Chapter 2
Regulating Metacognitive Processes—Support for a Meta-metacognitive Ability

Sandra Buratti and Carl Martin Allwood

Abstract Second-order judgments aim to regulate metacognitive judgments or at least to assess the accuracy of metacognitive judgments (first-order judgments). For this reason, second-order judgments can be seen as a form of meta-metacognition. In this chapter, we clarify the concept of meta-metacognition and how it relates to first-order metacognitive judgments. Furthermore, we explain why the concept of second-order judgments is an important addition to the research literature on metacognition and why it is an important concept in the context of learning and memory. We also present a new generalizable method for eliciting and measuring the accuracy (realism) of second-order judgments in the context of confidence judgments of semantic and episodic memory performance and suggest how this method can be computer implemented. An asset of this method is that it allows for fine-grained analyses of the strategies that people use when they make second-order judgments without reverting to think-aloud reports.

Keywords Confidence judgments · Second-order judgment · Regulation · Episodic memory · Semantic memory

Abbreviations

JOL Judgments of learning
M Mean

S. Buratti (✉) · C.M. Allwood
Psychology Department, University of Gothenburg,
Box 100, 40530 Gothenburg, Sweden
e-mail: sandra.buratti@psy.gu.se

C.M. Allwood
e-mail: cma@psy.gu.se

A. Peña-Ayala (ed.), *Metacognition: Fundaments, Applications, and Trends*,
Intelligent Systems Reference Library 76, DOI 10.1007/978-3-319-11062-2_2

2.1 Introduction

In recent years, a new line of research within the field of metacognition has emerged. This research concerns second-order judgments of first-order metacognitive judgments [for short reviews, see [1, 2]. By the term 'second-order metacognitive judgment', we mean a judgment that aims to regulate a first-order metacognitive judgment (e.g., the accuracy of confidence judgment, judgments of learning, etc.) or that at least aims to assess the accuracy of a first-order judgment. A second-order judgment thus has a first-order metacognitive judgment as its target and can therefore be seen as meta-metacognition.

The assessment of the accuracy of metacognitive judgments is important in many different contexts, such as in learning. For example, these judgments guide children in allocating study time, choosing the best search strategies during problem solving, and finding the best answer to a learning task [3]. It is clearly important to the learner that this guidance is correct. Another important context is forensic psychology, especially eyewitness psychology, in which eyewitnesses often are asked to give metacognitive judgments of their memory performances [4]. Thus, it is of importance that the witness has an accurate judgment of the trustworthiness of their memory reports.

A substantial amount of the research on second-order judgments has focused on a specific type of metacognitive judgment, namely confidence judgments of different kinds of performances. Confidence judgments may, according to the model presented by Koriat and Goldsmith [5], have a central role when a person decides whether to externally report information that he or she has retrieved covertly in memory.

According to Koriat and Goldsmith's model, when people covertly retrieve information in their memory, at approximately the same time they also generate a spontaneous feeling of confidence associated with the correctness of the retrieved memory. The model next assumes that the person's decision whether to externally report the retrieved memory or not is based on a comparison between the level of the person's spontaneous confidence and the cost of reporting an incorrect memory in the specific context the person is in. For example, in a court of justice, a person is likely to use a stricter report criterion than when making conversation with a friend at a cafeteria.

Confidence judgments are important in many settings, partly because of the central role confidence judgments play when people decide whether to report retrieved memories, but also for other reasons. For example, confidence judgments of semantic memory information are important for optimizing learning outcomes [6, 7].

The confidence judgments a student makes about a certain performance are part of a self-monitoring process that can result in different types of action. For example, if a student takes an oral exam and does not feel confident about a certain answer, the student might decide not to report the answer openly.

If, on the other hand, the student is confident about the answer, he is more likely to report the answer to the test question. Confidence judgments often are also made in decision-making contexts. For example, people make confidence judgments in their professions [8]. Thus, physicians may judge how likely it is that

their patient will have a heart attack based on the patient's symptoms. Judges in court make decisions about whether to release an offender or not, based on how confident they are that the offender will not commit new offenses.

Finally, confidence judgments are often made about episodic memory performance, such as by witnesses in the justice system. In the context of the justice system, research shows that the professionals involved in the process, such as the police, prosecutors, attorneys, judges, and jurors, often judge the credibility of witness testimony by how confident the witness appears [9–11]. Thus, such professionals and actors have been found to believe confident witnesses more than less confident witnesses [e.g., 12]. For this reason among others, it is important that the witness has an accurate understanding of the accuracy of their confidence.

A problem is that studies have shown that the accuracy of (first-order) confidence judgments in a number of different contexts is poor [13]. In many judgment situations, people are more confident in their performance than they are correct in that performance. This phenomenon is known as the "overconfidence effect". Although settings exist in which people display underconfidence [14–16], some research supports a persistent overconfidence phenomenon in many types of situations [17].

For example, the overconfidence effect is common for both episodic and semantic memory. It is therefore important to investigate whether people have the ability to make their confidence judgments more realistic. Numerous models and theories have been presented that have explained the overconfidence effect [13, 17]. This overconfidence phenomenon is troubling in different contexts. For the student in the example above, it is problematic because overconfidence may hinder him or her from correctly allocating study time to information that he or she has not learnt properly yet.

In line with this, Dunlosky and Rawson [18] asked undergraduate students to learn six definitions, each with different subcomponents, and after each trial asked students to report their confidence in their recalled definition at three qualitative levels. After a student had reported the highest confidence for their recalled definition in three trials in a row, it was not practiced anymore.

The results showed that students who exhibited the most overconfidence performed at a low level at a final recall session 2 days after the first experimental session. In the forensic system, overconfidence also is problematic [19]. In real life, the testimony of overly confident witnesses has often led to the conviction of innocent people [20]. Likewise, research has shown that people's confidence judgments in most of the researched professions are miscalibrated in the sense that they are overconfident.

In brief, confidence judgments can have an important impact on the person making the confidence judgments as well as on other people facing the consequences of the judgments. It is therefore very important that people's confidence judgments are as *realistic* as possible (i.e., as close as possible to what is actually the case with respect to the person's performance). More specifically stated, the realism of confidence judgments depends on their relation to the correctness of the actual performance. Confidence realism is in some research traditions also called *confidence accuracy*.

In this chapter, we first discuss the notion of metacognition and then explain our conception of second-order judgments that we relate to the concept of *meta-metacognition* and show that there are different types of meta-metacognitive judgments. Next, we review this new line of research of second-order metacognitive judgments and then present a new method that we have developed for research on second-order judgments of confidence judgments of both semantic and episodic memory. After this, we present our own research using this method in which we investigated if people have the ability to make successful second-order judgments of already given first-order confidence judgments.

