilitate a space for dialogue with peers, and provide explicit instruction regarding CSR and related constructs. Peer teachers may give guidance through collaboration, reciprocal teaching and peer critique.

An important aspect of professional development for university teachers is having the opportunity to practice generalising (thinking metacognitively) across experiences (research, teaching, and service) to recognise general thinking skills that may be applied in a variety of settings.

An avenue for developing CSR that arose in the review of the literature is actually teaching SRL. We noted, however, that a teacher’s ability to teach SRL may be limited by their self knowledge, capacity to generalise across the breadth of their academic work, approach to teaching, and goal orientation.

In a sophisticated professional development initiative, all of these elements may work in concert, for instance, collaborative inquiry into teaching students SRL, with an instructional or workshop component to the project.

5.1 Emergent Research Questions

The first question is ‘What professional learning situations are related to demonstrable change in the CSR of teachers?’ This is related to another, ‘What contexts encourage and equip university teachers to teach SRL explicitly and successfully to their students?’ In our review of the literature we note that little is known of the impact of SRL of teachers on student learning. McAlpine et al. (2004) and Zohar (2006) also indicate a lack of research in this area. McAlpine, Berdugo Oviedo, and Emrick (2008) outline how difficult it is to demonstrate links between the impact of academic development initiatives and student learning. “It involves tracking impact on lecturer thinking, through impact of that thinking on their teaching actions to impact on students they work with, and ideally comparison with a control group" (McAlpine et al., 2008, p. 671).

As we have already explained we conceive of three domains of knowledge for the university teacher – their discipline domain in which they conduct research, the pedagogical domain, and the applied pedagogical domain. It may be that successful development of CSR will focus on all three of these domains rather than just the two pedagogical areas – in the least these links and opportunities for transfer of knowledge should be examined through further research.

A hook for engaging university teachers in this learning about SRL is their desire for their students to learn thinking skills. This is noticed in the animated conversation regarding attributes of university graduates (see Barrie, 2007, for an overview) – many of these desired attributes echo skills and knowledge of self-regulated and metacognitive learners. Given this apparent interest by university teachers for their students to have these skills, they may be willing to teach them explicitly. And as we have all experienced in our own disciplines, we come to know something more deeply when wrestling with how best to present it to someone else.

Acknowledgements This chapter draws on a presentation given at the EARLI Metacognition SIG Conference in Ioannina, Greece (Bartimote-Aufflick & Brew, 2008).

Thanks to the reviewer for the feedback, and to Anastasia Efklides for helpful comments on the draft.

Kathryn Bartimote-Aufflick is Lecturer at the Institute for Teaching and Learning, the University of Sydney, Australia. Angela Brew is Professorial Fellow at the Learning and Teaching Centre, Macquarie University, Sydney, Australia. Mary Ainley is Associate Professor at the Department of Psychology, the University of Melbourne, Australia.

References

Baker, L., & Cerro, L. C. (2000). Assessing metacognition in children and adults. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 99–145). Lincoln, NE: Buros Institute of Mental Measurements.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Barrie, S. C. (2007). A conceptual framework for the teaching and learning of generic graduate attributes. *Studies in Higher Education, 32*, 439–458.

Bartimote-Aufflick, K., & Brew, A. (2008, May). University teachers' understanding of their own learning, *3rd Meeting of EARLI SIG 16: Metacognition*. Ioannina, Greece.

Bartimote-Aufflick, K., & Smith, L. (2008, July). Developing as academics through participation in learning communities, *Higher Education Research and Development Society of Australasia (HERDSA) Conference*. Rotorua, New Zealand.

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist, 26*, 369–398.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review, 54*, 199–231.

Boud, D. (1995). *Enhancing learning through self assessment*. London: Kogan Page.

Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education, 22*, 151–167.

Boud, D., Keogh, R., & Walker, D. (1985). What is reflection in learning? In D. Boud, R. Keogh, & D. Walker (Eds.), *Reflection: Turning experience into learning* (pp. 7–17). London: Kogan Page.

Brew, A. (2006). *Research and teaching: Beyond the divide*. Basingstoke, UK: Palgrave Macmillan.

Brown, A. L. (1978). Knowing when, where and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 77–165). Hillsdale, NJ: Erlbaum.

Brown, A. L. (1994). The advancement of learning. *Educational Researcher, 23*, 4–12.

Butler, D. L., Novak Lauscher, H., Jarvis-Selinger, S., & Beckingham, B. (2004). Collaboration and self-regulation in teachers' professional development. *Teaching and Teacher Education, 20*, 435–455.

Cornford, I. R. (2002). Learning-to-learn strategies as a basis for effective lifelong learning. *International Journal of Lifelong Education, 21*, 357–368.

Corno, L. (2001). Volitional aspects of self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed., pp. 179–212). Mahwah, NJ: Erlbaum.

Cranton, P., & Wright, B. (2008). The transformative educator as learning companion. *Journal of Transformative Education, 6*, 33–47.

Dean, D., & Kuhn, D. (2007). Direct instruction vs. discovery: The long view. *Science Education, 91*, 384–397.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*, 277–287.

Ertmer, P. A., & Newby, T. J. (1996). The expert learner: Strategic, self-regulated, and reflective. *Instructional Science, 24*, 1–24.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911.

Gordon, S. C., Dembo, M. H., & Hocevar, D. (2007). Do teachers' own learning behaviors influence their classroom goal orientation and control ideology? *Teaching and Teacher Education, 23*, 36–46.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*, 319–340.

Hoy, W. K. (2001). The pupil control studies. A historical, theoretical and empirical analysis. *Journal of Educational Administration, 39*, 424–441.

Hutchings, P., & Shulman, L. S. (1999). The scholarship of teaching: New elaborations, new developments. *Change, 31*, 10–15.

Kramarski, B. (2008). Promoting teachers' algebraic reasoning and self-regulation with metacognitive guidance. *Metacognition and Learning, 3*, 83–99.

Kreber, C. (2000). Becoming an expert university teacher: A self-directed process. In H. B. Long (Ed.), *Practice and theory in self-directed learning* (pp. 131–142). Schaumburg, IL: Motorola University Press.

Kreber, C. (2004). An analysis of two models of reflection and their implications for educational development. *International Journal for Academic Development, 9*, 29–49.

Kreber, C. (2005a). Charting a critical course on the scholarship of university teaching movement. *Studies in Higher Education, 30*, 389–405.

Kreber, C. (2005b). Reflection on teaching and the scholarship of teaching: Focus on science instructors. *Higher Education, 50*, 323–359.

Kreber, C., & Cranton, P. A. (2000). Exploring the scholarship of teaching. *Journal of Higher Education, 71*, 476–495.

