Jürgen Buder · Friedrich W. Hesse
Editors

Informational Environments

Effects of Use, Effective Designs

Springer

Informational Environments

Jürgen Buder • Friedrich W. Hesse
Editors

Informational Environments

Effects of Use, Effective Designs

Springer

Editors
Jürgen Buder
Leibniz-Institut für Wissensmedien (IWM)
Tübingen, Germany

Friedrich W. Hesse
Leibniz-Institut für Wissensmedien (IWM)
Tübingen, Germany

ISBN 978-3-319-64273-4 ISBN 978-3-319-64274-1 (eBook)
DOI 10.1007/978-3-319-64274-1

Library of Congress Control Number: 2017955219

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Contributors

Roger Azevedo North Carolina State University, Raleigh, NC, USA

Martina Bientzle Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Amanda E. Bradbury North Carolina State University, Raleigh, NC, USA

Katharina Brandelik Department of Psychology, Universität of Tübingen, Tübingen, Germany

Jürgen Buder Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Brett Buttliere Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Ulrike Cress Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Benjamin Fillisch University of Education Freiburg, Freiburg, Germany

Danny Flemming Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Peter Gerjets Knowledge Media Research Center, Tübingen, Germany

Annika Goeze Deutsches Institut für Erwachsenenbildung-Leibniz-Zentrum für Lebenslanges Lernen e.V., Bonn, Germany

Arthur C. Graesser Department of Psychology and Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA

Johannes Grapendorf Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Hannah Greving Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Martin Groß Department of Sociology, University of Tübingen, Tübingen, Germany

Torsten Grust Department of Computer Science, University of Tübingen, Tübingen, Germany

Andrew J. Hampton Department of Psychology & Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA

Jürgen Heller Department of Psychology, University of Tübingen, Tübingen, Germany

Friedrich W. Hesse Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Steffen Hillmert Department of Sociology, University of Tübingen, Tübingen, Germany

Institut für Soziologie, Eberhard Karls Universität Tübingen, Tübingen, Germany

Verena Hoehne Eberhard Karls Universität Tübingen, Hector-Institut für Empirische Bildungsforschung, Tübingen, Germany

Stefanie Jung Department of Psychology, University of Tübingen, Tübingen, Germany

Leibniz-Institut für Wissensmedien, Tübingen, Germany

Joachim Kimmerle Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Eberhard Karls University Tübingen, Tübingen, Germany

Viola Klotz Business School, University of Mannheim, Mannheim, Germany

Marie-Christin Krebs Leibniz-Institut für Wissensmedien Tübingen, Tübingen, Germany

Tanja Krumpe Department of Computer Engineering, Eberhard-Karls University Tübingen, Tübingen, Germany

Jasmin Leber Department of Psychology, University of Freiburg, Freiburg, Germany

Amna Liaqat University of Toronto, Toronto, ON, Canada

Michael P.C. Lin Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Anne M. Lippert Department of Psychology & Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA

Arita Liu Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Isabelle Mack Department of Psychosomatic Medicine and Psychotherapy, Tübingen University Hospital, Tübingen, Germany

Christian Marx Deutsches Institut für Erwachsenenbildung-Leibniz-Zentrum für Lebenslanges Lernen e.V., Bonn, Germany

Zahia Marzouk Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Tino G.K. Meitz Institute of Communication Research, Friedrich-Schiller-University of Jena, Jena, Germany

Garrett C. Millar North Carolina State University, Raleigh, NC, USA

Korbinian Moeller Department of Psychology, University of Tübingen, Tübingen, Germany

Leibniz-Institut für Wissensmedien, Tübingen, Germany

LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany

Zahra Mozaffari Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Nicholas V. Mudrick North Carolina State University, Raleigh, NC, USA

Nashwa Nashaat-Sobhy Universidad San Jorge, Zaragoza, Spain

John C. Nesbit Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Manuel Ninaus Leibniz-Institut für Wissensmedien, Tübingen, Germany

LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany

Hans-Christoph Nuerk Department of Psychology, University of Tübingen, Tübingen, Germany

Leibniz-Institut für Wissensmedien, Tübingen, Germany

LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany

Alexander Ort Department of Communication and Media Research (DCM), Empirical Communication Research, University of Fribourg, Fribourg, Switzerland

Azar Pakdaman-Savoji Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Alexandra Patzak Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Rolf Ploetzner University of Education Freiburg, Freiburg, Germany

Megan J. Price North Carolina State University, Raleigh, NC, USA

Mladen Rakovic Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Ilana Ram Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Alexander Renkl Department of Psychology, University of Freiburg, Freiburg, Germany

Stephanie Roesch Leibniz-Institut für Wissensmedien, Tübingen, Germany

Wolfgang Rosenstiel Department of Computer Engineering, Eberhard-Karls University Tübingen, Tübingen, Germany

Donya Samadi Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Kai Sassenberg Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Claudia Sassenrath Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany

Norbert Schäffeler Department of Psychosomatic Medicine and Psychotherapy, Tübingen University Hospital, Tübingen, Germany

Christian Scharinger Knowledge Media Research Center, Tübingen, Germany

Katharina Scheiter Leibniz-Institut für Wissensmedien Tübingen, Tübingen, Germany

Holger Schmidt Hochschule der Medien, Stuttgart, Germany

Bernhard Schmidt-Hertha Institute of Education, University of Tübingen, Tübingen, Germany

Josef Schrader Deutsches Institut für Erwachsenenbildung-Leibniz-Zentrum für Lebenslanges Lernen e.V., Bonn, Germany

Eberhard Karls Universität Tübingen, Institut für Erziehungswissenschaft, Tübingen, Germany

Anne Schüler Leibniz-Institut für Wissensmedien Tübingen, Tübingen, Germany

Mojtaba Soltanlou Department of Psychology, University of Tübingen, Tübingen, Germany

Leibniz-Institut für Wissensmedien, Tübingen, Germany

Graduate Training Centre of Neuroscience/IMPRS for Cognitive and Systems Neuroscience, Tübingen, Germany

Martin Spüler Department of Computer Engineering, Eberhard-Karls University Tübingen, Tübingen, Germany

Jason Stewart-Alonso Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Michelle Taub North Carolina State University, Raleigh, NC, USA

Thamar Voss Eberhard Karls Universität Tübingen, Hector-Institut für Empirische Bildungsforschung, Tübingen, Germany

Universität Freiburg, Institut für Erziehungswissenschaft, Freiburg, Germany

Jovita M. Vytasek Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Carina Walter Department of Computer Engineering, Eberhard-Karls University Tübingen, Tübingen, Germany

Hannes Weber Department of Sociology, University of Tübingen, Tübingen, Germany

Philip H. Winne Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Gottfried Zimmermann Hochschule der Medien, Stuttgart, Germany

Stephan Zipfel Department of Psychosomatic Medicine and Psychotherapy, Tübingen University Hospital, Tübingen, Germany

Guido Zurstiege Division for Empirical Media Research, Institute for Media Studies, University of Tübingen, Tübingen, Germany

About the Editors

Jürgen Buder studied psychology in Göttingen (diploma) and moved to Tübingen in 1995. There he worked at the German Institute for Research on Distance Education (DIFF; 1995–2000) and at the Department of Applied Cognitive Psychology and Media Psychology of the Institute for Psychology at Eberhard Karls Universität Tübingen (2000–2008). In 2002, he received a Faculty Award for his Ph.D. thesis on knowledge exchange. Since 2008, he has been at the Leibniz-Institut für Wissensmedien (IWM), where he became Deputy Head of the Knowledge Exchange Lab in 2012. He also coordinated the scientific development of the Leibniz-WissenschaftsCampus Tübingen *Informational Environments* (2010–2017) and the Leibniz-WissenschaftsCampus Tübingen *Cognitive Interfaces* (since 2017).

Friedrich W. Hesse studied psychology at the Universities of Marburg and Düsseldorf, received his doctorate at the RWTH Aachen, and qualified as Professor of Psychology at the University of Göttingen. He was research fellow at the Learning Research and Development Center (LRDC) and at the Carnegie Mellon University in Pittsburgh. He was Head of the Department of Applied Cognitive Science at the German Institute of Research for Distance Education (DIFF) and for 2 years he was director at the Laboratoire Européen de Recherche sur les Apprentissages et les Nouvelles Technologies (LERANT) in France funded by CNRS. He is the Founding Director of the Leibniz-Institut für Wissensmedien (IWM) and at present is Head of the Knowledge Exchange Lab. He is the Scientific Vice-President of the Leibniz Association (an umbrella organization for 91 research institutes in Germany) and holds the Chair of the Department for Applied Cognitive Psychology and Media Psychology at the Eberhard Karls Universität Tübingen. Friedrich Hesse has been an initiator and speaker for the first Virtual Graduate School *Knowledge acquisition and knowledge exchange with new media* funded by the Deutsche Forschungsgemeinschaft (German Research Society DFG), the DFG Priority Programme *Net-based Knowledge Communication in Groups*, the DFG Research Group *Analysis and Promotion of Effective Processes of Learning and Instruction*, the Leibniz-WissenschaftsCampus Tübingen *Informational Environments*, and currently the Leibniz- WissenschaftsCampus Tübingen *Cognitive Interfaces*.

Chapter 1
Informational Environments: Cognitive, Motivational-Affective, and Social-Interactive Forays into the Digital Transformation

Jürgen Buder and Friedrich W. Hesse

The Digital Transformation

Digital technologies have fundamentally transformed our lives. According to the Website internetlivestats, 3.6 billion people world-wide are using the Internet. With every passing second, more than 2.6 million e-mails are sent, more than 70,000 YouTube videos are watched, and more than 60,000 search queries are entered on Google. A study that tracked 94 smartphone users over several days found that, on average, people check their smartphone 76 times a day, performing 2,617 touch operations for a duration of 145 min (dscout, 2016). Apparently, we spend a lot of our time seeking information, pleasure, and companionship in a place that did not even exist 30 years ago.

Given how digital technologies permeate our lives, it is no small wonder that societal discourse has started reflecting on the various consequences that this digital transformation might have. While there are many angles from which the digital transformation can be scientifically investigated (e.g., sociological, political, clinical, legal), this book reflects on the digital transformation of society from a predominantly cognitive point of view. By cognitive we refer to the idea of looking at how individuals process information that they encounter in an environment, and how the information and the subsequent processing impacts the way people think, the knowledge they acquire, the attitudes that they have, and the judgments and decisions they make. Essentially, we investigate how people use information in order to get a better understanding of the world surrounding them.