In this context, we describe the different strategies used when making second-order judgments in the two investigated memory contexts. The method can easily be adopted to assess second-order judgments of different kinds of metacognitive judgments and can, as we describe, be implemented in computer-learning contexts. Finally, we suggest directions for future research within this field.

2.2 What Is Metacognition?

To clarify what we mean by meta-metacognitive ability, we first address what we mean by metacognition. Flavell provided one of the earliest definitions of metacognition in the 1970s. According to Flavell, metacognition is "one's knowledge concerning one's own cognitive processes or anything related to them" [21, p. 232].

Flavell [22] proposed that metacognitive processes do not necessarily differ from cognitive processes, although the target of metacognitive processes is different because the metacognitive processes target cognition itself. Over time, the definition of metacognition has been expanded. For example, Koriat [23, p. 261] noted that metacognition more broadly can be seen "to refer to cognition about cognition in general, as well as self-regulation processes that take cognitive processes as their object". Thus, metacognition is commonly seen to include the regulation of cognitive processes and may, for example, also include knowledge of other people's knowledge. Metacognition has become a highly multidimensional concept, and its definitions and components vary extensively among researchers [for a review, see 24].

One of the most well-known models of metacognition is the two-level model presented by Nelson and Narens [3, 25]. Nelson and Narens [3] noted that this model is abstract and referred for their terminology to texts written by the mathematician David Hilbert in the 1920s and by the philosopher Rudolf Carnap in the 1930s.

We discuss this model next because it includes an early framework for the discussion of the concept of meta-metacognition, without mentioning this term. In this model, what is called *the meta level* controls *the object level* and receives monitoring information from the object level, that is, the cognition level. Through *control processes*, the meta level modifies the object level, but not vice versa, and the control exerted by the meta level on the object level is said by Nelson and Narens to be

analogous to speaking into a telephone handset. *Monitoring* is said by Nelson and Narens to be "logically (even if not always psychologically) independent of the control component" (p. 127). They also state that monitoring means that "the meta-level is *informed* by the object level" (p. 127, Nelson and Narens' italics). Seen from the perspective of the monitoring level, this concept is said to be analogous to a person listening to a handset. Control can have three effects on the object level: (1) initiating an action; (2) continuing an action; or (3) terminating an action. Information from the object level may change the model at the meta level of the situation at hand.

To further explain Nelson and Narens' model, the following example is offered: For a student taking a test to answer a question concerning some topic, the student needs to self-direct his or her search for the answer and thus select a search strategy. This selection of search strategy and the termination of the search are control processes. The confidence the student expresses in the answer is based on information from the object level, communicated by a monitoring process. The confidence level will be heeded when the person determines whether the answer is at a sufficiently satisfactory level to be presented during the test (compare Koriat and Goldsmiths' model [5]), or if a new search for a better answer candidate should be initiated. Nelson and Narens' model has been highly influential within the field of educational metacognitive research.

However, it is, as noted above, somewhat abstract and seems primarily not to have been intended as psychological but rather as a more formal abstract model. In line with this, Nelson and Narens addressed few of the specific processes pertaining to the model. The somewhat abstract formulation is illustrated in the context of their claim that the two-level model they presented could easily be generalized to more than two levels, in such a way that the meta level may be the object level of a higher meta level. Thus, depending on context, it seems that something that is on the meta level on one occasion (e.g., a confidence judgment) can be on the object level on a following occasion.

2.3 How Does Metacognition Relate to Meta-Metacognition?

Although Nelson and Narens [3, 25] did not use the term meta-metacognition, they, as described above, proposed that their two-level model could be generalized to include more levels, in such a way that the meta level may be the object level of a higher meta level. In general, just as it is useful to separate metacognition from cognition, it is also relevant to separate meta-metacognition from metacognition.

This suggestion was presented by Renkl et al. [26] in a paper where they proposed the importance of having knowledge of one's own metacognitive knowledge. Similar to Renkl et al., Roberts and Erdos [27] proposed that there could be different levels of metacognitive awareness where cognition of metacognitive processes was defined as a possible form of meta-metacognition. In a somewhat different version of meta-metacognition, Efklides and Misailidi [28] suggested that "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 cognition" [28, p. 16].

In line with most of these ideas, we suggest that, just as metacognition can be referred to as "any knowledge or cognitive activity that takes as its object, or regulates any aspect of any cognitive enterprise" [29, p. 150], any activity that targets the regulation of a metacognitive judgment can be referred to as a form of meta-metacognition [e.g., 30]; more in general, meta-metacognitive judgments are judgments that include the evaluation of the accuracy of metacognitive judgments.

This definition illustrates that second-order judgments provide a possibility for individuals to improve their metacognitive performance without relying on external sources, such as social and environmental sources, focused on by Kim, Park, Moore, and Varma [31].

The difference between meta-metacognitive and metacognitive and cognitive processes is primarily the target of the processes. Meta-metacognitive processes target a metacognitive performance product.

Too little is presently known, however, to say more about exactly in which manner meta-metacognitive processes differ in nature from metacognitive processes, except formally as discussed above. In this chapter, we present a method for examining and regulating second-order judgments, which we see as a form a meta-metacognition because they target metacognitive judgments.

There are also, however, other forms of meta-metacognition, such as having knowledge about metacognitive knowledge, as proposed by Renkl et al. [26]. Moreover, as discussed below, some forms of meta-metacognition may be of a more mixed form, or different from what we have suggested above [28, 32]. Efklides and Misailidi [28] suggested using the term 'meta-metacognition' for metacognition that "has as its object the cognition of others" [28, p. 16]. This activity, although obviously of great interest, might perhaps better be called "social metacognition" or similar because it explicitly excludes meta-metacognition of *one's own* metacognitive performance.

Yet other usages of the term meta-metacognition exist that do not involve the regulation of metacognitive judgments but that include evaluation of the accuracy of a previous metacognitive judgment as a part of making a confidence judgment about a decision that also may include other considerations [e.g., 33]. This phenomenon is explained more in detail further below.

2.4 Recent Research on Second-Order Metacognitive Judgments

Recently researchers have investigated the making of second-order judgments [e.g., 1, 30, 33–40]. In general, the results of these studies show that people can make successful second-order judgments of different types of metacognitive judgments such as confidence judgments, confidence intervals, judgments of learning (JOLs), and prediction of exam scores.

Dunlosky et al. [37] investigated second-order judgments in the context of JOLs (i.e., a person's ratings of the likelihood that they will be able to recall a recently studied item). The authors defined second-order metacognitive judgments as "an individual's confidence in the JOLs themselves" (p. 335) and reported that their participants could make successful second-order assessment of their JOLs. More specifically, Dunlosky et al. asked their participants to rate the accuracy of their JOLs, and the results showed that the participants were more confident in JOLs that predicted recognition success than they were in JOLs that predicted recognition failure.