Lin, X., Schwartz, D. L., & Hatano, G. (2005). Toward teachers' adaptive metacognition. *Educational Psychologist, 40*, 245–255.
Malmberg, L.-E. (2008). Student teachers' achievement goal orientations during teacher studies: Antecedents, correlates and outcomes. *Learning and Instruction, 18*, 438–452.
Marsick, V., & Mezirow, J. (2002). New work on transformative learning, ID Number: 10876, *Teachers College Record*: http://www.tcrecord.org.
Masui, C., & De Corte, E. (2005). Learning to reflect and to attribute constructively as basic components of self-regulated learning. *British Journal of Educational Psychology, 75*, 351–372.
McAlpine, L., Berdugo Oviedo, G., & Emrick, A. (2008). Telling the second half of the story: Linking academic development to student experience of learning. *Assessment and Evaluation in Higher Education, 33*, 661–673.
McAlpine, L., Weston, C., & Beauchamp, C. (2002). Debriefing interview and colloquium: How effective are these as research strategies? *Instructional Science, 30*, 403–432.
McAlpine, L., Weston, C., Berthiaume, D., Fairbank-Roch, G., & Owen, M. (2004). Reflection on teaching: Types and goals of reflection. *Educational Research and Evaluation, 10*, 337–363.
Mezirow, J. (1990). Preface. In J. Mezirow (Ed.), *Fostering critical reflection in adulthood* (pp. xiii–xxi). San Francisco: Jossey-Bass.
Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco: Jossey-Bass.
Ozmon, H., & Craver, S. (1990). *Philosophical foundations of education* (4th ed.). New York: Macmillan.
Paris, S. G. (2002). When is metacognition helpful, debilitating, or benign? In P. Chambres, M. Izaute, & P.-J. Marescaux (Eds.), *Metacognition: Process, function and use* (pp. 105–120). Boston: Kluwer.
Paris, S. G., & Cunningham, A. (1996). Children becoming students. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 117–147). New York: Macmillan.
Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist, 36*, 89–101.
Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego, CA: Academic.
Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice, 41*, 219–225.
Postareff, L., & Lindblom-Ylänne, S. (2008). Variation in teachers' description of teaching: Broadening the understanding of teaching in higher education. *Learning and Instruction, 18*, 109–120.
Powers, W. T. (1973). *Behavior: The control of perception*. Chicago: Aldine.
Prosser, M., & Trigwell, K. (2006). Confrimatory factor analysis of the approaches to teaching inventory. *British Journal of Educational Psychology, 76*, 405–419.
Randi, J. (2004). Teachers as self-regulated learners. *Teachers College Record, 106*, 1825–1853.
Randi, J., & Corno, L. (2000). Teacher innovations in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 651–686). San Diego, CA: Academic.
Salomon, G., Globerson, T., & Guterman, E. (1989). The computer as a zone of proximal development: Internalizing reading-related metacognitions from a reading partner. *Journal of Educational Psychology, 81*, 620–627.
Samuels, M., & Betts, J. (2007). Crossing the threshold from description to deconstruction and reconstruction: Using self-assessment to deepen reflection. *Reflective Practice, 8*, 269–283.
Schön, D. A. (1995). *The reflective practitioner: How professionals think in action* (New ed.). Aldershot, UK: Arena, Ashgate Publishing.
Schunk, D. H., & Ertmer, P. A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 631–650). San Diego, CA: Academic.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*, 4–14.
Sinatra, G. M., & Pintrich, P. R. (2003). The role of intentions in conceptual change learning. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change* (pp. 1–18). Mahwah, NJ: Erlbaum.
Tillema, H. H., & Kremer-Hayon, L. (2002). "Practising what we preach" – Teacher educators' dilemmas in promoting self-regulated learning: A cross case comparison. *Teaching and Teacher Education, 18*, 593–607.
Trigwell, K., Prosser, M., & Ginns, P. (2005). Phenomenographic pedagogy and a revised approaches to teaching inventory. *Higher Education Research and Development, 24*, 349–360.
Trigwell, K., Prosser, M., & Waterhouse, F. (1999). Relations between teachers' approaches to teaching and students' approaches to learning. *Higher Education, 37*, 57–70.
Van Eekelen, I. M., Boshuizen, H. P. A., & Vermunt, J. D. (2005). Self-regulation in higher education teacher learning. *Higher Education, 50*, 447–471.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press (Original work published 1930).
Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology, 71*, 3–25.
Weiner, B. (2000). Intrapersonal and interpersonal theories of motivation from an attributional perspective. *Educational Psychology Review, 12*, 1–14.
Weston, C. B., & McAlpine, L. (2001). Making explicit the development toward the scholarship of teaching and learning. *New Directions for Teaching and Learning, 86*, 89–97.
Zamorski, B. (2002). Research-led teaching and learning in higher education: A case. *Teaching in Higher Education, 7*, 411–427.
Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–40). San Diego, CA: Academic.
Zimmerman, B. J. (2004). Sociocultural influence and students' development of academic self-regulation: A social-cognitive perspective. In D. M. McInerney & S. Van Etten (Eds.), *Research on sociocultural influences on motivation and learning: Vol. 4. Big theories revisited*. Greenwich, CT: Information Age.
Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*, 166–183.
Zimmerman, B. J., & Kitsantas, A. (1997). Developmental phases in self-regulation: Shifting from process goals to outcome goals. *Journal of Educational Psychology, 89*, 29–36.
Zimmerman, B. J., & Kitsantas, A. (1999). Acquiring writing revision skill: Shifting from process to outcome self-regulatory goals. *Journal of Educational Psychology, 91*, 241–250.
Zohar, A. (2004). Elements of teachers' pedagogical knowledge regarding instruction of higher order thinking. *Journal of Science Teacher Education, 15*, 293–312.
Zohar, A. (2006). The nature and development of teachers' metastrategic knowledge in the context of teaching higher order thinking. *Journal of the Learning Sciences, 15*, 331–377.
Zohar, A., & Ben David, A. (2008). Explicit teaching of meta-strategic knowledge in authentic classroom situations. *Metacognition and Learning, 3*, 59–82.

Chapter 20
Metacognitive Knowledge of Decision-Making: An Explorative Study

Barbara Colombo, Paola Iannello, and Alessandro Antonietti

1 Introduction

Research on human decision-making has shown that people's choices in various domains are affected by several biases, which lead them to make grave mistakes or, less dramatically, to make decisions that are far from the best ones. Such biases are produced by deceptive tendencies in the representation of the information about the available options (Arkes & Blumer, 1985); by cognitive (Tversky & Kahneman, 1973) and affective heuristics[1] (Slovic, Finucane, Peters, & MacGregor, 2002); by emotional states (Van den Bergh & Dewitte, 2006); by failures to activate inhibitory control processes needed in order one to be able to assess the adequacy of the impulsive or intuitive judgments and response tendencies and, if the case, to counteract them (Kahneman, 2003). The problem is that people not only decide in sub-optimal ways, but also that they are usually confident in their decisions and do not suspect the existence of these possibly misleading biases.

1.1 Decision-Making Processes

Psychological research on decision making support the view that people have limited, and sometimes inadequate, awareness concerning the way they decide. For instance, people's account of their judgments preceding a decision is often modified by the subsequent knowledge of what happened after the decision. The so-called hindsight bias (Fischhoff & Beyth, 1975) reflects the fact that the

[1] The "affect heuristic" is the tendency to derive judgments from the emotional reactions to stimuli. Subjective impressions of goodness and badness act as a heuristic, which is a source of fast, perceptual evaluations.

B. Colombo (✉)
Department of Psychology, Catholic University of the Sacred Heart,
Largo Gemelli, 1 20123, Milano, Italy
e-mail: barbara.colombo@unicatt.it

A. Efklides and P. Misailidi (eds.), *Trends and Prospects in Metacognition Research*,
DOI 10.1007/978-1-4419-6546-2_20, © Springer Science+Business Media, LLC 2010

retrospective reconsideration of previous mental operations is not reliable. In the same line, one can mention the outcome bias (Baron & Hershey, 1988). Also in this case the retrospective judgment of a decision is distorted by knowing the effect of the decision; if the decision is followed by the desired outcome, the decision is rated as good; if the outcome is negative, the decision is judged negatively. In both cases, however, the quality of the decisional process is the same, since, given some starting information, that decision was the one that had to be made if all data were taken into account in a proper way. In other words, persons seem to fail to distinguish between the adequacy of their process of deciding and the desirability of the effect following the decision. Both the hindsight and the outcome bias testify a lack of awareness concerning what the individual actually did in order to make a decision.

Furthermore, many findings in psychological research on decision making are counterintuitive. For instance, one might maintain that, in the absence of time pressure or impellent needs, analysing the features of the given options carefully and for a long time is the best way to go in order to make a decision. On the contrary, experiments have shown that in some circumstances (e.g., when the number of features to be evaluated is high) deciding without reflecting is the most appropriate approach (Dijksterhuis, 2004; Dijksterhuis, Bos, Nordgren, & van Baaren, 2006). Another example: in lay people's opinion, having a wide range of options is preferable to being constrained to choose among a small number of options. By contrast, research shows that, if the number of options is too large, decision-making is impaired (Iyengar, Wells, & Schwartz, 2006).

These remarks concerning the fact that people lack awareness of the decision making process and share wrong beliefs about it stress the need that they develop relevant metacognitive competences which can counteract the pernicious tendencies that drive people's behaviour. In fact, one can assume that better choices would follow from the increase of the level of awareness of one's own mental processes and from the acquisition of more adequate knowledge about such processes. Both awareness and knowledge of the psychological mechanisms underlying decisions are manifestations of metacognition and contribute to the enhancement of control over one's mental activity, and consequently over actual behaviour (Dinsmore, Alexander, & Loughlin, 2008).