J. Buder (✉) • F.W. Hesse
Leibniz-Institut für Wissensmedien (IWM), Schleichstr. 6, 72076, Tübingen, Germany
e-mail: j.buder@iwm-tuebingen.de

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_1

Informational Environments

The cognitive perspective on the digital transformation of society puts information and information processing at center stage. To begin our exploration, it might be a good start to show that investigating information and information processing touches on issues that are highly relevant to a person's life. Actually, all animal species make use of information in their environment to some degree, and they also communicate information to other animals (Witzany, 2014). Some of these forms of information use and information exchange involve communication with other animals of the same species, for instance by indicating prowess in mating rituals, by warning one's social group of approaching predators, or by the "waggle dance" of bees that indicates the direction and distance of food sources. Some animal communication is also trans-specific, for instance the rattle of a rattlesnake that keeps potential attackers at bay. From these examples it is evident that being properly informed about changes in a physical environment, or being able to transmit social signals to other individuals, provides a strong evolutionary advantage. Likewise, making use of information in the environment has played a crucial role in human evolution. In 1983, the psychologist George A. Miller coined the term "informavore"—higher animals (like humans) do not only forage and consume food in order to survive, but they also forage and consume information for the very same purpose: adequately perceiving social situations, learning about the environment in order to better adapt, and ultimately making decisions that might increase one's well-being and survival in a complex environment (Miller, 1983). As informavores, humans rely on knowledge passed down through generations, they consult others for advice, and they inform (and are informed) about new pieces of evidence that help with being aware of what's important or interesting in their physical and social surroundings.

The information foraging theory proposed by Pirolli and Card (1999) illustrates that these "biologistic" arguments are more than just a nice metaphor. The theory is built on the optimal foraging theory that was proposed in biology (McNamara, 1982), and which tries to predict the food foraging behavior of animals. In a similar fashion, the information foraging theory tries to predict how humans are seeking information in their environment. To some extent, this behavior is shaped by the informational needs of a person (the information diet). In addition, the behavior is shaped by the observation of cues in an environment that indicate how rich or relevant a particular piece of information will be. These cues are conveyed through an "information scent." Based on the information diet and the information scent, people perform cost-benefit analyses on which pieces of information in their environment they need to attend to, and when to move to another "information patch." Computational modeling has shown striking similarities between the food-foraging patterns of animals and the information-foraging behavior of humans. Among other things, the information foraging theory was applied to how people use information on the World Wide Web (Fu & Pirolli, 2007). For instance, the various hits that a query in a search engine yields provide a user with cues that conveys the "information scent" of a distal source. By formalizing the information scent and the goals of a user, Fu and Pirolli (2007) were able to computationally model users' decisions to click on a link or to return to a previous Web page. Taken together, "biologistic" models like the information-foraging theory describe the behavior of individuals in an environment as being shaped both by internal factors like the information diet and by external factors like the information scent.

The ideas put forward by the notions of informavores, information foraging, and other biologistic concepts heavily influenced our own development of the notion of "informational environments," which lent the title to the present book. Thus, we define informational environments as the set of informational resources that a person habitually taps into over a life-course in order to get a better understanding of the world.

This definition requires some elaboration, as it might differ in some respects from the way that many scientists, politicians, or practitioners think about the digital transformation. First, the emphasis on "information" and on "getting a better understanding of the world" suggests that our definition emphasizes cognitive aspects. While we acknowledge that people use digital technologies for many purposes, and while we acknowledge that using digital technologies has many consequences in people's lives, we focus on the use of information in order to acquire knowledge, in order to refine one's thinking, in order to form and apply attitudes, and in order to make informed decisions and judgments.

Second, we believe that it makes little sense to think about the digital transformation in isolation. Rather, we emphasize that digital technologies are just one part of a much larger and much more historically rooted informational infrastructure. Yes, people do "Google" an awful lot, and they write e-mails or read Websites, but of course there is an abundance of non-digital means to improve one's understanding of the world. For instance, people acquire information and build knowledge from books, newspapers, or TV. Moreover, in the physical world people are asking their parents and friends for advice, or they interact with experts like teachers, doctors, or lawyers who hold particular knowledge that will help to better understand the world surrounding them. Thus, our definition of informational environments encompasses both digital and physical information resources. People freely move in and out of these various contexts, and what they take up in their physical and social surroundings has an influence on how they use digital technologies, and vice versa.

Third, by suggesting that informational environments address patterns of behavior "over the life course," we emphasize our strong commitment to the notion of lifelong learning (Fischer, 2000). This is a particularly important point as many people might exclusively associate our cognitive angle with learning at schools. But it would be a mistake to only focus on information use in formalized learning settings. We began this chapter with the frequency of using Google, sending e-mails, watching an online video, or using a smartphone. Most of these episodes where people try to get a better understanding of the world surrounding them take place outside of dedicated learning institutions, they take place at the workplace or during people's free time. Consequently, our definition of informational environments pertains to formal, non-formal, and informal learning settings (Werquin, 2010).

Fourth, our definition focuses on the individual. Each person has his or her own informational environment, and in terms of informational resources to tap into, or in terms of topics of interest, the informational environments of any two individuals are likely to differ substantially. From this follows the idea that informational environments are highly idiosyncratic. There is no curriculum for lifelong learning, and we strongly believe that being interested in sports, jazz music, or rabbit breeding can be just as "valid," important, and knowledge-intensive to an individual as being interested in chemistry or other subject matters that are taught in schools and universities.

Fifth, our definition emphasizes individual autonomy. Even in formal learning settings like classrooms, the notion of self-regulation has taken a firm footing (Boekaerts, 1999). For lifelong learning, however, the notion that people are at their own discretion in deciding what informational resources to tap into, and whether to leave a resource at any time, becomes even more important. Therefore, we strongly believe that a purely cognitive perspective is not sufficient to understand the use of informational environments. Rather, research should also take motivational, affective, or social aspects into account.

Sixth, our definition comprises contexts with highly varying amounts of support. For instance, in a classroom setting, students are guided and mentored by highly skilled teaching professionals in order to maximize their learning gains. However, for many endeavors in lifelong learning (e.g., searching the Web for health-related information) such an explicit and dedicated support from professionals is lacking. Therefore, researchers and practitioners have tried to design informational resources in a way that the technology provides support for optimal information use (e.g., search engines that sort Web links by relevance; online "question and answer" forums for peer support).

In sum, our definition of informational environments is very broad by trying to capture the richness and range of settings where people try to get a better understanding of the world around them. It encompasses the "foraging" of physical and digital informational resources, it covers the entire life-span, it is not restricted to the classical subjects and topics covered in schools, and it pertains to settings with different amounts of support. While the main focus of the research presented in this book is on the digital transformation of society and its cognitive implications, we believe that this can be better understood by taking a much broader view. This includes analyzing how informational environments have developed over time, and how they were shaped through the invention of information and communication technologies, to which we will turn next.

A Very Brief History of Informational Environments

In prehistoric times, the individual informational environments of people were comprised of the social networks that they were embedded in. Knowledge was transmitted and acquired through interaction with parents, children, other relatives, elders, and priests. It is interesting to note that the cognitive function of getting a better understanding of one's surrounding was inextricably intertwined with social interaction, whether it is in the relation between parents and children, teachers and students, or masters and apprentices. Moreover, many forms of interaction were combining cognitive (learning) and motivational-affective (entertainment) elements, whether it was in the telling of stories, in games, or in communal rituals. Many of these elements carry on until the present day: we still rely on family and friends to seek or give advice. By interacting with others, we grow into a community that transmits values and ways of thinking—first through partaking in simple tasks in a community ("legitimate peripheral participation"; Lave & Wenger, 1991), and subsequently by taking up ever more central roles in a social network. The cognitive perspective of information use is embedded in a motivational and social fabric.

Technologies for information transmission and communication often introduced transformative extensions of individuals' informational environments. For instance, the invention of writing permitted a preservation of information and knowledge over time (and to some extent, also over places). Interestingly, the invention of writing (which led to scrolls and books) also stripped information use and sensemaking off its social component—individual acquisition of information through reading became a prevalent mode of learning and information use that continues until the present day. In fact, many people associate the term "learning" with individually acquiring information from books or other non-social media.

Printing paved the way for mass media, which provided access to information that was compiled not only at different times, but also originated from very different places. As a consequence, informational environments slowly began to contain elements on a more global scale as people got access to a much broader range of informational resources. Subsequently, access to information became more immediate (e.g., through newspapers), keeping people up to date about events that unfolded very recently. Moreover, through the invention of the telephone, radio, sound recordings, movies, and television, other codes of transmission than text and static pictures were employed to convey information.

With the advent of computers, the Internet, the World Wide Web, and social media, informational environments became ever more complex. In terms of the sheer amount of information that has become accessible, they outperform prior technologies by orders of magnitude. With the advent of mobile technologies, accessing and using information has become almost completely unrestrained in terms of time and place. Immediacy of information transmission is also unparalleled, for instance with Twitter messages sent on a global scale as events unfold in real-time. Interestingly, while many previous technological developments (e.g., books, newspapers) strengthened the individual use of informational environments at the cost of social-interactive functions, digital technologies have also "re-socialized" informational environments, thus taking up elements that were the only mode of information transmission in prehistoric times. Through social media, people can immediately talk back and exchange information with others, thus providing a platform for individual self-expression on a truly global scale.

It should be noted that more recent technological innovations often did not replace earlier ones, but rather added ever higher complexity to informational environments. We still rely on advice given by family members and friends, and we still consult experts in face-to-face contexts to get a better understanding of the world surrounding us. In Western societies, most people freely criss-cross between the physical and digital parts of their informational environments. Elements from the physical environment are transferred to the digital environment, and vice versa. For instance, we might share experiences that we've made with a product in the physical world by writing a user review on a digital portal. Conversely, many people who visit a doctor have previously searched for medical information online.

Of course, the advent of digital technologies should not be regarded as the final step in the evolution of informational environments. Though it is tricky to predict future trends, there is reason to believe that we are at the beginning of a phase where the physical and digital parts of informational environments merge. On the one hand, many physical objects are already connected to the digital realm (e.g., credit cards, modern TV sets), or are on the brink of being interfaced with the digital world (e.g., "smart homes," automated vehicles). The "Internet of things" is geared at softening the boundaries between the physical and the digital (Höller et al., 2014). At the same time, through sensors (e.g., activity trackers), or through advances in artificial intelligence (Markoff, 2016), or learning analytics (Baker & Inventado, 2014), natural language processing and deep learning (LeCun, Bengio, & Hinton, 2015), digital technologies have ever better capabilities to "be informed" about "their" environment—they "know" about our physical environments and about us much more than ever before, thus exhibiting cognitive properties ("cognitive interfaces"; Buder & Hesse, 2016).