A further finding showed that the participants had higher confidence for their delayed JOLs, that is, JOLs made some minutes after their item study session, compared to their immediate JOLs. Because the delayed JOLs were a better prediction of the participants' performance, this result is an indication of good realism of the second-order confidence judgments. Using a similar method, Serra and Dunlosky [40] compared the second-order confidence judgments of older adults (66 years) with those of younger adults (20 years).

For immediate JOLs, the older adults' second-order judgments were somewhat lower than those of the younger participants, but for delayed JOLs, the second-order judgments were about the same for the two age groups. The authors speculate that the reason may be that there were fewer salient cues (specifically: retrieval fluency) available in the context of immediate JOLs than after a delay and that the older participants for this reason were less willing to give high second-order judgments after the immediate JOLs compared with the delayed JOLs.

Meta-metacognitive judgments were also studied by Cesarini et al. [36], although the authors did not describe their research as lying within this framework. The participants were first asked to answer 10 numerical questions relating to economics by providing values for each question that constituted the upper and lower limits of a 90 % confidence interval for the correct answer. Next, they were asked to estimate how many of their 10 intervals included the correct answer. After this, all participants who had not estimated that 9 out of their 10 intervals included the correct answer were instructed to revise their intervals so that they thought that 9 of their intervals contained the correct answer. For the control group, the number of questions that included the correct answer was on average about 4.5 after the first round and nearly 6 questions after the second round. The adjustments in the confidence intervals thus resulted in better assessments because the goal was to construct intervals that included the correct answer for 9 of the 10 questions.

Miller and Geraci [39] investigated students' predictions of their exam scores. Students were first asked to make judgments of their number of correct answers to the exam (called *global judgments*) and then to make a confidence rating of the correctness of their predictions. Low-performing students were found to be more overconfident in their first-order predictions of their total exam performance than high-performing students, (i.e., they predicted higher scores than they had). However, the more interesting result was that the low-performing students were more accurate than the high-performing students in their second-order judgments

of their exam predictions. That is, they were better at evaluating the accuracy of their first-order global judgments compared with the high-performing students.

Händel and Fritzsche [38] studied the same issue and also investigated the effect of different types of five-step confidence scales on the success of the second-order judgments. In contrast to the study by Miller and Geraci [39], their results showed that the *high-performing* students made more successful second-order judgments of the correctness of their global judgments of their exam scores. It is not clear why this discrepancy in results between these two studies was found, and the authors did not relate their findings to those reported by Miller and Geraci.

In addition, the results presented by Händel and Fritzsche showed that bipolar non-numeric and non-verbal second-order confidence scales were more successful than unipolar versions of the same scale types and that the bipolar smiley scale (more or less sad–happy five-level smiley scale) was the best of the investigated scales.

In a recent study, Arnold [33] investigated a form of second-order judgments in a face-recognition task. In the task, the participants first saw 48 faces and then, after a filler task, performed a recognition test in which 24 of the already presented faces were mixed with 24 new faces. For each face, participants reported whether the face was *new* (they had never seen the stimuli before) or *old* (the stimuli had appeared during the study phase). The participants had to choose one of the options and were then, directly after the report for a face, asked to confidence rate how sure they were that the face was new/old. After this, the participants were asked if they wanted to report or withhold their answers, with the possibility of winning or losing points depending on the correctness of the reported answer.

The participants were then asked to rate their confidence in that they had made the right decision to go for points/withhold answer. In this way, Arnold separated the confidence a person had in an answer to the recognition task from the confidence he or she had in a decision about whether to withhold or report the answer and showed that these can differ. It is noteworthy that the confidence judgments of the decision to report probably did not solely pertain to the previous metacognitive performance because the decision likely would also have included figuring out the consequences of the specific reward structure the researcher had associated with the decision to report (compare with the Koriat & Goldsmith model [5]). Thus, the type of meta-metacognition investigated by Arnold may not be a pure form but rather a mixed or diluted form.

A majority of the studies investigating second-order judgments have focused on confidence as the means whereby a second-order judgment evaluates a first-order judgment of some kind (e.g., a confidence judgment, a global judgment, a confidence interval judgment, or a JOL). This relationship may be in line with how second-order judgments function in everyday life.

Such judgments may be triggered for various reasons; thus, for some reason, we may ask ourselves how confident we are in our metacognitive judgment—is it accurate (realistic) or does it need to be adjusted? One reason for such triggering may be that someone, for example, in court or a researcher, asks us to consider how realistic our confidence is. Another reason may be that we ourselves, more

or less automatically, note that some information that we heed contradicts our (for example) high confidence. This observation may lower our confidence in our first-order confidence and may make us reevaluate our confidence.

This thought is analogous to the process described in Koriat and Goldsmith's model [5] where, as described above, confidence judgments are used to regulate which information a person wants to report or withhold. If the person is so confident in a report that this confidence exceeds an accuracy criterion (which depends on the context), then the person will choose to report the information. If the confidence does not exceed the accuracy criterion, the person will not report the information. Second-order judgments of metacognitive judgments can be argued to function in a similar way. The second-order judgment is, or at least includes, an evaluation of the first-order confidence judgment and can be the basis for a person's decision to adjust a metacognitive judgment or keep it as it is. As noted above, the first-order metacognitive judgment does not, of course, need to be a confidence judgment but could be any type of metacognitive judgment, such as a JOL.

2.5 New Method for Studying Second-Order Judgments

Recently, we developed a new method for assessing second-order judgments, and with this method, we have shown that people have the ability to increase the realism in their confidence judgments for both episodic [1, 30, 34] and semantic memory tasks [35].

This method can be used to test a whole group of participants at the same time and, in brief, consists of an instruction about *confidence realism* and two tasks called the Confidence and the Adjustment tasks. The instructions on *confidence realism* explain the concepts of *realism* and *confidence judgments* in a way that makes them understandable to the participants. They address issues like *What is a confidence judgment?* and *How do confidence judgments relate to the actual memory report?*

The instructions also explain *What does it mean to show overconfidence, perfect realism, or underconfidence*? The instructions are 10 min long and help the person understand the concepts on a rather deep level. To check if the participants have understood the concepts explained, they are then provided with a test concerning these concepts.

After this, the participants continue to the Confidence task. In this task, the person gives a memory report of some kind and then provides confidence judgments for separate parts of this memory report. Depending on the type of memory tested, preparations for the Confidence task differ. For episodic memory, questions concerning some stimuli experienced by the participants are answered, and for this reason, the participants can, for example, be shown a short video clip about some event at the beginning of the test session. For semantic memory, general knowledge questions can be used, and no preparations before the confidence task thus may be needed. The format of the memory report can also differ. In our studies,

we have mostly used directed recall questions for which the participants are not given any answer alternatives but are told to provide their own answer.