There is a second consideration that supports the relevance of metacognition to decision making. Decisions sometimes involve taking into consideration other people's ways of deciding. Tasks such as the Ultimatum Game (Güth, Schmittberger, & Schwarze, 1982) demonstrate this case. Specifically, you are asked to split a sum of money with a partner, who can either accept or refuse the offer you made to him/her; if s/he accepts, you both keep the money (split according to your offer); otherwise you both lose the money. This task requires the person to make assumptions about how the partner's mind works, that is, "Why should the partner accept an unfair offer?", "Are persons opportunists, so that they will accept even a small part of the money since they think that less money is better than no money?", or "Do people care for their social reputation, so as to refuse unfair offers in order to prevent others to suspect that they need money at all costs?"

Also, the Beauty Contest game (Nagel, 1995) implicates folk conceptions about human decision-making. In this game participants are asked to simultaneously pick a number between 0 and 100; the winner will be the person whose number is closest to ½ the average of all numbers picked by the participants. Participants should reason as follows: the average of all numbers picked should be around 50; so the winning answer should be "25"; but, if all the other players follow the same reasoning, all will pick "25" and thus the winning response should be "12.5". This reasoning should be repeated recursively ad infinitum. When will people stop reasoning in this way? Naïve psychological assumptions about the limits of human recursive thinking are crucial to decide what will be the most frequent answer given by the others. These naïve assumptions are in essence theories people have about (recursive) thinking and therefore constitute metacognitive knowledge about thinking.

The aforementioned cases suggest that there might be metacognitive reasons for sub-optimal decision-making: if beliefs about the mental processes involved in choices do not correspond to the actual factors influencing them, persons are induced to approach decisions with wrong ideas in their mind and, consequently, with inappropriate attitudes. These remarks stress the need to take into account people's folk ideas about what occurs in the human mind when a decision has to be made – a piece of metacognitive knowledge which merits to be investigated and, if inadequate, to be considered in order to devise tools and procedures to modify it.

On the basis of these considerations, it seems important to analyse the metacognitive aspects of decision making. To apply proper strategies to control the decision-making processes, both an adequate awareness of one's own past, current and future mental operations and an accurate metacognitive knowledge of the usual and ideal ways in which people make decisions seem to be needed.

1.2 Metacognitive Knowledge

The term "metacognitive knowledge" is widely used to refer to beliefs about cognitive processes as well as persons, tasks, strategies, and goals (Flavell, 1979). It is well recognised that such beliefs – as well as metacognitive experiences concerning the feeling of familiarity, difficulty, confidence, and satisfaction (Efklides, 2002) – are coherently organised so as to constitute systemic conceptions, as shown in the case of problem solving (Antonietti, Ignazi, & Perego, 2000). Metacognitive knowledge refers not only to beliefs about one's own mental activity, but also to extra-personal cognitive processes, in general. Several studies have questioned whether these two aspects (i.e., the personal and the extra-personal) of the awareness of mental functioning and of metacognitive knowledge can be considered as one. For example, Vesonders and Voss (1985) and Jamenson, Nelson, Leonesio, and Narens (1993) showed that most people make the same use of metacognitive information, whether that information was self-relevant or general. In the same vein, Nelson, Kruglanski, and Jost (1998) argued that the general cognitive processes that drive one's sense of self-knowledge are fundamentally similar to the processes of deriving a sense of other people's knowledge. Likewise, in defining the concept of metacognition, Kluwe (1982) maintained that "the thinking subject has some knowledge about his own thinking and that of other persons" (p. 202). More recently, Efklides (2008) noted that metacognition is not only an individual phenomenon, but it develops in collaborative contexts and it is dynamically co-constructed by different agents. Finally, we must not neglect the fact that metacognition includes also an affective dimension (Efklides, 2006) and it is influenced by and influences emotions and motivation (Efklides, 2001).

Unfortunately, as yet little is known about metacognition in decision making. This chapter provides preliminary data on the issues already pinpointed. More precisely, the present study aimed to investigate the awareness people have about their decision-making processes in their working life, their conceptions of the decision-making processes, and how these aspects of metacognition are related to their profession, their level of professional expertise, and their personal decision-making style.

Metacognitive knowledge is analysed here in its various aspects. According to Flavell (1981), metacognitive knowledge encompasses beliefs about personal attributes and skills, task features and strategies. All these aspects were investigated by making reference to decision making. Furthermore, we were interested in studying both the personal and general aspects of metacognition, that is, what people believe about their own and other persons' mental processes, both by making reference to specific decisions and by considering the overall experience. Also, the affective dimension (feelings accompanying a decision and emotions following it) were taken into account. Finally, the beliefs about one's self-efficacy (Coutinho, 2008) in making decisions were also considered.

1.3 Aims: Hypotheses

The aims of the present study were the following: (a) Explore people's metacognitive knowledge about the way they make decisions, by analysing their awareness of type and number of decisions they make at workplace, their main characteristics as decision-makers, the processes involved in their decision-making, and type of strategies used – analytic vs. intuitive, global vs. specific, etc. (b) Describe the metacognitive knowledge shared by individuals, that is, their conceptions of decision making and of the strategies required for decision making. (c) Identify the possible associations between participants' professional field and their level of expertise with both awareness of their own way of making decisions and metacognitive knowledge about decision making. (d) Explore the relationships between participants' decision-making style (intuitive vs. analytical) and both metacognitive awareness of their own mode of decision making and conceptions of decision making.

It was predicted that people would be able to report various aspects of their metacognitive knowledge about the self as decision-maker (Hypothesis 1) and that profession, level of expertise, and decision-making style would affect the various aspects of metacognitive knowledge, both about the self and the decision-making processes (Hypothesis 2).

With respect to profession, it was expected that the decision-making processes are differentiated among professional categories, not because of the "technical" elements that are specific to each profession, since different jobs require different competences in decision making (Brown, James, & Mills, 2006; Iannello, 2007), but because of the different kinds of decisions that different professions require; for example, economists are used to making quick, substantial, and notable decisions; on the other hand, teachers and artists may have more time to reflect and be strategic about their decision, which usually are also less irrevocable. Therefore, the various professions were expected to be associated with different kinds of decisions and decision-making processes (Hypothesis 2a).

With respect to level of expertise, it was expected that people with more years of working life (high expertise) would be differentiated from those with few years (low expertise) mainly in the types of decisions they make and in the perceived effectiveness of their decisions and the regret expressed for mistaken decisions (Hypothesis 2b).

Finally, it was expected that associations would be found between decision-making style and both metacognitive awareness of one's personal way of deciding and conceptions of decision making (Hypothesis 2c).

2 Method

2.1 Participants

Participants in the study were 85 adults (35 men and 50 women), whose age ranged from 24 to 75 years (*M*=39.61, *SD*=13.18). Participants practised different types of professions (medical doctors, nurses, school teachers, psychologists, community educators, managers, economists, traders, workmen, clerks, drivers, artists, unemployed housewives and university students) and had distinct levels of expertise[2] which were quantified in terms of years participants practised their profession (*M*=14.23, *SD*=11.02). Students were included in the sample taking into consideration that attending a university is an occupation – analogous to a proper job – which requires cognitive and metacognitive competencies (with reference to decision-making processes) comparable with other professional categories (Bubany, Krieshok, Black, & McKay, 2008; Pietrzak, Duncan, & Korcuska, 2008). The selection of

[2] See Sects. 3 and 4 for exact categorisations of the level of expertise.

different subsamples was decided in order to ensure that the objectives of the study would be investigated in people who differed with respect to age, occupation, and level of expertise, thus allowing comparisons between them to be made.

To analyse the associations between professions and people's awareness and beliefs concerning the decision-making process, participants were collapsed into eight categories: (1) Medical professions: doctors and nurses ($n=6$); (2) School teachers ($n=18$); (3) Support professions: psychologists and educators ($n=10$); (4) Students ($n=13$); (5) Economists and related professions ($n=19$); (6) Routine professions: workmen and clerks ($n=5$); (7) Artists: musicians and illustrators ($n=11$); (8) Unemployed housewives ($n=2$).

2.2 Materials

2.2.1 The Solomon Questionnaire

The Solomon Questionnaire was devised for the needs of the present study. It aimed at investigating people's conceptions of decision making at the workplace, the beliefs about their decision-making processes and their decision-making competence. The instrument is composed of three sections fully reported in Appendices A and B.