In sum, today's informational environments offer a bewildering array of informational resources that individuals may tap into. Trying to make sense of all these different informational resources and to integrate them into a coherent role is the challenging task that people in modern societies face. How they go about doing this, and how this daunting task may be supported through technologies, is at the heart of this book.

About this Book

Thinking about cognitive aspects of the digital transformation, and developing the concept of informational environments was the starting point for a group of scientists located at the IWM (Leibniz-Institut für Wissensmedien) in Tübingen, Germany who all share an interest in understanding how digital technologies shape our daily lives. The sheer breadth of the research topic made us realize from very early on that understanding how people use their informational environments will benefit from a multi-disciplinary approach. Though most researchers at IWM have a background in psychology and thus have some expertise in the cognitive aspects of using digital media, it was evident that other disciplines can contribute a lot to this endeavor. Therefore, the present book does not provide an exclusively psychological point of view, but has contributions from a variety of disciplines, among them computer science, education, neuroscience, sociology, media sciences, and medical science. In 2010, we began to establish a research infrastructure in Tübingen, the ScienceCampus "Informational Environments," which brought researchers from IWM together with colleagues mainly from the Tübingen University, but also from other places in Germany. The present book reports findings from 7 years of research within the ScienceCampus, and it is complemented by contributions from several US-based colleagues that we have closely collaborated with over these years.

We suggest that a scientific investigation of informational environments has two components, a descriptive one and a prescriptive one. These two components are also reflected in the way this book is organized. The first component (effects of use) is descriptive-analytical in nature, and it revolves around the question of how people currently use digital (and non-digital) resources to extract information, and make sense of their informational environment in order to better understand the world surrounding them. For instance, we have investigated how internal personal factors like prior knowledge or prior attitudes influence information use. At the same time, we explored how information use is shaped by external factors, such as rules and regulations that govern the behavior in a given informational context.

The second component of research on informational environments is about effective designs, and is prescriptive-interventionist in nature. It revolves around the question of how environments can be created that provide opportunities for people to make even better sense of the world surrounding them. For instance, we have investigated which design features are most conducive to learning. And we have also explored the importance of being motivated, of having positive emotions, and of having the ability to closely interact with information and people for learning.

Those were initial and recurring questions that motivated our research, and ultimately also this book. In the following, we will discuss some issues that are associated with effects of use and effective designs, thereby providing a narrative that combines and integrates the subsequent chapters.

Effects of Use

The first broad question that fueled this book is about how individuals make use of their informational environments. We have already discussed informational environments at some length, providing a definition, discussing elements of this definition, and discussing how informational environments developed over the course of history. However, thus far we have not said very much about the "user," the individual person who traverses his or her informational environment. Therefore, here we will provide a description of central characteristics of the human mind as they pertain to the contents of this book. We have already maintained that a cognitive perspective of information use is paramount. However, we believe that to fully understand information use, we should also include a motivational-affective perspective, and a social-interactive perspective. These three perspectives are discussed here.

The Cognitive Perspective

The cognitive perspective that permeates current psychology holds first and foremost that humans are information processors. This was the basic tenet of the "cognitive revolution" that coined psychology from the mid-1950s on. Ever since, most scholars in psychology regard humans as cognitive systems that process information (Newell & Simon, 1972), thereby drawing an analogy between the functioning of the human mind and the functioning of computers. Information is perceived, selected, and taken up from sensory organs. The processing itself is believed to occur in a limited-capacity working memory (Baddeley, 2007). Here, relevant pieces of information are selected, organized, and integrated with knowledge that is retrieved and activated from long-term memory structures (Mayer, 2005).

Knowledge, defined as representations that reside in long-term structures of the human mind, is another concept that is crucial for the cognitive perspective. A common distinction is made between two types of knowledge representations, namely between declarative and procedural knowledge (Anderson, 1983). Declarative knowledge refers to mental representations of facts (e.g., knowing the capital of France). In contrast, procedural knowledge refers to skills that people have acquired, ranging from motor skills (e.g., doing a golf swing) to mental skills (e.g., knowing how to multiply two numbers). Declarative and procedural knowledge do not exist in isolated pieces. Rather, pieces form larger structures that are connected through associations. Larger structures of declarative knowledge can be referred to as conceptual knowledge (e.g., having general knowledge of chemistry), whereas larger structures of procedural knowledge can be referred to as competences (e.g., knowing how to solve a certain class of math problems, knowing how to teach a class).

A central conjecture for the cognitive perspective of the individual human mind is that there is no fixed directionality in the "flow" between information in the environment and knowledge in the head. Many psychological conceptualizations acknowledge this duality (Cress & Kimmerle, 2008; Melton, 1963; Piaget, 1970). On the one hand, there is a "bottom-up" flow that originates from information in the environment and subsequently might lead to changes of knowledge in the head. New knowledge might be formed or existing knowledge might be re-structured on the basis of incoming information (accommodation; Piaget, 1970). On the other hand, there is a "top-down" flow that impacts how people process information. Without prior knowledge we have no ability to understand and make sense of information coming in from the environment, so we make use of knowledge that is already stored in the mind and apply it to the world (assimilation; Piaget, 1970). For example, this kind of top-down processing is used to understand and make sense of language (Marslen-Wilson & Welsh, 1978).

The dual nature of top-down and bottom-up processing can also be illustrated when examining how people learn. One can conceive of learning as "bottom-up" as it typically involves a change in knowledge structures on the basis of incoming information. However, this bottom-up process is accompanied and self-regulated by "top-down" strategies and skills on how to learn (metacognition; Flavell, 1979). As a consequence, learning science is not only about getting knowledge into the head of individuals, but also about making individuals become metacognitive information processors that effectively deal with information in a self-regulated manner.

The cognitive system is driven both by similarities and by discrepancies. The ability to detect similarities in pieces of information is evident in mechanisms such as chunking (grouping together of pieces of declarative information; Miller, 1956) or priming (where the activation of one piece of knowledge speeds up the response toward semantically related pieces of information; Meyer & Schvaneveldt, 1971). On the other hand, the cognitive system is capable of registering discrepancies (e.g., between two pieces of information that reside in the head; Festinger, 1957). The ensuing cognitive conflicts call for attempts to resolve these discrepancies, and this focus on conflicts is regarded as a powerful mechanism to engender learning (Piaget, 1970).

The Motivational-Affective Perspective

While the main focus of our definition of informational environments is on cognitive aspects, we believe that we can better understand the effects of using informational environments when we also look at other factors that have an impact on human information processing. Declarative information from an environment is not always processed neutrally, but is frequently combined with an affective evaluation of the content (Olson & Zanna, 1993). For instance, it is almost impossible to just see or think of, say, Thai food without an affective evaluation ("I like Thai food" or "I don't like Thai food"). These affective evaluations of an object are the basis of attitudes. Attitudes refer to all kinds of objects (e.g., Thai food, the President of the USA, the concept of individual freedom), and they weave together cognitive components (the declarative information one encounters), affective components (the evaluations one assigns to these pieces of information), and behavioral components (the actions one performs on the basis of these evaluations) (Rosenberg & Hovland, 1960).

Consequently, the way people process information is infused by their attitudes, and there is an abundance of empirical evidence showing that human information processing is highly motivated and subjective (Kunda, 1990). An immediate consequence of motivated processing is the introduction of bias. The psychological literature abounds with examples showing that people process information in a highly selective manner: for instance, most people show a preference for information that supports their current attitudes (congeniality bias; Hart et al., 2009). Similarly, ambiguous information is interpreted in a way that confirms the current attitudes (biased assimilation; Lord, Ross, & Lepper, 1979). Many other cognitive biases are employed in order to interpret the world in a self-serving manner, for example to attribute success to one's internal abilities, whereas failure is attributed to external circumstances (Campbell & Sedikides, 1999).

Motivated processing is closely intertwined with affects. Attitudes are affective evaluations, and attitudinal biases can be conceived of as a kind of affective regulation—seeking information that is favorable with regard to one's attitudes, and defending against, avoiding, or derogating information that elicits unfavorable responses (Festinger, 1957). Similarly, self-serving biases attach positive information and events to oneself in order to bolster self-enhancement, while negative information and events are regarded as threatening to the self (Banaji & Prentice, 1994).

The Social-Interactive Perspective

In a wonderful quote, Edwin Hutchins (1995) reported looking at his office to seek for anything "that was not either produced or delivered to its present location by the cooperative efforts of individuals working in socially organized groups" (p. 175), concluding: "The only thing in my environment that meets this test is a striped pebble that I found at the beach and carried home to decorate my desk" (p. 176). Indeed, humans are social animals, and interacting with others is the glue that holds societies together and drives individual and societal advances. The psychological literature has also stressed the importance of forming strong and stable interpersonal relationships (Baumeister & Leary, 1995), and it consistently shows that belongingness to groups can be a crucial component of how people define and categorize themselves (Turner & Oakes, 1986).

Groups impact the information processing of individuals, and vice versa. An example of how groups can influence the thought and behavior of individuals is group conformity (Asch, 1951), a normative pressure to (at least) publicly agree with a majority viewpoint even if it is perceived to be wrong. However, there are also examples of how individuals impact groups, whether it is through participation in a collaborative community (Lave & Wenger, 1991), or through minority influence (Wood, Lundgren, Ouellette, Busceme, & Blackstone, 1994). Thus, interpersonal exchange has important implications of how individuals process information, leading to various outcomes such as individual learning, group learning, individual decision making, or group decision making (Buder, 2017).

Being in social contact with other persons may have many aspects (e.g., providing safety, instilling pride), but for the purposes of this book we regard social interaction predominantly as the opportunity of engaging oneself in a responsive environment. Thus, while many activities in the digital world (e.g., individually playing a computer game) are not inherently social, they share with social activities the element of interactive engagement with a responsive environment. For this reason, we use the term "social-interactive perspective" to simultaneously refer to the mutuality of interacting with others and/or the interaction with a responsive technology.

Summary

We suggested that to further our understanding of how the use of informational environments impacts the human mind, we should analyze the corresponding stakeholders (environments and individuals). With regard to environments, we emphasized their bewildering variety. They consist of a multitude of informational resources that persons tap into. Moreover, these resources encompass both the physical and the digital realm. The resources also cover an enormous topical breadth, addressing the various interests and needs that characterize a person over a lifetime. Informational resources in an informational environment also differ with regard to how much they are affording autonomy and providing guidance to an individual. In addition to that, each of these physical and digital information resources may have its own rules and regulations about who is allowed access to them, and of how to exhibit proper conduct in them.