When the participants have given their full memory report, for each memory question answered, they are asked to rate how confident they are in their answer on a confidence scale that ranges from 0 % ("I am absolutely sure that my answer is incorrect") to 100 % ("I am absolutely sure that my answer is correct").

As an introduction to the Confidence task, the participants are told to attempt to answer as many of the memory report questions correctly as possible and to be as realistic as possible in their confidence judgments. If they do not know the answer to a memory report question, they are told to make a guess. To have an incentive for the task, the participants can be told that the person with the most correct answers to the memory report questions will receive a reward (e.g., a movie ticket).

After the Confidence task, participants continue to the Adjustment task for which they are instructed to go back and change the confidence judgments they believe are unrealistic by making a new confidence judgment. The participants are also told that they are not allowed to change their responses to the memory report questions they answered during the Confidence task, only their confidence judgments. As an incentive, they can be told that the person with the best realism after this task will receive a reward (e.g., a movie ticket).

This method has been used for confidence judgments in episodic and semantic memory tasks, and can, as noted above, easily be adopted for other metacognitive judgments. For example, it can be used with JOLs. In that case, the participants are first asked to provide JOLs in some learning task for several items and then asked to go back and change the JOLs they find unreliable.

So far, we have used only recall questions with this method. However, it can also be used with recognition questions for which two or more answer alternatives are provided. In such cases, the confidence scale used should be altered to account for the possibility to be correct by chance, taking into account the number of options to choose from when answering.

2.6 The Making of Successful Second-Order Confidence Judgments

The above method has been used to investigate whether people can successfully improve their first-order confidence judgments by means of second-order judgments. In two studies, we showed that people can do this for an episodic memory task when being asked to go back and adjust their confidence judgments [1, 34]. In these studies, the participants first saw video clips concerning forensically relevant events (a theft in a park or a kidnapping).

After watching the video clip, they received instructions regarding the concept *realism of confidence* and performed a test that showed that they understood the concept. After this step, they completed the Confidence task, answering 40–50

directed recall questions (depending on the study) concerning the video clip and rating their confidence in their answers on a scale ranging from 0 % ("I am absolutely sure that my answer is incorrect") to 100 % ("I am absolutely sure that my answer is correct").

They then proceeded to the Adjustment task. The improvements in confidence realism after the Adjustment task showed that the participants on average significantly decreased their overconfidence, although the effect was small. These improvements were measured by using *absolute bias* and *calibration* scores, which are traditional realism of confidence measures [41]. Absolute bias is the absolute value of the bias score, and bias is calculated by subtracting the proportion correct from the average confidence level the person expresses in the task. Calibration is the average squared deviance between proportion correct and confidence for each separate confidence class (for example, for each of the 11 confidence classes: 0–9 %, 10–19 %, 20–29 %, …, 90–99 %, and 100 %). The results in a study by Buratti and Allwood [34] showed that the participants in the control condition improved their absolute bias score from $M = 0.14$ in the Confidence task to $M = 0.09$ in the Adjustment task.

In another study, using the same method, Buratti, Allwood and Kleitman showed that people also can adjust their confidence for semantic memory reports [35]. In that study, the participants answered 40 knowledge questions and rated how confident they were in their answers. The questions were directed recall questions. The results showed that the participants could successfully adjust their confidence judgments but that they accomplished this only by increasing the confidence for memory items that were correct. The confidence for correct items increased from $M = 0.88$ to $M = 0.90$.

The effect in that study was small, similar to studies of episodic memory [1, 34], but this outcome could result from the fact that the realism in this study was already high after the Confidence task. The value for absolute bias was 0.06 after the Confidence task, which is quite close to zero (the level for perfect realism).

This value is in line with earlier research in which less overconfidence is often found for recall questions than for recognition. Future studies should therefore investigate contexts in which first-order overconfidence is high in order to investigate contexts that are more sensitive to the participants' attempts to improve the realism of their confidence.

2.6.1 Strategies Used When Making Second-Order Judgments

When investigating if people have a meta-metacognitive ability, we found some surprising differences in strategies used for increasing realism of confidence. First, when people were asked to go back and try to increase the realism of their confidence judgments, they did not merely use a simple heuristic method in which they on average simply lowered their confidence across all items [1, 34].

In general, it should be noted that to increase the realism for correct items, the participants should increase the level of their confidence. The reverse is the case for incorrect items. Our analyses showed that selectively the participants were able to identify items in need of adjustment. For example, in the Buratti and Allwood study [1], analyses showed that the confidence judgments identified as in need of change were associated with lower realism than the confidence judgments that were not identified. That is, for correct items, the confidence level for chosen items ($M = 0.61$) was lower than the confidence level for items not chosen to be modified ($M = 0.84$). The reverse was true for incorrect items, namely that the items chosen for modification ($M = 0.65$) had a higher confidence level than the items not chosen for modification ($M = 0.51$).

Furthermore, in the Buratti and Allwood study [34], analyses also showed that the new confidence judgments (after change) were more realistic than the identified confidence judgments. For example, in the control condition, the participants' level of absolute bias was $M = 0.27$ for the chosen confidence judgments and $M = 0.14$ after the modification in the Adjustment task. This finding indicates that the participants chose the confidence judgments with the worst realism and increased the realism for these particular items.

Further, we found that depending on the memory task, the strategies used for increasing realism differed. When we analyzed the data separately for items the participants had answered correctly and items that they had answered incorrectly, we found for episodic memory that the participants were more successful with the incorrect items and lowered the confidence significantly for these items [1, 34]. For example, in the control condition in the Buratti and Allwood study [1], the participants decreased their confidence for incorrect items in the Confidence task from $M = 0.52$ to $M = 0.47$ in the Adjustment task. However, the confidence level for the correct items was the same across the Confidence and the Adjustment tasks.

In contrast, as mentioned earlier, the participants in the study investigating semantic memory increased the confidence for correct items, leaving the confidence level for incorrect items untouched [35]. These differences between episodic and semantic memory tasks could be explained by a divergence in encoding possibilities and feedback for episodic and semantic memory, respectively. Such variations have been suggested to explain why research has found differences in confidence judgments made for episodic and semantic memory [42, 43].

For example, the differences found in the regulation of first-order confidence judgments between episodic and semantic memory tasks could be due to variations in the form of feedback received. It may be that semantic memory information in general, compared with episodic information, is subjected to more confirmatory feedback. Over a lifetime, a person is expected to receive more confirmatory feedback that an answer is correct and therefore have more possibilities to encode correct answers. For example, learning that Rome is the capital of Italy is a fact that a person is likely to receive positive feedback on in several different contexts such as in school or when planning a holiday. Given this, it should be easier to know when a given answer in a general knowledge test is correct—i.e., when a low confidence judgment for a correct answer should be increased—than when a high confidence

judgment for an incorrect answer should be lowered. This scenario would explain why participants are more likely to increase the confidence for correct semantic memory items than they are to lower it for incorrect semantic memory items.