In the first section, labeled Direct Metacognitive Awareness, and comprising four questions with close- or open-ended items each, participants were requested to describe the kinds of decisions they make in their working life and the way in which they are making these decisions. The aim of this section was twofold. First, to gain a picture of prototypic decisions people make at the workplace (descriptive-behavioural aspect of metacognitive awareness/metacognitive knowledge of task). Second, to assess how aware people are of their decision-making processes and to describe personal affective reactions and cognitive strategies used when facing decisional problems, that is, the procedural (metacognitive knowledge of strategies) and emotional aspect of their metacognitive knowledge of the self. Also, to assess metacognitive knowledge of person, that is, the general experience of the respondent as decision-maker and memory of specific decisions the person has made.

The second section, labeled Indirect Metacognitive Awareness, comprised two multiple-choice questions. It required participants to imagine themselves as decision-makers. To do that, participants were asked to choose among different analogies (which correspond to distinct decision-making approaches). In essence, this section also taps metacognitive knowledge of the self as decision-maker; however, it is investigated in an indirect way by means of analogies, with the aim of integrating these data with those gained by direct and explicit questions, as those included in the first section of the Solomon Questionnaire (Lustie, 1998; Sillman, 1999). The analogies used in this section were built and tested in a previous study

(Colombo & Iannello, 2010), which showed their potential to measure indirect metacognitive awareness of the self as decision-maker.

In the third section, labeled Metacognitive Knowledge of Processes, which comprised four close- and open-ended questions, conceptions of the decision-making processes were investigated. Participants were asked to identify which features are peculiar to good decision-makers and to describe how people can become good at making decisions.

2.2.2 Preference for Intuition and Deliberation Scale (PID)

The PID scale was developed to assess preferences in making decisions intuitively or deliberatively (Betsch, 2004). It assesses individual inclinations towards intuitive decision-making (based on affective reactions towards the decision options) and deliberate decision-making (based on evaluations and reasons). The PID scale is made up of two subscales, the Preference for Intuition (PID-I) and the Preference for Deliberation (PID-D) subscales. Nine items assess the PID-D subscale and nine items assess the PID-I subscale. Responses are given on a 5-point Likert-type scale, ranging from 1 (minimum) to 5 (maximum). Examples of items of the two subscales are “Before making decisions I first think them through” (PID-D) and “I listen carefully to my deepest feelings” (PID-I).

In the present study the Italian version of the PID scale was employed. The data collected in an Italian sample (Iannello, 2008) supported the structure of the original instrument (Betsch, 2004). Moreover, both subscales of the PID scale showed acceptable reliabilities (Cronbach’s $\alpha = 0.73$ and $\alpha = 0.78$ for the PID-I and PID-D, respectively). The pattern of inter-correlations between the two subscales showed a negative relationship although it did not reach a significant level, as previously found by Richetin, Perugini, Adjali, and Hurling (2007).

2.3 Procedure

The Solomon Questionnaire and the PID scale were put online by using a website providing a direct link to the survey and collecting the data in a specific database anonymising the logfiles generated by participants’ computers. The first contact mail addressed to possible participants described briefly the research and asked them to join the study. To recruit participants, we chose among a database of people who during the last years took part in other research projects promoted by the Department of Psychology of our university and gave their e-mail address in order to be contacted again for future studies. We chose persons who had not participated before in studies associated in any way to the topic of the present study. Afterwards the direct link to the research tools was provided to allow participants to fill them in.

2.4 Scoring

Content analysis was used to analyse the answers given by participants to the open questions 4, 7, 8, and 10 of the Solomon Questionnaire. All answers were initially read by two judges, who, through content analysis and recursive categorisations, agreed on a number of basic categories. Responses were then re-examined by each judge and assigned to each of the different categories. The total agreement between the two judges was 98%. For each question Cohen's kappa was computed (Question 4a: $k=0.72$; Question 4b: $k=0.78$; Question 4c: $k=0.73$; Question 4d: $k=0.80$; Question 4e: $k=0.79$; Question 4f: $k=0.79$; Question 4g: $k=0.80$; Question 7: $k=0.81$; Question 8: $k=0.83$; Question 10: $k=0.82$). Ambiguous cases were evaluated by the judges, until agreement was reached. The questions and respective categories are given in Table 20.1.

2.5 Data Analyses

In the analyses, the focus was on the level of expertise variable rather than on both age and level of expertise, because the correlation between the two variables was very high ($r=0.80$).

Next, the participants who showed an intuitive or analytical decision-making style based on their scores on the PID were identified. According to the scoring procedure of the instrument, after calculating the difference between the intuitive and deliberative PID scales, only those participants who obtained an absolute difference value higher than 5 were identified as intuitive or analytical (e.g., PID-difference = −10: the participant was classified as analytical; PID-difference = 6: the individual was classified as intuitive; but PID-difference = −3: the participant was classified neither as analytical nor as intuitive). This procedure led to the selection of 29 participants with an analytical decision-making style (age: $M=35.31$ years, $SD=14.50$; level of expertise: $M=11.00$ years, $SD=11.02$) and five showing an intuitive decision-making style (age: $M=41.20$ years, $SD=13.70$; level of expertise: $M=17.20$ years, $SD=11.45$). Age and the level of expertise were not significantly different in the two subgroups; for age, $t(32)=-0.99$, $p>0.05$, and for level of expertise, $t(32)=-1.14$, $p>0.05$, respectively.

Then the association between the individual decision-making style and the decisions the participants reported making in their work, the awareness of their own decision making, and the conceptions of decision making was investigated. The most interesting data which emerged from the analyses are reported and discussed below. It is worth noting that the numbers of participants vary in different analyses because some respondents failed to answer specific questions.

Table 20.1 Categories of responses to the open ended questions of the Solomon Questionnaire

Question	Categories
4a. Describe the general situation, that is, the context in which you are requested to make this specific decision	Organizing something
	Choosing what to do
	Choosing how to do something
	Strategic planning
	Facing an emergency
	Evaluating something
4b. Which is your first thought?	Analysis and evaluation of the situation
	What to do as a first thing
	Nothing (passive attitude)
4c. How do you feel when you make this kind of decision?	Calm
	Worried
	Stressed
	Involved
	Sad
	Unsatisfied
	Ineffective
4d. What do you do to make this decision?	Evaluate different options
	Confrontation with others
	Apply a strategy
	Refrain from acting
	Intuition
4e. Do you face the situation by yourself or do you ask others for help/advice?	By myself
	Ask others
	It depends
	First by myself – then I ask others
	First I ask others – then I decide by myself
4f. Do you basically employ solutions that turned out to be effective in the past, or do you tend to try out new solutions?	Effective in the past
	Try out new ones
	It depends
	Half and half
4g. Once you have made the decision, do you follow it or do you modify it (entirely or partly)?	Stick to the decision
	Modify in progress – with the help of others
	Modify after reflection – with the help of others
	Modify in progress – alone
	Modify after reflection – alone
	It depends

(continued)

Table 20.1 (continued)

Question	Categories
7. In your opinion, which qualities characterise those people who are effective in making their decisions?	Self-confidence Bravery Experience Intuition Foresight Balance Intelligence
8. A good decision-maker is someone who never regrets his/her decision?Why?	Mistakes are possible Mistakes as learning experience You cannot make mistakes
10. If you believe that the competence of making good decisions can be learned or improved, how do you think a person can become a good decision-maker?	Experience Training Good teachers Observation Bravery Metacognitive awareness Cannot be learned

3 Results

3.1 *Associations Between Direct Awareness of Decision Making and Profession*

First, the data from the Direct Metacognitive Awareness section of the Solomon Questionnaire that aimed at investigating the metacognitive knowledge of the self as regards the task (i.e., number and type of workplace decisions), strategies (employed in specific decisions), emotions (i.e., feelings accompanying one's decisions), and the processes involved in decision making were analysed.

3.1.1 Number of Decisions and Profession

In Question 1 economists ($M=15.76$, $SD=1.08$) and medical professionals ($M=11.67$, $SD=2.89$) reported making more decisions than the other professions; school teachers ($M=5.07$, $SD=2.56$), routine workers ($M=4.75$, $SD=5.19$), and housewives ($M=3.00$; $SD=1.32$) were the ones who reported making less decisions. Even though the differences were not significant, as shown by the ANOVA with profession as between subjects factor, $F(7, 62)=1.87$, $p=0.09$, the tendency outlined suggests a direction in the data, that is, more "quiet" and routine professions require people to make smaller number of decisions.