In our discussion of the human mind we focus on three perspectives. The first (and central) perspective is cognitive, and it refers to the common psychological conceptualization of humans as information processors. Various forms of knowledge have been discussed. The flow of information between individual and environment is bidirectional, and it is driven both by similarities and by discrepancies. However, a purely cognitive perspective is not sufficient to describe the use of informational environments, as information processing is infused with affective and attitudinal evaluations. Motivated processing introduces biases and selectivity that help explain how individuals constitute and make use of their informational environments. Finally, though our definition of informational environments pertains to individuals, individuals rarely act in isolation, but are embedded in social contexts of real, implied, or imagined others. Therefore, our cognitive angle on the informational environment should also be complemented by a social-interactive angle.

Consequently, our investigation of how people use their informational environments looks both at characteristics of an environment that affect an individual, and at characteristics of the human mind (cognitive, motivational-affective, social-interactive) that affect the behavior in an environment. In Chaps. 2–5 of this book we will highlight parts of this interactive dance between individual and environmental factors.

Outline of Part I of this Book (Effects of Use)

Chapter 2 (Hillmert, Groß, Schmidt-Hertha, & Weber, 2017) provides an integrative view on the interaction between individuals and their environments. In particular, it investigates how individual and environmental factors may impact the life-altering decision of students to drop out from university. Universities constitute a significant part of the informational environment of many people, and they exhibit many characteristics that we addressed when discussing the definition of informational environments (they are comprised of physical and digital contexts, and they provide a mixture of learner autonomy and professional guidance). As to the environmental factors that may impact student dropout from university, Chap. 2 focuses on rules and regulations that constrain information use, in particular as they pertain to grading procedures. Grading procedures might differ in several respects (e.g., whether students can participate in the grading, or in how far grades actually are valid indicators of performance), and the chapter tracks how far the perceived justice of grading procedures impacts student dropout. Moreover, the chapter takes a look at individual precursors of student dropout. While dropout in itself reflects an outcome of a motivational variable, the chapter investigates how far informational competences of students (cognitive perspective) and social integration with other students (social-interactive perspective) contribute to dropout decisions. The chapter also reports on a large online survey that tracked individual and environmental factors across various academic fields in Tübingen University, and the influence of these factors on student dropout.

Chapter 3 (Buder, Buttliere, & Cress, 2017) focuses on how individual factors impact the way that people use their informational environments. In particular, it suggests that the use of informational environments is driven by cognitive conflicts, which can manifest as knowledge discrepancies (cognitive perspective) or attitudinal discrepancies (motivational-affective perspective). Research on cognitive conflict is discussed from the viewpoint of two academic disciplines, learning sciences and social psychology. In addition, Chap. 3 reports on a set of studies that investigate how cognitive conflicts play out in social media settings where people are afforded an opportunity to act on their environment (social-interactive perspective).

Chapter 4 (Kimmerle et al., 2017) investigates the use of informational environments in lifelong, informal contexts, particularly in the way that people seek and process health-related information on the Internet. Similar to Chap. 2, it highlights individual factors that impact information use. In particular, the chapter reports on studies showing how information use is influenced by conceptual knowledge that people have about health (biomedical vs. biopsychosocial health concept). The impact of epistemological beliefs (conceptual knowledge about knowledge and knowing) and the impact of prior opinions is also discussed in this chapter. Finally, Chap. 3 covers how far information seeking and information use of individuals is shaped by affective variables (mild vs. severe health threats). All of these individual factors lead to certain biases in information use, thus underscoring the general notion that humans are motivated information processors.

Similar to Chap. 4, Chap. 5 (Zurstiege et al., 2017) is also focusing on health-related issues, thus addressing lifelong, informal learning outside of formalized institutions. In particular, it investigates how obesity-related health campaigns that were embedded in various online platforms (Facebook, blog, online news sites) may affect assessments of self-relevance (cognitive perspective) and self-perception (motivational-affective perspective). Results provide evidence that one and the same content will be contextualized and processed differently, depending on the platform on which it occurs, thus providing interesting insights into the question of how to frame health campaigns. The chapter also discusses how health communication campaigns that contain fear appeals interact with individual coping skills to have either supportive or disruptive effects. Finally, Chap. 5 leads into the second part of the book by reporting on the design of a serious game (motivational-affective perspective) that provides health-related information (cognitive perspective), and involves physical (inter-)activity (social-interactive perspective).

Effective Designs

The second part of the book is dedicated to effective designs. While the first part focuses on the entirety of a person's informational environment (e.g., various University contexts, use of various social media environments, use of various health-related platforms), research in the second part predominantly addresses the design of single "information patches" in a person's informational environment. Moreover, most of the research in the second part of the book focuses on learning activities, with other sense-making activities (judgment, decision making, attitude formation and application) playing only a minor role.

In order to identify principles for the effective design of learning activities, it might be useful to have a yardstick against which to measure effectiveness. We would suggest that this yardstick is presented by the relationship between a teacher and a learner. Until the present day, the one-on-one interaction between teacher and student has been hailed as the gold standard to improve learning, and some of the most sophisticated technologically-driven learning environments are measured up against the performance of one-on-one teaching (VanLehn, 2011). What is it that makes teacher-student interactions so effective? We believe that the answer to this question can be found by examining the three perspectives that we have discussed in the previous section on effects of use. First and foremost, good teachers address the cognitive perspective, not only by transmitting knowledge "bottom-up," but by assisting learners in developing "top-down" strategies. Second, good teachers motivate learners by giving positive feedback and providing constructive ways to deal with negative learning experiences and negative learning performances, thus addressing the motivational-affective perspective. And third, the interaction with teachers caters to the social-interactive perspective by engaging students in a dialogical interplay, prompting for answers and ideas, and helping to shape and refine one's worldview (a technique that was already practiced in the Socratic method of ancient Greece).

Of course, while one-on-one teaching can be regarded as the gold standard to effectively support learning, it might be difficult to achieve. Obviously, one-on-one teaching for every individual learner is not feasible in institutional contexts. Moreover, lifelong learning in informational environments often unfolds in informal settings where teachers simply are not available. Therefore, the idea to support learning through the use of technologies (often in the absence of a teacher) has a long history. However, early "learning technologies" like books only catered to parts of the cognitive perspective by providing "bottom-up" access to content that a learner could acquire. In terms of "bottom-up" access to information, we are now living in unparalleled times, with the World Wide Web providing incredible amounts of information at one's fingertips. However, just providing access to resources does not address other parts of the cognitive perspective (e.g., managing workload, supporting self-regulation, and metacognition) and does not cater to the motivational-affective and social-interactive needs. Having vast amounts of knowledge at one's disposal is an important step in technological development, but it is far from providing the teacher-like richness of mutual interaction, which is the hallmark of effective learning.

It took until the twentieth century before technologies were developed that at least remotely had some teacher-like qualities. In the 1920s, Sidney Pressey developed a mechanical "teaching machine" that administered multiple-choice questions to learners and would only display the next question once the correct answer was picked (Pressey, 1926). Later on, B. F. Skinner continued to use and refine teaching machines (Skinner, 1961). Ever since, the support of learning through mechanical (and later on, digital) technologies has become a central theme in educational research and development. The PLATO system at the University of Illinois had become a hotbed for new technological teaching and learning tools from 1960 onward. Not only did it provide online courses for learners, it also was among the first systems to cater to social-interactive functions of digital media by providing e-mail, messaging, and chat functionalities as well as multiplayer games since the early 1970s. In the meantime, a lot has been learnt about how to provide effective designs for learning. In the following, we discuss some of these principles, using our tripartite distinction between a cognitive perspective, a motivational-affective perspective, and a social-interactive perspective.

The Cognitive Perspective

With regard to the cognitive perspective of effective designs for learning, we highlight three areas: how to ensure that the limited capacity of the working memory is used effectively; how to ascertain that declarative knowledge becomes conceptual knowledge; and how to support the development of skills in self-regulated learning.

The limited capacity of human working memory has become a central research topic in educational psychology and instructional design since the 1990s. John Sweller (1988) has coined the term cognitive load to describe the various factors that impinge on working memory capacity. Cognitive load theory distinguishes between three types of cognitive load: an intrinsic load, which is taken up by processing the complexity of learning material; an extraneous load, which is taken up by processing the particular representation of information; and a germane load, which is used to elaborate learning content, thus fostering the acquisition and integration of material (Sweller, Van Merriënboer, & Paas, 1998). The main idea is that the design of learning environments should try to minimize intrinsic load (e.g., through breaking down complex material into simpler elements, where possible), and try to minimize extraneous load (through proper design and representation of learning content), so that a lot of working memory capacity is free for germane load. One way to reduce extraneous load is to design information in a way that learners do not have to switch too much between different representations (e.g., text and pictures), as these "split-attention effects" inhibit learning performance (Chandler & Sweller, 1992). Other effects that will reduce extraneous cognitive load involve combining visual and auditory information in non-redundant ways (Chandler & Sweller, 1992). Cognitive load theory paved the way for the cognitive theory of multimedia learning (Mayer, 2005), which provided a theoretical model, and developed additional design principles. Most importantly, Mayer's approach and subsequent work investigated the best way that text and pictures should be combined and integrated to support learning. While cognitive load theory and the cognitive theory of multimedia learning had a huge impact on instructional design and educational psychology, the former was plagued by the problem that cognitive load could only be measured indirectly (through self-reports, measurements of reaction times, or test scores). This has led to novel ways on how to measure how learners process information. For instance, more recent research has focused on eye movements as indicative of learning processes (van Gog & Scheiter, 2010), or on the way that mental workload can be captured by electrical activity of the brain (Murata, 2005).

The second area we will briefly discuss refers to the question of how declarative knowledge can be transformed into rich conceptual knowledge structures. Simple declarative knowledge of facts can be acquired via rote learning. However, the goal of effective learning should rather be on gaining an associatively linked body of conceptual knowledge, and to be able to flexibly use the knowledge one has acquired in a variety of contexts. One way to measure whether learners have conceptually mastered knowledge is to test for knowledge transfer, using tasks that require learners to apply their knowledge to new situations (Clark & Voogel, 1985). There are various design principles that can foster the transformation of declarative to conceptual knowledge. For instance, cognitive flexibility theory (Spiro, Collins, Thota, & Feltovich, 2003) suggests that the ability to fluidly apply knowledge in different contexts and from multiple perspectives can be fostered by presenting learning content as hypertexts or hypermedia rather than through linear text.