Episodic memory tasks lack the encoding and feedback possibilities that are common for semantic memory, which may result in a focus on the answers that are incorrect. The reason may be that people target answers that stand out in consciousness because they seem unlikely and that the person therefore may believe are incorrect. These experiences of answers as "unlikely" or "odd" events can been seen as a form of cognitive interruption because the information attended to in the answers does not follow schema-related expectations [44].

Likewise, such experiences may be associated with a decrease in the *Feeling of Rightness* (sense of correctness) suggested by Thompson et al. [45] to accompany answers or performance in problem solving. In line with this reasoning, Allwood [46], in a study using the think-aloud protocol, found that a fruitful way for first-year university students to detect errors that they had made when solving statistical problems was to follow up a noted difference between their *expectations* about some property of the generated answer and the answer they had devised.

2.6.2 Cues for Making Second-Order Confidence Judgments

In our studies we also investigated which cues were used when making second-order confidence judgments. Different types of cues have been assumed to affect confidence judgments [47]. Koriat et al. divided such cues into theory-based and experience-based cues, in which theory-based cues are related to beliefs one might have about one's competence and knowledge whereas experience-based cues are related to features of the person's phenomenological experience during the making of the confidence judgments. In one of our studies, we investigated two such possible experienced-based cues in the context of episodic memory, namely processing fluency and phenomenological memory quality [34]. Processing fluency is the subjective feeling of ease a person experiences when performing a cognitive task such as recalling a certain item [48]. Studies have shown that confidence judgments are to a large extent based on processing fluency so that high processing fluency is associated with high confidence judgments in semantic knowledge tasks [49, 50] as well as in eyewitness situations [51, 52].

Because processing fluency plays such an important part in making first-order confidence judgments, we hypothesized that processing fluency may also be an important cue when making second-order judgments. Thus, in Buratti and Allwood [34], the participants in the processing fluency condition, directly after each memory report question in the Confidence task, also made a rating of how fluently the answer came to mind and in another condition rated the memory quality of the item.

The results in the Buratti and Allwood study [34] showed that the confidence judgments participants tended to modify in the Adjustment task were associated with lower levels of experienced fluency ($M = 3.10$) than the confidence

judgments that were not modified ($M = 4.51$). Given that incorrectly remembered items are likely to be associated with lower fluency than correctly remembered items, this finding is in line with the fact that the participants, for episodic memory, mostly opted to change their confidence for incorrect answers. It is also in line with the idea that items that did not cohere with the participants' implicit or explicit expectations may have been associated with lower fluency and a lower Feeling of Rightness.

As mentioned above, we also investigated the experience-based cue of phenomenological memory quality [34]. Two types of phenomenological memory qualities are *remember* and *know* [53]. A memory has a "remember" quality if the person experiencing the memory recollects concrete details about the event or subject at hand. If, on the other hand, the person has only a vague feeling of familiarity during the retrieval, than this memory is associated with the quality "know".

In a study [54] investigating the realism of confidence in an eyewitness situation, a higher degree of realism was found for answers associated with the memory quality "remember" than answers associated with the memory quality "know" [54]. In our study [34], we showed that phenomenological memory quality could also be used as a possible cue when making second-order judgments. Answers associated with the memory quality "know" had worse realism (absolute bias $M = 0.24$) than the "remember" items (absolute bias $M = 0.13$), and the "know" items were chosen to be modified more often (52 % of the "know" items) than answers associated with the memory quality "remember" (20 % of the "remember" items).

2.6.3 Individual Differences in Second-Order Judgment

Our studies also analyzed individual differences associated with the making of successful second-order judgments. One such individual difference is cognitive ability, or more specifically, short-term memory. People who can hold more information activated might be better at regulating their metacognitive judgments than people with low ability in this regard.

However, our results showed no correlation between performance on the digit span task and how successful individuals were at making second-order judgments [30]. In one study, we also investigated the link between several personality and cognitive styles and second-order judgments [35]. Personality factors such as the Big five factors [55] and narcissism [56] were investigated together with cognitive styles such as Need for cognition [57] and Need for closure [58].

The only factor found to predict the confidence level in incorrect items for both first and second-order judgments was Openness [35]. People high in Openness tend to be open to possibilities and solutions. They have an active imagination and have a high intellectual curiosity [55]. They often perform well on cognitive tasks, which may lead them to not question and doubt their abilities, possibly providing the reason why they were significantly more confident for incorrect items than people low in Openness.

The Extraversion/Narcissism factor predicted confidence level in correct items for the first-order judgment only. This outcome is in line with earlier research showing that people high in extraversion [59, 60] and narcissism [61] tend to show overconfidence when performing different types of tasks. The weak relationship between personality and cognitive style variables on the one hand and second-order judgments on the other is not surprising because only weak relationships between personality and first-order judgments have been found [e.g., 62].

We also investigated if there was a relationship between willingness to make second-order judgments that involved changing one's first-order confidence judgments and different personality traits. We believed that people high in self-doubt would be more willing to question their ability and thus more prone to making second-order judgments that involved changing their first-order confidence judgments.

We also believed that people high in narcissism would not be as willing to doubt their ability, which would lead this group to make fewer second-order judgments of the just-described kind. However, no such relations were found between willingness to make second-order judgments and different personality traits [35]. The reason for this lack is unclear and should be investigated in future research.

2.7 Does Research Really Support that We Have a Meta-Metacognitive Ability?

What distinguishes a first-order judgment from a second-order judgment? Is the making of a second-order judgment merely redoing a metacognitive judgment? The process by which people regulate their first-order confidence judgments can be divided into two parts. The first part is an *identification* of the confidence judgments with poor realism that are in need of modification.

In this stage, the person needs to consider not only the confidence judgment per se but also the relation between the memory report that forms the target for the first-order judgment and the confidence judgment. This step may also include reconsidering the correctness of the memory report itself. In brief, the cues from the memory report and the level of the confidence judgment for each item need to match; thus, the whole metacognitive judgment needs to be evaluated. This questioning and evaluation of metacognitive judgments that happen during the identification part are, we believe, partly a process beyond just making another first-order judgment.

The second part in the regulation of first-order confidence judgments is the actual *adjustment* during which a new metacognitive judgment is made, and this process need not differ from a first-order metacognitive judgment [35].

We do not know yet which processes specifically underlie the regulation of confidence judgments. Critics might claim that the improvement in realism after the Adjustment task merely is an effect of redoing the confidence judgment.

One reason for this claim is that studies investigating error models, i.e., models assuming that all judgments are associated with measurement error, have reported that when people make multiple confidence judgments of the same item, their random error decreases, causing the realism to improve [63, 64].