3.1.2 Number of Decisions About One's Own Self and Profession

A strong difference emerged with respect to the number of decisions people reported making about themselves (Question 2), $F(7, 73)=6.56$, $p<0.001$, partial $\eta^2=0.39$. Post hoc LSD tests showed that students reported significantly more decisions of this kind ($M=74.92$, $SD=16.71$), differing from medicals ($M=20.83$, $SD=23.75$), teachers ($M=47.50$, $SD=30.88$), support professionals ($M=27.10$, $SD=20.97$), economists ($M=27.89$, $SD=24.11$), and housewives ($M=7.50$, $SD=3.53$). Artists ($M=56.00$, $SD=22.71$) and routine workers ($M=53.00$, $SD=35.64$) occupied an intermediate position. These findings appear to be consistent with the kind of profession, namely students are required to manage their own learning, personal and occupational life and, hence, are expected to make decisions about themselves. The same is true especially for artists (who were the ones to score highest after students) who are often managers of themselves. Routine workers' quite high number of self-related decisions seems unusual. Maybe they tend to feel their work as disconnected from the other links of the work chain and, hence, they report making decisions mostly about themselves rather than about their work.

3.1.3 Number of Regretted Decisions and Profession

In so far as regretted decisions are concerned (Question 3), differences among the professions emerged, $F(7, 76)=2.53$, $p<0.05$, partial $\eta^2=0.19$. Students ($M=35.00$, $SD=17.32$) and professionals practising support jobs ($M=23.00$, $SD=14.18$) were those who reported more regretted decisions, although the LSD post hoc tests showed that only the subsample of students differed significantly from the other subsamples. Understandably, students tend to regret more their decisions, because they are still learning and trying different strategies, which presumably are not always optimal. Their working condition is probably experienced as a "gym" and even the failures, and the subsequent regret, are part of the learning experience. On the other hand, those who work as supporters to other people likely experience more regret since taking care of another person requires paying attention to occasional undesired consequences of one's own actions. Yet, it is interesting that medicals reported regretting fewer decisions ($M=11.67$, $SD=9.83$); this finding can be ascribed to the medical training, which leads medicals to develop specific procedures ("sure procedures") that prevent the probability of human mistakes in decision making. They are also aware that making a mistaken decision is a possibility associated with their job, but this is within certain limits, thus decreasing regret.

3.1.4 Difficult Decisions and Profession

As regards the description of difficult decisions (Question 4), differences emerged only when people were asked to report what they do to make decisions (Question

4d; see Table 20.2), Pearson's contingency coefficient $C(N=84)=0.64$, $p<0.01$. Other aspects concerning the context in which participants are asked to take a decision appeared not to differ among subsamples, thus highlighting how the metacognitive control aspects are to be influenced by professional differences more than the awareness aspects. As Table 20.2 shows, economists, medical professionals, and teachers reported being more strategic than the other professions. Artists showed a higher tendency towards evaluating different options than the other professions; furthermore, they were also the only subsample who conceived intuition as a good way of making a decision. People practicing a supporting job appeared to be the ones to select more often the confrontation with others, presumably because of the team-working characterising their profession. Finally, teachers were the ones refraining from acting, probably because this is helpful in their profession.

3.1.5 Number of Modified Decisions and Profession

Professions appeared to also differ with respect to the extent people reported they modify to their decision (Question 4g; see Table 20.3), Pearson's contingency coefficient $C(N=84)=0.70$, $p<0.001$. As can be seen in Table 20.3, economists, teachers, and students had a similar pattern of behaviour. They tended to equally stick to their decisions and to modify them while in progress and by themselves. This can be explained considering that such professions often require them to work alone. Artists also work often alone, but they appear to be more disposed to modifying their decisions, both in progress and after reflection. The difference between these two subgroups can be explained by considering the actual time at their disposal (artists on the average have more time to experiment with their work) and the consequences of the decisions made (which often are harder for professionals of the first group). People who practice support professions tended to modify their decisions alone and after reflection. This attitude can be explained by the nature of their work, that is, psychologists and educators have to analyse what the persons they are assisting are saying or doing to calibrate their reactions accordingly.

3.2 Associations Between Decision Making and Expertise

To consider the effect of different degrees of expertise on decision making, the sample was split into four categories reflecting different levels of professional expertise (1 = low, i.e., from 1 to 7 years of professional experience; 2 = medium-low, i.e., from 8 to 12 years; 3 = medium-high, i.e., from 13 to 19 years; 4 = high, from 20 to 44 years).

Table 20.2 Frequencies of responses to the Question 4d "What do you do to make this decision?" as a function of profession

Profession	Evaluate different options	Confrontation with others	Apply a strategy	Refrain from acting	Intuition
Medicals	1	0	5	0	0
Teachers	3	2	7	4	0
Supporting jobs	2	3	5	0	0
Students	3	1	3	1	0
Economists	1	1	14	2	0
Routine jobs	2	0	3	0	0
Artists	5	0	3	1	2
Housewives	1	1	0	0	0

Table 20.3 Frequencies of responses to the Question 4g "Once you have made the decision, do you follow it or do you modify it (entirely or partly)?" as a function of profession

Profession	Stick to it	Modify in progress – with the help of others	Modify after reflection – with the help of others	Modify in progress – alone	Modify after reflection – alone	It depends
Medicals	1	2	1	1	1	0
Teachers	7	3	0	8	0	0
Supporting jobs	0	2	1	3	4	0
Students	4	0	0	4	0	0
Economists	6	3	2	5	2	0
Routine jobs	1	0	2	2	0	0
Artists	1	0	1	4	3	2
Housewives	0	0	0	0	1	1

3.2.1 Direct Metacognitive Awareness and Expertise

In the case of the Direct Metacognitive Awareness section of the Solomon Questionnaire, people with low expertise appeared to differ from the other three categories in the number of regretted decisions (Question 3), $F(3, 81) = 3.28$, $p < 0.05$, partial $\eta^2 = 0.11$. Post-hoc LSD tests showed that, in terms of number of regretted decisions, people with low expertise reported higher regret ($M = 26.81$, $SD = 17.52$) than the other three categories. In more detail, the difference was bigger between those with low expertise and those who had a medium-low expertise ($M = 11.83$, $SD = 14.38$) and tended to decrease in the other two categories (for medium-high experience: $M = 17.93$, $SD = 19.37$; for high experience: $M = 17.33$, $SD = 12.05$). It can be argued that, after a first phase during which learning a new job is linked to more mistakes and insecurity (and hence regret), people go through a second phase in which the practicing of new knowledge and skills brings more self-confidence, and hence less regret. More experience, as a third phase, brings in more accurate metacognitive awareness and more objective self-evaluation. The latter phase is highlighted by an increase of regret that continues in the high expertise category – showing that people become more aware not only of their competences but also of their weaknesses – and are able to regret mistakes they now fully recognise and understand.

With respect to the hard decision-making situations reported by participants (Question 4a), expertise was found to influence the typologies of cited situations (see Table 20.4), Pearson's contingency coefficient $C(N = 85) = 0.51$, $p < 0.05$. As shown in Table 20.4, the lowest category of expertise was associated with more situations concerning simple organisation tasks (such as deciding an appropriate timetable, which is an activity mostly described as related to routine days), whereas higher categories of expertise were associated with situations involving strategic planning and concrete operations ("how to do" something). This difference may depend on the assignments associated with different levels of expertise, that is, people with low expertise are faced with simpler and more routine tasks than people from the other three categories.

It is worth noting that no other associations of expertise with the aspects of direct metacognitive awareness were found.