A third active research area pertaining to cognitive aspects of learning refers to procedural knowledge and skills, particularly to skills in self-regulated learning (Boekaerts, 1999; Winne & Hadwin, 1998). The model proposed by Winne and Hadwin (1998) regards self-regulated learning as a four-stage process where learners define a task for themselves, then engage in goal-setting and planning, then enact these plans, and finally adapt their procedural knowledge structures in order to be prepared to master future tasks. The model of self-regulated learning by Boekaerts (1999) does not focus on stages of a self-regulated task, but rather conceptualizes self-regulated learning occurring on three different layers. On a central layer, learners regulate their general mode of how to interact with information (e.g., processing information on a surface level vs. processing through deeper elaboration). On a second level, learners regulate their learning process through metacognitive activities that involve monitoring one's learning progress and controlling one's behavior. The third level is about choosing general goals and resources, and according to Boekaerts (1999) it has stronger motivational than cognitive components (see below). These models of self-regulation and metacognition (Nelson & Narens, 1994) have spurred considerable research on how to foster learners' abilities to develop these relevant skills. For instance, self-regulation can be facilitated through the use of step-by-step instructions on how to solve a task (worked-out examples; Renkl, 1997), by metacognitive prompts (Bartholomé & Bromme, 2009), or by capturing eye movements of advanced learners and making these movements visible to less advanced learners (Mason, Pluchino, & Tornatora, 2015).

The Motivational-Affective Perspective

In contrast to the cognitive perspective on learning, much less research has been devoted to motivational and affective aspects of learning. In this section, we will briefly touch on one of the few exceptions, a meta-analysis by D'Mello (2013) integrating 24 studies on the emotions that students experience while learning with technology. The emotions that were most commonly reported were engagement/flow, boredom, and confusion, followed by curiosity, happiness, and frustration. The occurrence of some of these emotions (engagement/flow, boredom, frustration) also seems to be strongly dependent on contextual factors. Boredom and frustration were reported more often in scientific studies that were conducted in the laboratory, and which used self-reports and relatively unsophisticated learning technologies. The more interesting finding for current purposes, however, was that the positive emotion of engagement/flow was reported more often in authentic contexts and when advanced learning technologies were employed. The analyses by D'Mello (2013) investigated three types of advanced learning technologies that all seem to have some positive impact on motivational and affective dimensions: intelligent tutoring systems that provide one-on-one teaching to a learner (Aleven, Mclaren, Sewall, & Koedinger, 2009); animations and simulations (Ainsworth, 2008); and educational games (Plass et al., 2013). It should be noted that all three types of learning technology have a strong interactive and/or social component, to which we turn next.

The Social-Interactive Perspective

Findings reported in the previous section suggest that technologies that cater to social-interactive needs have a strong impact on the motivation to learn. This further substantiates our idea that mutual and highly responsive social exchange is a key feature that makes one-on-one teacher-student interactions the gold standard for individual learning. In other words, if technologies cater to the social-interactive needs of learners, and if they come close to the ideal of an interpersonal communication between teachers and learners, this is likely to unlock the power to support learning. Consequently, many technological approaches have implemented social-interactive components. We would argue that there are two very broad approaches on how to create this kind of interactivity. The first is to employ technologies that (among other things) enable learners to get into interpersonal contact with other learners. The second broad approach is to implement social and interactive features in the way that technologies themselves respond to a learner. We will briefly discuss each of these strategies.

One way to cater to social-interactive needs is to design technologies in a way that engenders interpersonal interaction. Cooperative learning has become a hot topic in the learning sciences since the 1970s (Johnson & Johnson, 1975), and since the 1990s the research field of computer-supported collaborative learning (CSCL) has investigated how digital technologies can be used as a platform and a facilitator for learning groups. CSCL approaches are researched and implemented in a variety of settings and for all kinds of interpersonal interaction, ranging from dyadic learning to “mass collaboration” (Cress, Moskaliuk, & Jeong, 2016). One important feature of CSCL is that groups are not only seen as settings where learners mutually “deliver” knowledge to each other (Roscoe & Chi, 2007). Instead, collaborative learning has the potential to create a situation where learners take up ideas from each other and develop and refine their understanding of content. This process, which is similar to scientific progress, is referred to as knowledge building (Scardamalia & Bereiter, 2006). Several methods have been developed to facilitate collaborative learning and knowledge building, for instance by structuring collaborative episodes through so-called collaboration scripts (Kollar, Fischer, & Hesse, 2006), or by providing contextual information about the knowledge of one’s learning partners (Engelmann, Dehler, Bodemer, & Buder, 2009).

While CSCL promotes learning by bringing technologies into collaborative and interpersonal settings, a second broad approach is to instill social and interactive features into the individual interaction between humans and computers. In other words, the second strategy is to make interaction with a technology more similar to the social interaction with a teacher. A basic requirement to achieve this goal is to provide learners with an opportunity to actively engage with the technology. For instance, the system needs to respond to activities that are initiated by the learner. A simple word processor or a search engine already fulfils this most basic requirement. However, interactivity cuts both ways, and a teacher-like system should also initiate activities to which the learner responds (e.g., in an online quiz). Truly interactive systems like games or simulations enable both learner-initiated and system-initiated activities. However, interactivity is still a far cry from the gold standard of teacher-like interaction, as most interactive systems adhere to a “one-size-fits-all” philosophy that does not take differences among learners into account. This has led to the development of more sophisticated, so-called adaptive learning technologies that are tailored to the current knowledge, the current workload, or the main interests of a learner (Brusilovsky & Peylo, 2003).

In order to be able to cater to the specific needs of a learner, adaptive learning systems need to form and constantly update some kind of model or representation of a learner. In other words, assessment of the current situation becomes a central feature of adaptive learning systems. There are various ways in which such variables can be assessed, e.g., by explicitly prompting learners about their current knowledge (Kalyuga & Sweller, 2005), or more implicitly through capturing the online behavior of learners through learning analytics (Siemens & Long, 2011) or educational data mining techniques (Romero & Ventura, 2007). The data gained from these forms of assessment can then be used to adapt to the current situation of a learner. For instance, intelligent tutoring systems like the cognitive tutor (Koedinger & Corbett, 2006) build a model of the current level of understanding that a learner has, and adapt feedback and support to this assessment. Another example of adaptivity is provided through recommender systems for learning (Buder & Schwind, 2012), which suggest information sources based on the inferred interests and tastes of a learner.

We believe that there are two classes of adaptive learning technologies, which we will refer to as “directive adaptivity” and “assistive adaptivity”. Directive adaptivity means that a technology tailors its behavior to the needs of a person without the person actually noticing it. For instance, when a learner interacts with an intelligent tutoring system and fails on a given task, the tutoring system might provide an easier task or problem in the next trial. While directive adaptivity can be highly sophisticated, it leads to a situation where self-regulation of a learner and external regulation by the technology might be in conflict. In contrast, a recommender system for learning would be an example of assistive adaptivity. It is based on assessments of a person and a situation, but preserves learner control and autonomy. Many recent technologies (not only in learning) are based on this assistive metaphor. They capture information that might be helpful or might be of interest to a person, and offer this information, either on demand or based on the current situation. Thus, assistive adaptivity creates communicative feedback loops with a learner, quite similar to the “dialogical” way an effective teacher would respond.

Summary

When discussing principles, ideas, and theories about effective designs, we argued that the interaction between teachers and learners is regarded as the gold standard of how to promote learning in informational environments. While the main goal of teaching is to help learners get a better understanding of the world surrounding them (cognitive perspective), we further argued that effective teaching also involves motivational-affective aspects, as well as social-interactive aspects. Consequently, many learning technologies try to address these three perspectives. The main goals of supporting the cognitive aspects of learning are to keep cognitive workload manageable, to ensure that declarative knowledge will be transformed into larger conceptual knowledge structures, and to foster self-regulation and metacognitive skills. In order to address the motivational-affective perspective, learning technologies should try to keep motivation high and evoke positive affect (e.g., through the introduction of game-like elements). Finally, in order to cater to social-interactive needs, technologies should either enable interpersonal, social contact with other learners (CSCL), or they should try to implement “teacher-like” features in the way the system interacts with a learner. At a minimum, this would involve interactivity, a constant interplay of activities that are initiated by the learner and by the technology. However, in order to truly achieve a teacher-like interaction, learning technologies should be adaptive and cater to the situational needs of a learner. This could be achieved through directive adaptivity where the technology adapts its behavior without the learner noticing the adaptation, or through assistive adaptivity where the technology adapts its behavior by providing the learner with relevant information just in time.

The chapters in the second part of the present book take a prescriptive-interventionist stance by trying to describe how informational resources in an informational environment should be designed to promote individuals to obtaining a better understanding of the world surrounding them, with a particular focus on learning scenarios. There we will take up many of the concepts discussed so far. Chapter 6 is heavily inspired by the cognitive perspective (promoting conceptual knowledge of teachers), and Chap. 7 builds on the motivational-affective perspective by investigating a game-based environment. Subsequently, Chaps. 8–12 all address the social-interactive perspective by providing various forays into the design of adaptive learning environments.

Outline of Part II of this Book (Effective Designs)

Chapter 6 (Voss et al., 2017) investigates the learning of school teachers and adult educators. This chapter reports on research with regard to the assessment and the promotion of teacher competences. More specifically, in the assessment part of the chapter it is investigated how digital video technologies can provide new avenues to measure the pedagogical-psychological knowledge of teachers (conceptual knowledge; cognitive perspective). The promotion part of the chapter describes the development of a learning and testing environment that employs annotated video cases of classroom interactions. This environment is based on principles of cognitive flexibility theory, thus providing an opportunity to transform declarative into conceptual knowledge (cognitive perspective) in an interactive fashion (social-interactive perspective).

While Chap. 6 focuses quite a lot on the cognitive perspective of learning, Chap. 7 (Soltanlou et al., 2017) is heavily inspired by the motivational-affective perspective. In particular, it is about a game-based learning environment that helps learners in an informal context to train their orthography and numeracy skills. The development of these skills (cognitive perspective) often requires a significant amount of training, and can easily become boring or tiresome. However, by designing a game-based training environment, motivation is likely to stay high. The chapter also addresses the notion of adaptivity (social-interactive perspective), by discussing how a game can match its difficulty level to the current skills of a learner, or how it can select a competing learner of a similar skill level for multi-player gaming. Finally, the chapter makes important contributions to learning assessments by tracking the neurocognitive changes that are invoked through the use of arithmetic learning environments.

Chapter 8 (Spüler et al., 2017) reports on the use of adaptive brain-computer interfaces in learning contexts. These interfaces try to assess the mental workload of a learner during task performance through online measurement of brain activity (electroencephalography). On the basis of these "neural signatures," the system then automatically adapts the difficulty level of arithmetic exercises to the assessed mental workload. As these adaptations take place without the learner noticing them, the brain-computer interfaces described in this chapter provide a nice example of directive adaptivity.