Against this conclusion as the full explanation of the improvement in realism in the Adjustment task is the study of Allwood et al. [65] in which subjects in one condition were told to simply redo their confidence judgments. The results showed that there was no improvement in realism for this group. Another study investigating the dialectical bootstrapping model found that when participants were told to merely redo their estimations for different questions, they did not improve these estimations [66]. Rather, the improvement occurred in the condition in which the participants were told to actually question their first estimation. This outcome supports the idea that there needs to be a questioning of the first-order judgment that calls for a deeper or different processing than merely redoing a first-order judgment.

2.8 Illustration of Computer Implementation of the New Method for Improving Second-Order Judgments in a Learning Context

The new method presented earlier in the chapter could easily be implemented in a computer program, for use in classrooms, for example. Below, an illustration is provided for how this can be done, but various aspects of this illustration can obviously be altered to fit the specific needs of the situation. In the computer implementation, the instructions may be provided in writing on the screen.

If the session where the method is used is planned to involve new learning material, this material is presented first by the teacher in the classroom or on the computer screen. This material could be, for example, language information, some type of text (e.g., concerning events in history), or a short film on some topic. If the goal is to test material that has been presented to the students earlier, the students can start directly at the first step of the method.

This step presents the instructions on the concept of *confidence realism* in a lively way on the screen. Following this presentation, the test evaluating whether participants have understood the instructions on confidence realism could be given in a multi-alternative answer format so that the results can be automatically corrected by the computer program. If a student does not do well on this test, the computer could present further instructions and tests until the student has reached a suitable level of understanding. After this step, the program would allow the teacher to prepare for different types of memory reports from the students, either directed questions, with or without answer alternatives, or free recall instructions.

In the latter case, the student would be asked to enter his or her free recall in the form of one short sentence at a time delimited by use of the Return key. Each answer to a directed recall question and each free recall sentence delimited by a press on the Return key is seen as an “item”.

When the memory report task is finished, the student is instructed to confidence rate each answered memory report item on a confidence scale presented just below the reported item. This scale would be fitted to the number of answer alternatives provided in the context of the memory report items. For example, if two answer alternatives were provided, the scale would start at 50 % ("I'm guessing") and run to 100 % ("I'm sure the answer is correct"). When no answer alternatives are provided, the scale could, for the sake of simplicity, go from 0 % ("I'm guessing") to 100 % (I'm sure the answer is correct").

When the students have entered their confidence ratings on the scales, either a pause with another activity could commence or the next phase, the Adjustment phase, can start directly. In the Adjustment phase, the students are instructed to look over their confidence ratings (shown on the screen together with their associated memory reports) and change those ratings that they consider to be lacking in realism. When the students have finished the Adjustment task, they press the "Finished" key.

If the memory questions answered by the student are in recognition form, that is, with two or more provided answer alternatives, the computer can easily be programmed to correct the memory questions and then compute measures of the realism of the students' first- and second-order confidence ratings. When memory report questions without answer alternatives or free recall reports are used, much more advanced programming would be needed to correct the memory reports, and such programming may not be practically possible for many learning materials.

In situations where programming that automatically corrects the students memory reports has been done, however, the computer can next provide the students with feedback on how the realism of their new confidence ratings compares with the realism of their first confidence ratings. Such feedback on first-order confidence tasks has been suggested and tested by Leclercq [67], and similar means of providing feedback can be used in the present context.

For example, the feedback could be provided by means of two curves in a calibration diagram showing the different confidence levels of the confidence scale (e.g., from 0–100 %, in increments of 10 %) on the x-axis and the proportion correct items in each confidence class on the y-axis. As the students should be told when they are instructed on the concept of confidence realism, the diagonal in the calibration diagram shows perfect confidence realism, where, for example, 60 % of the items at the confidence level 60 % are correct. One of the two curves would show the student's first confidence levels and the second curve, in another color, would show their confidence ratings after they have been (possibly partly) changed in the Adjustment task. Multiple rounds of adjustments, each separated by a day or two or more, may be useful for the students.

2.9 Future Research

Because the field of second-order judgments is young, many different venues exist for future research projects. An important aspect that should be explored is how meta-metacognitive processes differ from metacognitive ones. This exploration

might be a difficult task because of still-unanswered questions concerning what separates metacognitive from cognitive processes. It might be that meta-metacognitive processes do not differ in nature from metacognition or cognition, but it is yet too early to say.

On a similar note, it would be interesting to further investigate which cues underlie meta-metacognitive judgments. So far, as mentioned above, one study has investigated the use of two examples of experienced-based cues [34]. It would also be interesting to research, however, how theory-based cues may affect second-order judgments. For example, an individual might believe that he or she is often too confident when it comes to their memory about events and thus will be more prone to making second-order judgments of their confidence judgments than people who do not have this theory about their confidence level in relations to their performance.

As described above, differences in strategies between doing successful second-order judgments for episodic and semantic memory have been found [1, 30, 34, 35]. It would be interesting to further investigate the actual reason for these differences in strategies. As suggested above, the divergence could be because of differences in feedback and encoding possibilities for the different types of memory reports.

It would be of interest to investigate whether people can make successful second-order judgments of performances in which individuals show *underconfidence*. Do people have the ability to increase their realism in these cases, as well? It would also be interesting to investigate which strategies the individuals use in these contexts.

Future research endeavors could additionally investigate if individuals using the method proposed earlier in this chapter can regulate other types of first-order judgments than just confidence judgments. Few of the second-order judgments have focused on the regulating aspect of second-order judgments and have instead focused on the assessment of first-order judgments. It would be interesting to investigate whether our method can help students to regulate their first-order JOLs, for example.

Finally, it would be of great theoretical and practical interest to combine research on meta-metacognition in the more individual sense described in this chapter with social versions of meta-metacognition [31, 32]. How could, for example, instructions to small groups of individuals be constructed to enhance their ability to improve the realism of their earlier first-order confidence judgments?

In this context, a study by Fraundorf and Benjamin [68] is of relevance. These authors investigated the possibility of using several first-order estimates of 12 numerical-estimate questions provided by the participants to optimize their performance, and the researchers also in one experiment incorporated the answers from other individuals. Although this study did not investigate second-order judgments because the participants were not asked to target first-order metacognitive judgments, it is still interesting because of the similarities between this type of multiple estimates study and second-order judgment studies in that both types investigate individuals' ability to improve their performance by making several judgments.

Likewise, research on integrating the different individual judgments of people in a larger group is relevant in this context [e.g., 69]. In all of these various contexts, ideas about how the association frame, or decision frame, of the individual or group deliberating can be broadened are important [e.g., 70, 71].