3.2.2 Indirect Metacognitive Awareness and Expertise

In the case of the Indirect Metacognitive Awareness section of the Solomon Questionnaire, in which people were asked to use analogies to describe their preferred modality for decision making, significant differences emerged in the choice of the animal (Question 5; see Table 20.5), Pearson's contingency coefficient, $C(N = 85) = 0.58$, $p < 0.05$. As Table 20.5 shows, people with low expertise tended to see themselves almost exclusively as sheep, whereas those with high expertise chose the lion and the dog. The former, when asked to make a decision, are probably more inclined to follow advice or examples from more experienced people, whereas after years of experience people are more independent

Table 20.4 Frequencies of the types of situations requiring difficult decision-making (Question 4a) as a function of the level of expertise

Level of expertise	Types of situations						
	Organizing something	Choosing what to do	Choosing how to do	Strategic planning	Facing and emergency	Evaluating something	Missing
Low	18	6	0	3	0	3	1
Medium-low	3	2	1	5	1	0	0
Medium-high	2	6	3	4	0	0	0
High	8	7	5	5	0	2	0

Table 20.5 Frequencies of selected analogies (Question 5) of decision-makers as a function of the level of expertise

Level of expertise	Analogies										
	Giraffe	Sheep	Tortoise	Eagle	Lion	Crocodile	Dog	Shark	Seal	Other	Missing
Low	1	11	3	1	2	3	6	1	1	0	2
Medium-low	1	0	2	2	1	0	1	0	0	1	4
Medium-high	2	1	2	2	0	0	1	1	0	1	5
High	3	0	3	1	4	0	8	0	1	1	6

Table 20.6 Frequencies of the reasons of regret/nonregret (Question 8) as a function of the level of expertise

	Reasons of regret/nonregret			
Level of expertise	Mistakes are possible	Mistakes as learning experience	You cannot make mistakes	Missing
Low	9	14	7	1
Medium-low	7	3	2	0
Medium-high	7	8	0	0
High	13	7	2	5

(the dog was associated with relying with absolute sureness in strategies that had proved to be useful in the past, without considering the actual situation when making a decision). Crocodile was chosen mostly by less experienced people: the choice seems to be consistent with responses to Question 3 since this animal was associated with a hard decision-making process, which often brings to regretted decisions.

3.2.3 Metacognitive Knowledge of Processes and Expertise

In the case of the Metacognitive Knowledge of Processes section of the Solomon Questionnaire, differences emerged in the reasons participants gave about why good decision-makers regret a decision (Question 8; see Table 20.6), Pearson's contingency coefficient $C(N=85)=0.42$, $p<0.05$. People with higher level of expertise believed that mistakes are possible, probably because they had experienced them and learned that they are part of working life. On the other hand, workers with low level of expertise saw mistakes as learning opportunities, considering them as a way to highlight missing competences and learn new strategies. People's opinions concerning the possibilities and the reasons of regret appear thus to be coherent with the actual role of mistakes in working life.

On the other hand, no significant findings were obtained by analysing responses to Question 7 (asking people to enunciate the qualities of good decision makers) and 10 (asking how a person can become a good decision maker). This suggests that people appear to be more effective in verbalizing their metacognitive knowledge when they are asked to reflect on a specific aspect (such as regret) than when on a general one.

3.3 Relationships Between Decision Making and Individual Decision Style

3.3.1 Direct Metacognitive Awareness and Individual Decision Style

In the case of the Direct Metacognitive Awareness section of the Solomon Questionnaire, there were no significant differences between intuitive and analytical people, thus allowing the conclusion that metacognitive awareness that people have concerning the frequency and the specific type of decisions they make in their working life are not affected by the individual decision-making style. As for the description of a specific situation demanding a hard decision, the analysis failed to reach the level of statistical significance, although it could be noticed that in Question 4c intuitive people tended to report feeling more serene than analytical people when making a difficult decision, whereas analytical people tended to be more stressed and felt more responsible than intuitive people. Moreover, in order to make a difficult decision, intuitive people tried to imagine and to look for several different alternatives, whereas analytical people tended to activate a strategic planning and to examine the details of the situation (Question 4d). As numerous studies have highlighted, people with an intuitive style are creative problem-solvers who imagine and generate several options among which to choose, whereas people with an analytical style engage in a detailed analysis of the situation at hand by taking into consideration all the features of the situation (Scott & Bruce, 1995; Torrance, Reynolds, Riegel, & Ball, 1978).

3.3.2 Indirect Metacognitive Awareness and Individual Decision Style

In the case of the Indirect Metacognitive Awareness section of the Solomon Questionnaire, a significant association was found between the analogies selected to represent participants' own way of deciding (Question 6; see Table 20.7), Pearson's contingency coefficient $C(N=34)=0.54$, $p<0.05$. Specifically, analytical people opted for the "well organized travel", whereas intuitive people endorsed the forest and the hide-and-seek analogies. It seems that, while analytical people conceptualize decision making as a process which has to be well organized and planned in order to be successful, intuitive people represent the process as an "adventure" which does not require precise and detailed points of reference but, rather, can be seen as something that requires quick decisions to face unexpected circumstances.

3.3.3 Metacognitive Knowledge of Processes and Individual Decision Style

In the case of the Metacognitive Knowledge of Processes section of the Solomon Questionnaire, which explores people's conceptions of decision making, differences emerged in Question 8 regarding the beliefs that good decision-makers experience regret and the reasons for it (see Tables 20.8 and 20.9), Pearson's contingency

Table 20.7 Frequencies of the analogies representing the decision-making process (Question 6) as a function of the decision-making style

	Analogies						
Decision-making style	Mountain	Forest	Rapids	Cards	Desert	Hide-and-seek	Trips
Intuitive	0	2	0	0	0	2	1
Analytical	3	7	3	4	0	0	12

Table 20.8 Frequencies of the responses on whether a good decision-maker regrets or not a decision (Question 8) as a function of the decision-making style

	Good decision-maker		
Decision-making style	Regretting decisions	Not regretting decisions	It depends
Intuitive	0	1	3
Analytical	12	12	2

Table 20.9 Frequencies of the reasons for regret/nonregret (Question 8) as a function of the decision-making style

	Reasons for regret/nonregret			
Level of expertise	Mistakes are possible	Mistakes as learning experience	You cannot make mistakes	Missing
Intuitive	0	1	3	1
Analytical	12	12	2	3

coefficient $C(N=34)=0.52$, $p<0.05$ and $C(N=34)=0.50$, $p<0.05$, respectively. While analytical people believed that good decision makers can regret their decisions since mistakes are not only possible but they are also learning opportunities, intuitive people, on the contrary, thought that a good decision-maker cannot regret what he/she has decided because it is not possible to make mistakes for him/her.

These results are consistent with the idea that the conceptualization of intuitive/analytical styles and the distinction between maximisers/satisficers (Schwartz et al., 2002) are partially overlapping (Iannello, 2008). According to some authors (Parker, Bruine de Bruin, & Fischhoff, 2007), maximisers engage in rational decision making which reflects their perception of systematic deliberation about their choices, more dependence on others (which indicates the interpersonal comparisons and the quest for information they usually activate) and more avoidant decision making which reveals their tendency to postpone decisions to search for more information. Satisficers, on the contrary, look for the alternative that is over the threshold of acceptability and as soon as they find it, they opt for it. As a consequence, whereas maximising tendencies turned out to be related to perfectionism, need for cognition and regret, satisficing tendencies were found to be connected with happiness, optimism, and satisfaction with life. Probably, since maximisers always strive for the best solution and examine all the possible alternatives, they experience a greater feeling of regret after the decision has been made as compared to the satisficers.

No differences between intuitive and analytical respondents were found concerning the idea that being a good decision-maker is innate or learned (Question 9). Hence, people, even when taking their decisions more intuitively, do not believe more than analytical persons that their ability is innate. This implies that intuition is conceived to reflect knowledge and experience, at least for a good portion of our subsample.

4 Discussion

The present study was a first attempt to investigate people's metacognitive awareness and conception of decision making, a neglected area within the domain of metacognition. We followed an ecological approach by focusing our attention on decisions people often make in their professional life. A self-report instrument, the Solomon Questionnaire, was developed with the aim both to induce people to describe the types of decisions they usually make at work and to test to what extent they are aware of the emotions experienced and of the processes and strategies applied during decision making. An additional goal of the questionnaire was to assess individuals' beliefs about the optimal way to make a decision, the capacities which support it and the possibility to improve the decision-making skills. The questions included in the Solomon Questionnaire tried to encompass the various aspects of metacognitive knowledge mentioned in Sect. 1 by connecting them to decision making. More precisely, the questionnaire took into account aspects of metacognitive knowledge related to the task, the strategies, the emotions, the skills, and the personal attributes (Flavell, 1981). It included questions tapping awareness of one's own mental processes as well as beliefs about those of others.