Chapter 9 (Scheiter et al., 2017) focuses on the support of self-regulated learning (cognitive perspective) through an adaptive multimedia environment (social-interactive perspective). The chapter describes various studies in which information processing of learners was captured through tracking of a learner's eye movements and through rapid, online assessments of current knowledge. Based on these assessments, the multimedia environment could then adapt to the situational needs of the learner. The system provided both directive adaptivity (automatic changes in the design of the learning environment) and assistive adaptivity (displaying instructional prompts that help to improve self-regulation skills).

Chapter 10 (Azevedo et al., 2017) introduces a conceptual framework that focuses on emotion regulation during learning (motivational-affective perspective). The chapter discusses various ways to capture online assessments of cognition and metacognition (e.g., eye tracking, log files, screen recordings) and various ways to capture emotions (e.g., galvanic skin responses, facial expressions). It is also discussed how an environment can be designed that captures all these data online and adaptively feeds them back to a learner in order to foster emotion regulation. Providing learners with a suite of data about their own cognitions and emotions is an example of assistive adaptivity.

Chapter 11 (Winne et al., 2017) has a unique focus in investigating technologies that support learners over longer stretches of time (up to several months) while they are traversing their informational environments. The concrete learning context presented in this chapter is the assignment to create a term paper, an activity that is referred to as information problem solving. The adaptive learning environment is based on nStudy, a tool that captures the complete and detailed history of a learner's online activities. Through learning analytics, aggregated data from nStudy can then be fed back to the learner in order to assist and support self-regulated activities over various stages of information problem solving (assistive adaptivity).

Last but not least, Chap. 12 (Graesser, Lippert, & Hampton, 2017) provides insights into the most literal implementation of systems that provide a teacher-like social interaction between learner and technology. This chapter focuses on conversational agents that have the capability to engage in natural language dialogues with a learner and/or other agents. In particular, it provides several examples of conversational agents who effectively support the development of so-called deeper learning. For instance, conversational agents can improve learners' abilities to ask "deep" questions (cognitive perspective), the ability to deal with cognitive conflict and disagreements (motivational-affective perspective), or the ability to solve a problem collaboratively (social-interactive perspective).

Together, the chapters that make up Part II provide various forays into how effective designs can help to address the cognitive perspective of learning, and cater to the motivational-affective and the social-interactive needs of learners.

References

Ainsworth, S. (2008). How do animations influence learning? In D. H. Robinson & G. J. Schraw (Eds.), *Recent innovations in educational technology that facilitate student learning* (pp. 37–67). Charlotte, NC: Information Age.

Aleven, V., Mclaren, B. M., Sewall, J., & Koedinger, K. R. (2009). A new paradigm for intelligent tutoring systems: Example-tracing tutors. *International Journal of Artificial Intelligence in Education*, *19*, 105–154.

Anderson, J. R. (1983). *The architecture of cognition*. Hillsdale, NJ: Lawrence Erlbaum.

Asch, S. E. (1951). Effects of group pressure on the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men* (pp. 177–190). Pittsburgh, PA: Carnegie Press.

Azevedo, R., Millar, G. C., Taub, M., Mudrick, N. V., Bradbury, A. E., & Price, M. J. (2017). Using data visualizations to foster emotion regulation during self-regulated learning with advanced learning technologies. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 225–247). New York, NY: Springer.

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford: University Press.

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics* (pp. 61–75). New York, NY: Springer.

Banaji, M. R., & Prentice, D. A. (1994). The self in social contexts. *Annual Review of Psychology*, *45*, 297–332.

Bartholomé, T., & Bromme, R. (2009). Coherence formation when learning from text and pictures: What kind of support for whom? *Journal of Educational Psychology*, *101*, 282–293.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, *31*, 445–457.

Brusilovsky, P., & Peylo, C. (2003). Adaptive and intelligent web-based educational systems. *International Journal of Artificial Intelligence in Education*, *13*, 159–172.

Buder, J. (2017). A conceptual framework of knowledge exchange. In S. Schwan & U. Cress (Eds.), *The psychology of digital learning: Constructing, exchanging, and acquiring knowledge with digital media* (pp. 105–122). Cham, Switzerland: Springer International Publishing.

Buder, J., Buttliere, B., & Cress, U. (2017). The role of cognitive conflicts in informational environments: Conflicting evidence from the learning sciences and social psychology? In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 53–74). New York, NY: Springer.

Buder, J., & Hesse, F. W. (2016). Designing digital technologies for deeper learning. In M. Spector, B. B. Lockee, & M. Childress (Eds.), *Learning, design, and technology: An international compendium of theory, research, practice, and policy*. New York, NY: Springer.

Buder, J., & Schwind, C. (2012). Learning with personalized recommender systems: A psychological view. *Computers in Human Behavior*, *28*, 207–216.

Campbell, W. K., & Sedikides, C. (1999). Self-threat magnifies the self-serving bias: A meta-analytic integration. *Review of General Psychology*, *3*, 23–43.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*, 233–246.

Clark, R. E., & Voogel, A. (1985). Transfer of training principles for instructional design. *ECTJ*, *33*, 113.

Cress, U., & Kimmerle, J. (2008). A systemic and cognitive view on collaborative knowledge building with wikis. *International Journal of Computer-Supported Collaborative Learning*, *3*, 105–122.

Cress, U., Moskaliuk, J., & Jeong, H. (Eds.). (2016). *Mass collaboration and education* (Vol. 16). Cham: Springer International Publishing.

D'Mello, S. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology*, *105*, 1082–1099.

dscout (2016, June 15). *Mobile touches: dscout's inaugural study on humans and their tech*. Retrieved from https://blog.dscout.com/mobile-touches

Engelmann, T., Dehler, J., Bodemer, D., & Buder, J. (2009). Knowledge awareness in CSCL: A psychological perspective. *Computers in Human Behavior*, *25*, 949–960.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Fischer, G. (2000). Lifelong learning—More than training. *Journal of Interactive Learning Research*, *11*, 265–294.

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

Fu, W.-T., & Pirolli, P. (2007). SNIF-ACT: A cognitive model of user navigation on the World Wide Web. *Human-Computer Interaction*, *22*, 355–412.

Graesser, A. C., Lippert, A. M., & Hampton, A. J. (2017). Successes and failures in building learning environments to promote deep learning: The value of conversational agents. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 273–298). New York, NY: Springer.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*, 555–588.

Hillmert, S., Groß, M., Schmidt-Hertha, B., & Weber, H. (2017). Informational environments and college student dropout. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 27–52). New York, NY: Springer.

Höller, J., Tsiatsis, V., Mulligan, C., Karnouskos, S., Avesand, S., & Boyle, D. (2014). *From machine-to-machine to the Internet of things: Introduction to a new age of intelligence*. Amsterdam: Elsevier.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Johnson, D., & Johnson, R. (1975). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Englewood Cliffs, NJ: Prentice-Hall.

Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology Research and Development*, *53*, 83–93.

Kimmerle, J., Bientzle, M., Cress, U., Flemming, D., Greving, H., Grapendorf, J., … Sassenberg, K. (2017). Motivated processing of health-related information in online environments. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 75–96). New York, NY: Springer.

Koedinger, K., & Corbett, A. (2006). Cognitive tutors: Technology bringing learning science to the classroom. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61–78). Cambridge, MA: University Press.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts—A conceptual analysis. *Educational Psychology Review*, *18*, 159–185.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, MA: Cambridge University Press.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098–2109.

Markoff, J. (2016). *Machines of loving grace: The quest for common ground between humans and robots*. New York, NY: Harper Collins.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63.

Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology*, *41*, 172–187.

Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. Cambridge, MA: University Press.

McNamara, J. (1982). Optimal patch use in a stochastic environment. *Theoretical Population Biology*, *21*, 269–288.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *2*, 1–21.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Miller, G. A. (1983). Informavores. In F. Machlup & U. Mansfield (Eds.), *The study of information: Interdisciplinary messages* (pp. 111–113). New York, NY: Wiley.

Murata, A. (2005). An attempt to evaluate mental workload using wavelet transform of EEG. *Human Factors*, *47*, 498–508.

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.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Olson, J. M., & Zanna, M. P. (1993). Attitudes and attitude change. *Annual Review of Psychology*, *44*, 117–154.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. 1, pp. 703–732). New York, NY: Wiley.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, *106*, 643–675.

Plass, J. L., O'Keefe, P. A., Homer, B. D., Case, J., Hayward, E. O., Stein, M., & Perlin, K. (2013). The impact of individual, competitive, and collaborative mathematics game play on learning, performance, and motivation. *Journal of Educational Psychology*, *105*, 1050–1066.

Pressey, S. L. (1926). A simple apparatus which gives tests and scores—And teaches. *School and Society*, *23*, 373–376.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, *21*, 1–29.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, *33*, 135–146.

Roscoe, R. D ., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutor's explanations and questions. *Review of Education Research*, *77*, 534–574.

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1–14). New Haven, CT: Yale University Press.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 97–118). New York, NY: Cambridge University Press.

Scheiter, K., Fillisch, B., Krebs, M.-C., Leber, J., Plötzer, R., Renkl, A., ... Zimmermann, G. (2017). How to design adaptive multimedia environments to support self-regulated learning. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 203–223). New York, NY: Springer.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, *46*, 30–32.

Skinner, B. F. (1961). Why we need teaching machines. *Harvard Educational Review*, *31*, 377–398.

Soltanlou, M., Jung, S., Roesch, S., Ninaus, M., Brandelik, K., Heller, J., ... Moeller, K. (2017). Behavioral and neurocognitive evaluation of a web-platform for game-based learning of orthography and numeracy. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 149–176). New York, NY: Springer.

Spiro, R. J., Collins, B. P., Thota, J. J., & Feltovich, P. J. (2003). Cognitive flexibility theory: Hypermedia for complex learning, adaptive knowledge application, and experience acceleration. *Educational Technology*, *43*, 5–10.

Spüler, M., Krumpe, T., Walter, C., Scharinger, C., Rosenstiel, W., & Gerjets, P. (2017). Brain-computer interfaces for educational applications. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 177–201). New York, NY: Springer.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.

Sweller, J., Van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296.

Turner, J. C., & Oakes, P. J. (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology*, *25*, 237–252.

van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, *20*, 95–99.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*, 197–221.

Voss, T., Goeze, A., Marx, C., Hoehne, V., Klotz, V., & Schrader, J. (2017). Using digital media to assess and promote school and adult education teacher competence. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 125–148). New York, NY: Springer.