References

1. Buratti, S., Allwood, C.M.: The effects of advice and "try more" instructions on improving realism of confidence. Acta Psychol. **144**(1), 136–144 (2013)
2. Serra, M.J., England, B.D.: Magnitude and accuracy differences between judgments of remembering and forgetting. Q. J. Exp. Psychol. **65**(11), 2231–2257 (2012)
3. Nelson, T.O., Narens, L.: Metamemory: a theoretical framework and new findings. In: Bower, G.H. (ed.) The Psychology of Learning and Motivation, pp. 1–45. Academic Press, New York (1990)
4. Allwood, C.M.: Eyewitness Confidence. In: Granhag, P.A. (ed.) Forensic Psychology in Context: Nordic and International Approaches, pp. 281–303. Willan Publishing, Devon (2010)
5. Koriat, A., Goldsmith, M.: Monitoring and control processes in the strategic regulation of memory accuracy. Psychol. Rev. **103**(3), 490–517 (1996)
6. Kleitman, S., Gibson, J.: Metacognitive beliefs, self-confidence and primary learning environment of sixth grade students. Learn. Individ. Differ. **21**(6), 728–735 (2011)
7. Wang, M.C., Haertel, G.D., Walberg, H.J.: What influences learning? a content analysis of review literature. J. Educ. Res. **84**(1), 30–43 (1990)
8. Allwood, C.M., Granhag, P.A.: Feelings of confidence and the realism of confidence judgments in everyday life. In: Juslin, P., Montgomery, H. (eds.) Judgment and decision making: Neo-Brunswikian and process-tracing approaches, pp. 123–146. Lawrence Erlbaum Associates, Mahwah (1999)
9. Cutler, B.L., Penrod, S.D., Stuve, T.E.: Juror decision making in eyewitness identification cases. Law Hum. Behav. **12**(1), 41–55 (1988)
10. Lindsay, R.C., Wells, G.L., Rumpel, C.M.: Can people detect eyewitness-identification accuracy within and across situations? J. Appl. Psychol. **66**(1), 79–89 (1981)
11. Wells, G.L., Ferguson, T.J., Lindsay, R.C.: The tractability of eyewitness confidence and its implications for triers of fact. J. Appl. Psychol. **66**(6), 688–696 (1981)
12. Boyce, M., Beaudry, J.L., Lindsay, R.C.L.: Belief of eyewitness identification evidence. In: Lindsay, R.C., Ross, D.F., Don Read, J., Toglia, M.P. (eds.) Handbook of Eyewitness Psychology, vol. 2 Memory for People, pp. 501–525. Lawrence Erlbaum, Mahwah (2007)
13. Griffin, D., Brenner, L.: Perspectives on probability judgment calibration. In: Koehler, D.J., Harvey, N. (eds.) Blackwell Handbook of Judgment and Decision Making, pp. 177–198. Blackwell Publishing, Malden (2004)
14. Baranski, J.V., Petrusic, W.M.: Realism of confidence in sensory discrimination. Percept. Psychophys. **61**(7), 1369–1383 (1999)
15. Björkman, M., Juslin, P., Winman, A.: Realism of confidence in sensory discrimination: the underconfidence phenomenon. Percept. Psychophys. **54**(1), 75–81 (1993)
16. Stankov, L.: Calibration curves, scatterplots and the distinction between general knowledge and perceptual tasks. Learn. Individ. Differ. **10**(1), 29–50 (1998)
17. McClelland, A.G.R., Bolger, F.: The calibration of subjective probability: Theories and models 1980–94. In: Wright, G., Ayton, P. (eds.) Subjective Probability, pp. 453–482. Wiley, Oxford (1994)
18. Dunlosky, J., Rawson, K.A.: Overconfidence produces underachievement: inaccurate self evaluations undermine students' learning and retention. Learn. Instruct. **22**, 271–280 (2012)
19. Brewer, N., Wells, G.L.: Eyewitness identification. Curr. Dir. Psychol. Sci. **20**(1), 24–27 (2011)
20. Wells, G.L., Small, M., Penrod, S., Malpass, R.S., Fulero, S.M., Brimacombe, C.A.E.: Eyewitness identification procedures: recommendations for lineups and photospreads. Law Hum. Behav. **22**(6), 603–647 (1998)

21. Flavell, J.H.: Metacognitive aspects of problem-solving. In: Resnick, L. (ed.) The Nature of Intelligence, pp. 231–236. Erlbaum Associates, Hillsdale (1976)
22. Flavell, J.H.: Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. Am. Psychol. **34**, 906–911 (1979)
23. Koriat, A.: Metacognition research: An interim report. In: Perfect, T.J., Schwartz, B.L. (eds.) Applied Metacognition. Cambridge University Press, Cambridge (2002)
24. Lai, E.R.: Metacognition: a literature review. Technical Report. Pearson Research Report http://www.pearsonassessments.com/hai/images/tmrs/Metacognition_Literature_Review_Final.pdf (2011)
25. Nelson, T.O., Narens, L.: Why investigate metacognition? In: Metcalfe, J., Shimamura, A.P. (eds.) Metacognition: knowing about knowing, pp. 1–25. MIT Press, Cambridge (2011)
26. Renkl, A., Mandl, H., Gruber, H.: Inert knowledge: ANALYSES and remedies. Educ. Psychol. **31**(2), 115–121 (1996)
27. Roberts, M.J., Erdos, G.: Strategy selection and metacognition. Educ. Psychol. **13**, 259–266 (1993)
28. Efklides, A., Misailidi, P.: Introduction: the present and the future in metacognition. In: Efklides, A., Misailidi, P. (eds.) Trends and Prospects in Metacognition Research. Springer, New York (2010)
29. Flavell, J.H., Miller, S.A., Miller, P.H.: Cognitive development, 3rd edn. Prentice-Hall, Englewood Cliffs (1993)
30. Buratti, S., Allwood, C.M.: The accuracy of meta-metacognitive judgments: Regulating the realism of confidence. Cogn. Process. **13**(3), 243–253 (2012)
31. Kim, Y.R., Park, M.S., Moore, T.J., Varma, S.: Multiple levels of metacognition and their elicitation through complex problem-solving tasks. J. Mathematical Behavior **32**, 377–396 (2013)
32. Efklides, A.: Metacognition: defining its facets and levels of functioning in relation to self-regulation and co-regulation. Eur. Psychol. **13**(4), 277–287 (2008)
33. Arnold, M.M.: Monitoring and meta-metacognition in the own-race bias. Acta Psychol. **144**, 380–389 (2013)
34. Buratti, S., Allwood, C.M.: Improved realism of confidence for an episodic memory event. Judgment Decis. Making **7**(5), 590–601 (2012)
35. Buratti, S., Allwood, C.M., Kleitman, S.: First- and second-order metacognitive judgments of semantic memory reports: THE influence of personality traits and cognitive styles. Metacognition Learn. **8**(1), 79–102 (2013)
36. Cesarini, D., Sandewall, Ö., Johannesson, M.: Confidence interval estimation tasks and the economics of overconfidence. J. Econ. Behav. Organ. **61**(3), 453–470 (2006)
37. Dunlosky, J., Serra, M.J., Matvey, G., Rawson, K.: Second-order judgments about Judgments of learning. J. Gen. Psychol. **132**(4), 335–346 (2005)
38. Händel, M., Fritzsche, E.S: Students' confidence in their performance judgments: a comparison of different response scales. Educational Psychology: Int. J. Exp. Educ. Psychol. (In Press) (2014)
39. Miller, T.M., Geraci, L.: Unskilled but aware: reinterpreting overconfidence in low-performing students. J. Exp. Psychol. Learn. Mem. Cogn. **37**(2), 502–506 (2011)
40. Serra, M.J., Dunlosky, J.: Do older adults show less confidence in their monitoring of learning? Exp. Aging Res. **34**, 379–391 (2008)
41. Yates, J.F.: Subjective probability accuracy analysis. In: Wright, G.A., Ayton, P. (eds.) Subjective Probability, pp. 381–410. John Wiley and Sons, Oxford (1994)
42. Perfect, T.J.: The role of self-rated ability in the accuracy of confidence judgements in eyewitness memory and general knowledge. Appl. Cogn. Psychol. **18**(2), 157–168 (2004)
43. Perfect, T.J., Watson, E.L., Wagstaff, G.F.: Accuracy of confidence ratings associated with general knowledge and eyewitness memory. J. Appl. Psychol. **78**(1), 144–147 (1993)
44. Touroutoglou, A., Efklides, A.: Cognitive interruption as an object of metacognitive monitoring: feeling of difficulty and surprise. In: Efklides, A., Misailidi, P. (eds.) Trends and Prospects in Metacognition Research, pp. 171–208. Springer, New York (2010)