In so far Hypothesis 1 is concerned, the analyses of the responses given to the questions included in the instrument we applied showed that the interviewees were able to report their own mental processes involved in making decisions and to express their beliefs about such processes. The answers given to the open questions appeared to be likely and reasonable. In addition, the descriptions of what people perceived to occur in their mind during decision making were always coherent with the justifications accompanying such descriptions. Finally, responses covered a relatively wide range of strategies, feelings, and self-attributions, and each respondent employed personal ways of expressing his/her own perceptions and opinions, making reference to his/her actual metacognitive experience, but not to widespread, superficial opinions. The overall impression is that people possess the ability to perceive and monitor, at least partially, the mental processes involved in decision making and report about them when asked, through both direct and metaphorical questions. For the same reasons the beliefs about the personal features and skills required to make good decisions also appeared to constitute not stereotypic opinions, but the results of genuine metacognitive reflections.

With respect to Hypothesis 2, the frequencies of the types of decisions spontaneously mentioned by the respondents varied consistently according both to the kind of profession they practiced (Hypothesis 2a) and to the level of expertise they had (Hypothesis 2b). Also, when asked to identify or to reconstruct the mental processes occurring when making decisions, participants provided responses which were differentiated depending on profession (Hypothesis 2a) and level of expertise (Hypothesis 2b). The same was true for the aspects of metacognitive knowledge investigated in the third section of the Solomon Questionnaire. The fact that the above mentioned differences were coherent with both the profession and the level of expertise can be considered as indirect support to the claim that the metacognitive competence tested by our questionnaire is valid (as suggested by the first hypothesis).

It is worth noting that distinct patterns of metacognitive knowledge depending on profession and professional expertise mainly emerged in relation to decision-making failures. A reflective attitude is often triggered when we have to face a problem, that is, when well-established automatisms or familiar approaches are not viable or have been proven to be ineffective and so we have to look for alternative strategies (Mecacci & Righi, 2006). Otherwise we have no reason to slow down the action by reflecting on what we are doing. The data of the present study showed that this also happens when people are involved in decision making: awareness of mental processes and metacognitive knowledge seem to be enhanced in response to mistakes or in unsuccessful outcomes. Also this is indirect evidence of the validity of the Solomon Questionnaire (Hypothesis 1) and of the influence of profession and expertise on metacognition about decision making (Hypotheses 2a and 2b).

The decision-making style (intuitive vs. analytical) also plays a distinctive role in modulating both metacognitive awareness and metacognitive knowledge (Hypothesis 2c). The awareness of the strategies used when making a decision appeared to be different in intuitive and analytical decision-makers. The decision process is perceived and described in terms which are consistent with one's personal decision-making style. Hence, style and metacognitive knowledge appear to be coherently connected.

The emotional aspects of metacognition (Efklides, 2006) investigated in the Solomon Questionnaire were found to be consistent with the general picture. Respondents were aware of the association between post-decisional regret and decision failures. Furthermore, the awareness of regret varied according to professional and personal characteristics. Regret was perceived mostly by people engaged in jobs which expose the worker to the likelihood of making mistakes, by less experienced individuals, and by persons who show maximisation tendencies. Similar trends emerged with reference to the perceived self-efficacy. Therefore, the cognitive and affective facets of metacognition seem to be interwoven within an integrated mental framework which individuals develop in reference to the way they make decisions.

The present study showed that both the personal awareness and the metacognitive knowledge about decision making are linked to professions and expertise and that personal style modulates this facet of metacognition, thus confirming our hypotheses. What are the implications of these findings? They concern mostly education. Biases that people tend to have while making decisions (see Sect. 1) could potentially be overcome by training them to develop and practice more efficient metacognitive strategies (Batha & Carroll, 2007; Glasspool & Fox, 2005; Kuiper, 2002; Lonie & Dolinsky, 2002). To do this, according to the results of our study, the constraints of profession, level of expertise, and decision-making style should be taken into consideration. The findings of the present investigation stress the need to keep in mind, when we try to enhance decision-making skills through metacognition, that the level of awareness and the metacognitive knowledge people possess concerning decision making are functionally connected to the actual decision tasks they have to face. Persons are prevalently aware of the aspects of the decision-making processes which are relevant to their goals and share metacognitive beliefs which mirror their habits. We know that people often are reluctant to abandon their spontaneous decisional strategies, even though not optimal, since they are convinced that they are the best ones for them (Cohen, Freeman, & Thompson, 1997). However, when explicitly asked to pay attention to the way they make a decision, individuals are less overconfident toward the choices they did (Kvidera & Koutstaal, 2008) and judge more correctly such choices (Sanna & Schwarz, 2006). This suggests that it is better to induce people to reflect upon their decision-making processes by taking into account the link with their actual jobs and their personal features than trying to convince them to apply new decision-making strategies and attitudes with no reference to the concrete professional situations and to the individual preferences.

4.1 Limitations of the Study

The main limitations of our study are two. First, we had no way to check whether what respondents reported about how they make decisions matches the actual processes they activate during decision making. This is a limit of many self-report procedures (Nisbett & Wilson, 1977). However, the internal coherence of the responses given by the interviewees and the coherence of the answers with respect to the profession, the level of expertise, and the styles of the participants provide indirect support to the validity of the Solomon Questionnaire. Second, only general aspects of metacognitive knowledge involved in decision making were investigated. The present study aimed to give only a preliminary contribution to the topic of metacognition in decision making. However, even a general overview of decision-makers' metacognitive competence, as that provided by this investigation, can allow us to test the applicability of the metacognitive perspective to the field of decision making and to set some reference points which might be useful to design further research which should address this topic in more depth. For instance, this could be done by considering the role of individual differences while analysing aspects of metacognitive awareness and knowledge associated to specific decisional biases. Another promising direction might be the investigation of the relationships between metacognition and decisional skills or the actual behaviour held in making a decision. Finally, the effects of different methods to improve metacognitive competence in decision making should merit attention.

Acknowledgements The authors would like to thank Guglielmo Puglisi for his assistance with the scoring of the Solomon Questionnaire and the PID.

5 Appendix A: The Solomon Questionnaire

The Solomon Questionnaire is designed to investigate how people make decisions at work. It takes about 15 min to complete. It is anonymous (you will be requested to give only some generic personal data) and it has no evaluative aims: there are no right or wrong answers. Please, answer sincerely to the questions.

1. How many decisions connected with your occupation do you make during a day on average?

2. Think about the decisions you make at work:
   (a) How many of them concern exclusively or mainly yourself? ....%
   (b) How many of them concern also other people? ....%
3. Thinking about the decisions you make during your working day:
   (a) How many times do you regret your decisions? ....%
   (b) How many times you don't regret your decisions? ....%
4. Keep on thinking about your working day. Identify a typical situation, or at least a situation that you often experience, in which making a decision is really demanding and difficult.
   (a) Describe the general situation, that is, the context in which you are requested to make this specific decision
   (b) Which is your first thought?
   (c) How do you feel when you make this kind of decision?
   (d) What do you do to make this decision?
   (e) Do you face the situation by yourself or do you ask others for help/advice?
   (f) Do you basically employ solutions that turned out to be effective in the past, or do you tend to try out new solutions?
   (g) Once you have made the decision, do you follow it or do you modify it (entirely or partly)? On the basis of which thoughts/reflections do you modify/do not modify your decision?
5. Thinking of yourself as a decision maker, how would you depict your specific way of making a decision? Identify, among the animals described below, the ones that best represent the way you usually make your decisions. Choose the three animals that best depict it, then arrange them in increasing order (from 1 to 3) depending on the extent to which they actually represent you.
   (a) Giraffe (when you make a decision you like to have a general overview of the situation. You give more importance to the whole than to single details)
   (b) Tortoise (you prefer to make your decisions calmly, paying attention to every single detail)
   (c) Eagle (you usually make your decision starting from a general overview, and being guided from you intuition in order to reach your goal as quick as you can)
   (d) Lion (you are very self-confident every time you make a decision, hence you rarely feel the need of discussing your decisions with others)
   (e) Crocodile (making decisions is hard for you, and you often regret them)
   (f) Dog (when you make a decision you tend to make use of those strategies that proved to be useful in the past, leaving aside the specific situation you are facing)
   (g) Shark (when you make a decision you focus on your final goal, without bothering about possible consequences on other persons, even if they could be negative)
   (h) Seal (once you make a decision, you tend to leave it behind, without defending it in front of other people)

(i) Sheep (while making decision you always need support from and comparison with the group)
(j) Other (please specify)

6. Which one of the following analogies do you believe to be the most suitable to depict the way you make your decisions?
   Making a decision is like… (one choice only)
   (a) Climbing a mountain
   (b) Walking across a forest without a map
   (c) Descending a river's rapids
   (d) Playing cards with friends
   (e) Losing oneself in the desert
   (f) Playing hide-and-seek
   (g) A well organised trip
7. In your opinion, which qualities characterise those people who are effective in making their decisions?
8. A good decision-maker is someone who never regrets his/her decision?
   (a) Yes
   (b) No

   Why?
9. According to your opinion, the competence of being "a good decision-maker" is:
   (a) Innate
   (b) Learned
   (c) Partly innate, partly learned
10. If you believe that the competence of making good decisions can be learned or improved, how do you think a person can become a good decision-maker?