Werquin, P. (2010, February). *Recognition of non-formal and informal learning: Country practices*. Paris: OECD.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Lawrence Erlbaum.

Winne, P. H., Vytasek, J. M., Patzak, A., Rakovic, M., Marzouk, Z., Pakdaman-Savoji, A., … Nesbit, H. C. (2017). Designs for learning analytics to support information problem solving. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 249–272). New York, NY: Springer.

Witzany, G. (Ed.). (2014). *Biocommunication of animals*. Dordrecht: Springer.

Wood, W., Lundgren, S., Ouellette, J. A., Busceme, S., & Blackstone, T. (1994). Minority influence: A meta-analytic review of social influence processes. *Psychological Bulletin*, *115*, 323–345.

Zurstiege, G., Zipfel, S., Ort, A., Mack, I., Meitz, T. G. K., & Schäffeler, N. (2017). Managing obesity prevention using digital media—A double-sided approach. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 97–123). New York, NY: Springer.

Chapter 2
Informational Environments and College Student Dropout

Steffen Hillmert, Martin Groß, Bernhard Schmidt-Hertha, and Hannes Weber

Introduction

Problems of informational environments are ubiquitous in life. Heterogeneous informational environments and aspects of lacking or distorted information play a major role for the generation of relevant social problems. This can be the case, for instance, if different environments—such as peers, institutions, and digital media or offline sources—present conflicting information to the individual, complicating cognitive adaption. In this paper we discuss the example of students who prematurely leave college. Student dropout from institutions of higher education has been a common phenomenon for many years. It has also been increasingly regarded as a problem for society as it typically entails negative individual and collective consequences. With regard to individual costs, potential qualification deficits and prolonged training careers are the most obvious risks. With regard to collective costs, a high degree of dropout means a misallocation of resources for the institutions of higher education, and these institutions do not adequately fulfill their qualification function in society.

We use the example of leaving college and conceptualize it as being characterized (also) by information problems. We argue, first, that deficits in the

S. Hillmert (✉)
Department of Sociology, University of Tübingen, Tübingen, Germany

Institut für Soziologie, Eberhard Karls Universität Tübingen, Wilhelmstr. 36, 72074, Tübingen, Germany
e-mail: steffen.hillmert@uni-tuebingen.de

M. Groß • H. Weber
Department of Sociology, University of Tübingen, Tübingen, Germany

B. Schmidt-Hertha
Institute of Education, University of Tübingen, Tübingen, Germany

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_2

individual processing of information represent an essential part of the process leading to dropout. Second, we look at the heterogeneity of relevant informational contexts. Specific institutions (such as different academic programs) may provide different environments, and different types of students may have different levels and configurations of information. The complexity of available information due to the interplay between different information contexts can lead to cognitive conflicts with detrimental outcomes since the environment can hardly be changed in our example (see Buder et al., Chap. 3 in this volume). Finally, information strategies may be actively used for promoting academic success. On the basis of this concept, we ask how specific aspects of informational environments and informational behavior mediate the process of academic dropout.

Our paper is structured as follows: We first provide a basic concept of college dropout as an individual process with a special emphasis on informational aspects. The following section discusses various information-related predictors of college dropout. We then test the derived hypotheses using specifically collected student data. After a brief description of our data base and empirical methods, we present and discuss our empirical results. We conclude with a number of practical suggestions.

Integration into College and Problems of Information

Following the seminal work of Tinto (1993), relevant determinants of college dropout are neither exclusively individual nor purely environmental. Rather, college careers can be regarded as status passages among different communities—e.g., between school and college but also between family and former peers and new personal networks of fellow students—where multiple problems of adjustment have to be solved. We can regard academic dropout as a process of increasingly lacking integration. Deficits become manifest in two major areas, referring to *intellectual* or *academic* integration and to *social* integration. Academic integration is achieved predominantly via academic performance. Negative feedback in the form of low grades or failed examinations reduces the level of individual confidence in academic success. Social integration is achieved through personal ties with peers. Lacking or unsatisfactory contacts with fellow students or relevant others will again decrease the subjective likelihood of success and increase the likelihood of dropout.

We can imagine having manifest indicators for integration in these two dimensions such as academic grades or measurements of personal friendship networks. While the lack of academic performance may be regarded as a legitimate reason for college dropout, this is much less the case for lacking social support. However, even the first aspect is not without ambiguity. For example, the personal level of achievement may in fact be unclear to individuals when they receive only diffuse feedback about their level of performance. An adequate processing of information is therefore a necessary link between individual academic potential and performance and the biographical consequences that follow from them. This example also

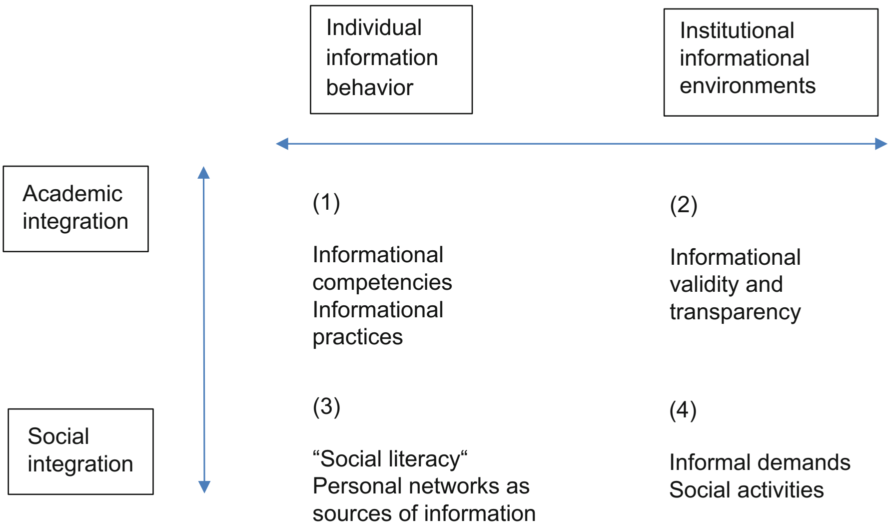

Fig. 2.1 Informational aspects relevant for college integration

suggests a high degree of *insecurity* that is typically involved in higher education, starting with basic issues of organization and including questions of biographical insecurity. We therefore focus on mediating processes between potential deficits in academic and social integration and the individual intentions or decisions to leave academic training.

As illustrated in Fig. 2.1, various aspects of information are involved in processes of integration into the college system and possible forms of disintegration. Ideal-typically speaking, we can distinguish two dimensions: on the one hand, institutional and individual characteristics, and, on the other hand, aspects relevant primarily for *academic* integration and aspects relevant primarily for *social* integration. Regarding the institutional side, not only the college system itself, but also specific educational institutions and even specific programs and courses represent specific informational environments. For individuals, information is a necessary resource for advancing their college careers. The institutional and the individual side are closely connected, and a high risk of dropout may result from a mismatch between these two sides of the ubiquitous information problems. A dichotomization of the two dimensions results in four ideal types of informational aspects that are related to the integration into college. Let us have a brief look at each of them:

1. *Individual aspects of academic integration*: Adequate informational competencies and practices are the basis for individual learning. In particular, they may compensate for existing deficits in individual knowledge.
2. *Institutional aspects of academic integration*: It is the core function of educational institutions to provide not only adequate substantive information (about the content of the subject or organizational details), but also to measure and

to certify individual performance. Of course, this implies that this information should be objective, reliable and valid, and the procedures should be transparent and comprehensible.

3. *Individual aspects of social integration*: Besides Tinto (1993), many sociologists have emphasized that focusing on academic achievement alone provides a very incomplete picture of school life, including aspects of individual advancement, status, and satisfaction. Rather, aspects of "adequate" behavior and inter-personal relationships play a crucial role (Bourdieu, 1996; Coleman, 1961). To be successful within a specific institution, individuals need to be competent in also correctly reading the signs of informal rules of behavior ("hidden curriculum"). Social origin is a well-known predictor for observable differences in this regard. There is no clear borderline towards academic achievement. Social resources that individuals can draw upon may again compensate for existing deficits in individual knowledge. For example, personal social networks may also be used as sources of information about academic questions.
4. *Institutional aspects of social integration*: Social integration is by no means a purely individual phenomenon, but the specific environments may appear to individuals as either more friendly or more hostile in social terms. In this sense they may pose informal demands on students that are not immediately obvious to them but that need to be decoded. In this sense, the lack of information can also be a decisive factor in mechanisms leading to college dropout. On the other hand, institutions of higher education increasingly offer activities that are explicitly designed to facilitate social integration—and hence, potential exchange of information—among their students, for example in the form of special freshmen weeks.

Determinants of Student Dropout (Intentions)

Subsequent to the analytical distinctions developed in the previous section, we now ask about the role of various predictors of dropout (intentions). Using comprehensive survey data (see the next section) we can cover several aspects of the conceptual model presented.

Informational Competencies and Behavior

In the current "information age" (Castells, 2011), students in higher education are expected to acquire appropriate knowledge by familiarizing themselves quickly with new information and evaluating it (Tippelt & Schmidt, 2006). This also involves the use of modern information and communication technology (ICT) for different study-related tasks such as doing research, writing, presenting, or using the computer for communication (Stauder, 2013). Depending on the theoretical background, terminology for skills related to such tasks include “media competence,” “information literacy,” or “digital competence” (for an overview see Gutiérrez & Tyner, 2011; Zylka, Müller, & Martins, 2011). We use the term “media competence” as it is often understood as an aspect of general communication competence that enables a person to orientate oneself in a mediatized world and to get to know the world actively with the assistance of media (Baacke, 1996). However, little is known about the relevance of media competence for student attrition. There is a strong consensus about the necessity of using digital media effectively (Hesse, Gaiser, & Reinhardt, 2006; Kerres & Voß, 2006; Schmidt-Hertha & Strobel, in press) and about the potential of digital media for learning (Meister & Meise, 2010). Studies have shown a regular increase in digital media use among students (e.g., the HISBUS studies carried out in Germany: Kleimann, Göcks, & Özkilic, 2008, or the ECAR study in the USA: Smith, Salaway, & Caruso, 2009). An interrelation between media competence and student dropout seems plausible when considering the significance of information research and the necessary evaluation of sources in many study programs (HRK, 2012). At the same time the spread of a broad range of online learning arrangements in higher education programs and online communication require a higher level of media literacy from both students and university staff (Schäffer, 2015). However, studies have also shown that the majority of students use online platforms and social media for making and sustaining social contacts, but to a much lesser extent for activities directly related to university such as learning (Madge, Meek, Wellens, & Hooley, 2009; Margaryan, Littlejohn, & Vojt, 2011).