45. Thompson, V.A., Prowse Turner, J.A., Pennycook, G.: Intuition, reason, and metacognition. Cogn. Psychol. **63**, 107–140 (2011)
46. Allwood, C.M.: Error detection processes in sta-tisti-cal problem solving. Cogn. Science **8**, 413–437 (1984)
47. Koriat, A., Nussinson, R., Bless, H., Shaked, N.: Information-based and experience-based metacognitive judgments: evidence from subjective confidence. In: Dunlosky, J., Bjork, R.A. (eds.) A Handbook of Memory and Metamemory, pp. 117–136. Lawrence Erlbaum, Mahwah (2008)
48. Alter, A.L., Oppenheimer, D.M.: Uniting the tribes of fluency to form a metacognitive nation. Pers. Soc. Psychol. Rev. **13**(3), 219–235 (2009)
49. Kelley, C.M., Lindsay, D.S.: Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. J. Mem. Lang. **32**(1), 1–24 (1993)
50. Koriat, A.: How do we know that we know? The accessibility model of the feeling of knowing. Psychol. Rev. **100**(4), 609–639 (1993)
51. Robinson, M.D., Johnson, J.T., Herndon, F.: Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. J. Appl. Psychol. **82**(3), 416–425 (1997)
52. Robinson, M.D., Johnson, J.T., Robertson, D.A.: Process versus content in eyewitness metamemory monitoring. J. Exp. Psychol. Appl. **6**(3), 207–221 (2000)
53. Tulving, E.: Memory and consciousness. Can. Psychol. /Psychologie Canadienne **26**(1), 1–12 (1985)
54. Seemungal, F.V., Stevenage, S.V.: Using state of awareness judgements to improve eyewitness confidence-accuracy judgements. In: Chambres, P., Izaute, M., Marescaux, P.J. (eds.) Metacognition: Process, Function and Use, pp. 219–231. Kluwer Academic Publishers, Dordrecht (2002)
55. Goldberg, L.R.: The development of markers for the big five structure. Psychol. Assess. **4**, 26–42 (1992)
56. Raskin, R., Terry, H.: A principal components analysis of the narcissistic personality inventory and further evidence of its construct validity. J. Pers. Soc. Psychol. **54**, 890–902 (1988)
57. Cacioppo, J.T., Petty, R.E.: The need for cognition. J. Pers. Soc. Psychol. **42**(1), 116–131 (1982)
58. Webster, D.M., Kruglanski, A.W.: Individual differences in need for cognitive closure. J. Pers. Soc. Psychol. **67**(6), 1049–1062 (1994)
59. Pallier, G., Wilkinson, R., Danthir, V., Kleitman, S., Knezevic, G., Stankov, L., Roberts, R.D.: The role of individual differences in the accuracy of confidence judgments. J. Gen. Psychol. **129**(3), 257–299 (2002)
60. Schaefer, P.S., Williams, C.C., Goodie, A.S., Campbell, W.K.: Overconfidence and the big five. J. Res. Pers. **38**(5), 473–480 (2004)
61. Campbell, W.K., Goodie, A.S., Foster, J.D.: Narcissism, confidence, and risk attitude. J. Behav. Dec. Making **17**(4), 297–311 (2004)
62. Dahl, M., Allwood, C.M., Rennemark, M., Hagberg, B.: The relation between personality and the realism in confidence judgements in older adults. Eur. J. Ageing **7**(4), 283–291 (2010)
63. Johnson, T.R., Budescu, D.V., Wallsten, T.S.: Averaging probability judgments: Monte Carlo analyses of asymptotic diagnostic value. J. Behav. Dec. Making **14**(2), 123–140 (2001)
64. Wallsten, T.S., Diederich, A.: Understanding pooled subjective probability estimates. Math. Soc. Sci. **41**, 1–18 (2001)
65. Allwood, C.M., Granhag, P.A., Johansson, M.: Increased realism in eyewitness confidence judgments: The effect of dyadic collaboration. Appl. Cogn. Psychol. **17**, 545–561 (2003)
66. Herzog, S.M., Hertwig, R.: The wisdom of many in one mind: Improving individual judgments with dialectical bootstrapping. Psychol. Sci. **20**(2), 231–237 (2009)
67. Leclercq, D.: Validity, reliability, and acuity of self-assessment in educational testing. In: Leclercq, D., Bruno, J. (eds.) Item Banking: Interactive Testing and Self-Assessment, pp. 114–131. Springer Verlag, Berlin (1993)

68. Fraundorf, S.H., Benjamin, A.S.: Knowing the crowd within: metacognitive limits on combining multiple judgments. J. Mem. Lang. **71**, 17–38 (2014)
69. Surowiecki, J.: The Wisdom of the Crowds. Abacus, London, UK (2004)
70. Larrick, R.P.: Debiasing. In: Koehler, D.J., Harvey, N. (eds.) Blackwell Handbook of Judgment and Decision Making, pp. 316–337. Blackwell Publishing, Malden (2004)
71. Larrick, R.P.: Broaden the decision frame to make effective decisions. In: Locke, E.A. (ed.) Handbook of Principles of Organizational Behavior, 2nd edn, pp. 461–480. Wiley, Chichester (2009)