6 Appendix B: Structure and sections of the Solomon Questionnaire

Section	Label	Number of items per section	The aim of each section was to measure:
1	Direct metacognitive awareness	4	Metacognitive knowledge of: Number and type of work decisions Strategies employed Emotions felt Processes involved
2	Indirect metacognitive awareness	2	Metacognitive knowledge of the self as a decision-maker (analogies)
3	Knowledge of processes	4	Metacognitive knowledge of processes involved in decision making

References

Antonietti, A., Ignazi, S., & Perego, P. (2000). Metacognitive knowledge about problem-solving methods. *British Journal of Educational Psychology, 70*, 1–16.

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes, 35*, 124–140.

Baron, J., & Hershey J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology, 54*, 569–579.

Batha, K., & Carroll, M. (2007). Metacognitive training aids decision making. *Australian Journal of Psychology, 59*, 64–69.

Betsch, C. (2004). Präferenz für Intuition und Deliberation. Inventar zur Erfassung von affekt und kognitionsbasiertem Entscheiden [Preference for intuition and deliberation. An inventory to understand decisions based on affect and cognition]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 25*, 179–197.

Bubany, S. T., Krieshok, T. S., Black, M. D., & McKay, R. A. (2008). College students' perspectives on their career decision-making. *Journal of Career Assessment, 16*, 177–197.

Brown, D. E., James, G. D., & Mills, P. S. (2006). Occupational differences in job strain and physiological stress: Female nurses and school teachers in Hawaii. *Psychosomatic Medicine, 68*, 524–530.

Cohen, M. S., Freeman, J. T., & Thompson, B. B. (1997). Training the naturalistic decision maker. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 257–268). Hillsdale, NJ: Erlbaum.

Colombo, B., & Iannello P. (2010). *L'uso della metafora per indagare la consapevolezza metacognitiva negli adulti* [Using metaphors to investigate adults' metacognitive awareness]. Manuscript submitted for publication.

Coutinho, S. (2008). Self-efficacy, metacognition, and performance. *North American Journal of Psychology, 10*, 165–172.

Dijksterhuis, A. (2004). Think different: The merits of unconscious thought in preference development and decision making. *Journal of Personality and Social Psychology, 87*, 586–598.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science, 311*, 1005–1007.

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review, 20*, 391–409.

Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation, and self-regulation. In A. Efklides, J. Kuhl, & R. M. Sorrentino (Eds.), *Trends and prospects in motivation research* (pp. 297–323). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2002). The systemic nature of metacognitive experiences. In P. Chambres, M. Izaute, & P.-J. Marescaux (Eds.), *Metacognition: Process, function and use* (pp. 19–34). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*, 277–287.

Fischhoff, B., & Beyth, R. (1975). "I knew it would happen": Remembered probabilities of once-future things. *Organizational Behavior and Human Performance, 13*, 1–16.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34*, 906–911.

Flavell, J. H. (1981). Cognitive monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills* (pp. 35–60). New York: Academic.

Glasspool, D. W., & Fox, J. (2005). Knowledge, argument, and metacognition in routine decision-making. In T. Betsch & S. Haberstroh (Eds.), *The routines of decision making* (pp. 343–358). Mahwah, NJ: Erlbaum.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization, 3*, 376.

Iannello, P. (2007). *Stili cognitive e decisionali: Il ruolo dell'attività lavorativa* [Cognitive and decision styles: The role played by work activities]. Imparare, 4, 39–62.

Iannello, P. (2008). *Intuitive and analytical thinking in decision making: The role of mindreading and cognitive style in a strategic interactive context*. Unpublished doctoral dissertation, Catholic University of the Sacred Heart, Milan, Italy.

Iyengar, S. S., Wells, R. E., & Schwartz, B. (2006). Doing better but feeling worse: Looking for the "best" job undermines satisfaction. *Psychological Science, 17*, 143–150.

Jamenson, A., Nelson, T. O., Leonesio, R. J., & Narens, L. (1993). The feeling of another person's knowing. *Journal of Memory and Language, 32*, 320–335.

Kahneman, D. (2003). A perspective on judgment and choice. *American Psychologist, 58*, 697–720.

Kluwe, R. H. (1982). Cognitive knowledge and executive control: Metacognition. In D. R. Griffin (Ed.), *Animal mind–human mind* (pp. 201–224). New York: Springer.

Kuiper, R. (2002). Enhancing metacognition through the reflective use of self-regulated learning strategies. *Journal of Continuing Education in Nursing, 33*, 78–87.

Kvidera, S., & Koutstaal, W. (2008). Confidence and decision type under matched stimulus conditions: Overconfidence in perceptual but not conceptual decisions. *Journal of Behavioral Decision Making, 21*, 253–281.

Lonie, J. M., & Dolinsky, D. (2002). Enhancing metacognitive skills using written narratives: An analysis of pharmacy student's negative health behaviors in a behavioral pharmacy class. *American Journal of Pharmaceutical Education, 66*, 273–276.

Lustie, S. J. D. (1998). Metaphorical thinking: Constructing cognitive classrooms. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 59*(6-A), 1891.

Mecacci, L., & Righi, S. (2006). Cognitive failures, metacognitive beliefs and aging. *Personality and Individual Differences, 40*, 1453–1459.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review, 85*, 1313–1326

Nelson, T. O., Kruglanski, A. W., & Jost, J. T. (1998). Knowing thyself and others: Progress in metacognitive social psychology. In V. Y. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Metacognition: Cognitive and social dimensions* (pp. 69–89). London: Sage.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259.

Parker, A. M., Bruine de Bruin, W., & Fischhoff, B. (2007). Maximizers versus satisficers: Decision-making styles, competence, and outcomes. *Judgment and Decision Making, 6*, 342–350.

Pietrzak, D., Duncan, K., & Korcuska, J. S. (2008). Counseling students' decision making regarding teaching effectiveness: A conjoint analysis. *Counselor Education and Supervision, 48*, 114–132.

Richetin, J., Perugini, M., Adjali, I., & Hurling, R. (2007). The moderator role of intuitive versus deliberative decision making for the predictive validity of implicit and explicit measures. *European Journal of Personality, 21*, 529–546.

Sanna, L. J., & Schwarz, N. (2006). Metacognitive experiences and human judgment: The case of hindsight bias and its debiasing. *Current Directions in Psychological Science, 15*, 172–176.

Schwarzt, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology, 83*, 1178–1197.

Scott, S. G., & Bruce, R. A. (1995). Decision-making style: The development of a new measure. *Educational and Psychological Measurement, 55*, 818–831.

Sillman, K. A. (1999). Metaphor: A tool for monitoring prospective elementary teachers' developing metacognitive awareness of learning and teaching science. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 60*(1-A), 0102.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 397–420). New York: Cambridge University Press.

Torrance, E. P., Reynolds, C. R., Riegel, T, & Ball, O. E. (1978). *Revised norms-technical manual for your style of learning and thinking. Forms A and B*. Athens, GA: Georgia Studies of Creative Behavior.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.

Van den Bergh, B., & Dewitte, S. (2006). Digit ratio (2D:4D) moderates the impact of sexual cues on men's decisions in ultimatum games. *Proceedings of the Royal Society, B, 273*, 2091–2095.

Vesonders, G. T., & Voss, J. F. (1985). On the ability to predict one's own responses while learning. *Journal of Memory and Language, 24*, 363–376.

Index