Hence, the general use of digital media has to be distinguished from specialized skills relevant for university. Empirical studies have found that only the latter is correlated with academic success (Tien & Fu, 2008), while a general impact of ICT use on students’ learning is disputed (Cox & Marshall, 2007). Following Baacke (1997), we can distinguish between different facets of media competence, such as skills for using software or information resources related to the field of study, the ability to reflect critically on media content, and a general open-mindedness for digital media. Given the omnipresence of digital content in contemporary academic life, we expect study-related media competencies to be positively related to academic success (and thus negatively to dropout). This is also true with regard to the ability to critically evaluate content that, according to several studies, many undergraduates still lack (Timmers & Glas, 2010). No significant “digital divide” can be found within our sample with regard to the general use of digital media. For instance, nine out of ten respondents report to have a Facebook account and a similarly high proportion uses search engines such as Google to prepare for examinations or homework. A general rejection of ICT might therefore have negative effects on both social integration and the ability to find relevant information or to deal with study-related tasks that are organized online. On the other hand, we do not expect extensive use of digital media to have a positive effect on academic integration. As previous research suggests, spending a large amount of time online is usually not accompanied by an increase in study-related activities (Madge et al., 2009).

A related central concept in our analysis is information behavior, understood as strategic search for and evaluation of information (Wilson, 1999). Knowing where to look for information and being able to determine which information is valuable and whether it comes from trustworthy sources are central competencies required in academic life. Studies using standardized tests of information seeking competencies have shown that many students enter university with insufficient skills (Gross & Latham, 2012). Obviously many students are unaware of academic databases or search engines and any more advanced form of search queries. Evaluation-related competencies can manifest themselves, for example, in the ability to differentiate between information found in an academic journal as opposed to information found on a random blog. We look at two factors related to information behavior: Consultation of online versus offline sources when preparing a paper and self-reported behavior with regard to the evaluation of sources. We expect critical reflection on sources of information found on the internet to be positively correlated with academic integration and thus negatively with dropout. For types of sources, we cannot postulate a general superiority of online over offline sources, or vice versa. Moreover, today only few students rely exclusively on offline information. Therefore, doing research in the library or actively seeking advice from the lecturer in addition to using Google or Wikipedia may increase the chances of finding valuable information, which in turn may positively affect academic success.

However, there are probably significant differences between subject groups (Grosch & Gidion, 2011). In accordance with the theory of situated cognition (Greeno, 1998), we expect that there is a strong link between media-related skills and information strategies and specific contexts of application. For instance, learning environments are typically thought to be less standardized in humanities and social sciences as opposed to the natural sciences and medicine, which induces the need for "deep-learning" strategies (Baeten, Kyndt, Struyven, & Dochy, 2010). A simple Google search may thus suffice in one context while a more sophisticated search strategy is needed in another. The effectiveness of online versus offline search strategies can also be related to subject-specific differences on whether the relevant publications can mostly be found online, or, as in the humanities, still in books (Engels, Ossenblok, & Spruyt, 2012). This suggests that doing research online as opposed to going to the library does not have equal effects across all academic fields.

Informational Transparency and Sense of Fairness

The idea that the fairness of grading procedures should affect dropout intentions stems from the role of grades in determining the students' perceived probability of success. The better the grades a student receives, the greater the chances that they will at some point successfully graduate from university. Conversely, students who continue to receive exceedingly poor grades can take this as an indicator that their chances to succeed are too low to justify further investments, thus increasing the viability of alternative pathways.

However, it is not enough to view grades as a product. Rather, it is important to consider the procedures from which the grade resulted. At its core, the most basic function of the grading process is to take an input in the form of student knowledge and assign a quantitative measure to it. Thus, grades should ideally reflect student knowledge. But if the grading process is not fair, this function is impeded (Burger & Groß, 2016). The strategy of investigating the fairness of a procedure rather than the fairness of an outcome goes back to the works of Thibaut and Walker (1975) and Leventhal (1980) on procedural justice. We distinguish two aspects of procedural justice: control-related procedural justice and validity-related procedural justice.

Control-related procedural justice addresses the extent to which students can participate in the grading process. In the present context, *process control* implies that students are given a voice during the grading process, for example by deciding on the grading criteria together with the instructor. *Decision control* refers to influencing the grade itself (Colquitt, 2001). In turn, *correctability* means that students should have the option to appeal a grading decision they consider to be erroneous (Leventhal, 1980). The more students can get involved in the grading process, the greater their responsibility for the final results. The very existence of this possibility is enough to reassure students that they can still be successful even if they have trouble in future assignments. If, on the other hand, all power lies with the instructor, students can feel that they are kept from receiving the grades they themselves feel they deserve. Since this increases uncertainty regarding the probability of success, it is expected that low control-related procedural justice increases dropout intentions.

In contrast to that, *validity-related* procedural justice does not address questions of student involvement, but rather the extent to which grading procedures are capable of producing valid results. We define grading to be fair with regard to validity-related procedural justice if it fulfills three criteria: bias suppression, consistency, and accuracy. These criteria are important elements of Leventhal's (1980) definition of procedural justice. A procedure can be said to be *free of bias* if it is not guided by self-interest. Applied to the grading process, this means that instructors must not base their grading decisions on their personal sentiments toward individual students, regardless of whether this would result in better or in worse grades. *Consistency* requires that instructors judge the quality of a student's work according to dependable standards. This means that similar efforts should be awarded similar grades. *Accuracy* demands that instructors need to gather all the information necessary to make an informed decision with regard to the grade. The further grading procedures deviate from these ideals, the lower the match between student knowledge and grade. Since these aspects of instructor conduct cannot be influenced by the students, a lack of control-related procedural justice increases uncertainty with regard to the probability of success. Therefore, it is expected that it increases dropout intentions whereas high levels of validity-related procedural justice signal that students are likely to receive an appropriate return on their investments.

Social Integration and Personal Networks

It has long been acknowledged that social networks can be a crucial resource for the formation of human capital (Coleman, 1988). This is not only true for labor market outcomes (Granovetter, 1973), but also earlier in the life-course for educational success (Sacerdote, 2001; Zimmerman, 2003). For both secondary as well as tertiary education, studies have shown that interacting with high-achieving co-students can positively affect academic achievements (Hanushek, Kain, Markman, & Rivkin, 2003; Lavy, Paserman, & Schlosser, 2012; Lomi, Snijders, Steglich, & Torló, 2011). When preparing for examinations or writing homework, friends, roommates, or learning partners can, for example, explain difficult theories or recommend literature (Hasan & Bagde, 2013). This helps to close knowledge gaps and should be positively related to grades, which in turn are associated with lower risks of dropout.

Besides these peer effects, which are directly related to academic performance, social integration can also be important with regard to more informal aspects of college life. Being part of a larger social network increases the chances that information flowing through this network is received (Calvó-Armengol, Patacchini, & Zenou, 2009). This could be information on organizational issues, but also on, say, available jobs as a student assistant. Since the transition to university often involves relocating to another city, off-campus activities such as finding flats or hobbies are also easier in case of successful social integration. Thus, we expect personal networks to be associated with decreased chances of student dropout.

Further Determinants

In addition, several factors have long been established as determinants of student attrition. Studies show the relevance of intellectual capabilities, study motivation, and self-efficacy of students for success in higher education but also underline the meaning of fit between students and study program as a critical factor for student attrition (Heublein & Wolter, 2011; Pascarella & Terenzini, 1983; Kolb, Kraus, Pixner, & Schüpbach, 2006; Robbins et al., 2004; Sarcletti & Müller, 2011; Stinebrickner & Stinebrickner, 2014). In addition, social background has long been related to dropout rates (Bean, 1980; Wolter, Diem, & Messer, 2014). On the other hand, social background appears to be of lower relevance for success in tertiary education because selectivity with regard to parental education or status typically matters more in earlier stages and transitions of education (Hillmert & Jacob, 2010).

Socio-demographic and other factors need to be considered since they could also be correlated with our main variables of interest; for instance, media competencies might be stronger among high-achieving students. Besides, we focus on subject group-specific differences that may, as described above, matter for factors such as media use as well as for dropout rates. Differences in dropout rates may stem from, among other factors, differences in admission procedures between study courses. Where these are very selective, students typically show better performance (Delaney, Harmon, & Redmond, 2011) and lower dropout rates (Scherfer & Weber, 2014). Several pitfalls of research into student dropouts can also be related to differences in administrative regulations and practices between faculties and study programs. For instance, some study programs with no restrictions of admission are known to attract students who do not intend to graduate in the respective program, but rather intend bridging the time gap until their application for another course is decided upon. These types of dropout are usually not associated with the factors in our model and can thus complicate empirical analysis into the causes of dropout based on standardized surveys.

Data, Measurements, and Methods

We use data from *CampusPanel* (Burger, 2015; Lang & Hillmert, 2014) that we collected at a large German university. This online survey focused on various aspects of academic behavior as well as biographical information. Our sample consists of students from all academic programs at the Bachelor and Master level. Around 3,800 students participated in the first wave, carried out in the winter term of 2013/2014. Of those, around 700 took part in a follow-up panel survey in mid-2015.

Our measurement of dropout is twofold. First, we use dropout intentions as stated in the first wave of the survey as the main dependent variable for our analyses. We construct a scale consisting of three items (validated by confirmatory factor analysis). A list of all items along with descriptive statistics can be found in the Appendix. Second, we want to compare these findings to an analysis of actual dropouts instead of intentions. We therefore investigated whether first-wave participants were still active in their study course in summer 2015, around 1.5 years after the first survey. The panel participants provided us with the information of whether they were still enrolled, or had already graduated, dropped out, or changed subjects. In addition, we gathered administrative data from the university's register on some of the first-wave participants who did not participate in the second wave. This was only possible for a sub-sample of participants since not all gave their student e-mail addresses and agreed to be contacted again. In total, we have information on the status of 1,478 first-wave participants 1.5 years after the study. Among them, 127 left the university without a degree, another 102 changed their major subject but remained within the university, 150 graduated, and the remaining 1,099 were still enrolled in their previous study program. Here, we define dropout as leaving the university without a degree, and hence we have 127 "actual dropouts" in our sample. There are some ambiguities about this definition that necessarily arise when defining dropout. For instance, from the view of a particular study program, changing a subject within university could also be interpreted as a "dropout," whereas from the viewpoint of society, leaving college but subsequently enrolling at another college might not be viewed as dropping out since the respective person is still in tertiary education (as opposed to, say, the labor market). Since we have incomplete information on study course changes and current status after de-registration from university, we work with the definition given above.

Our central predictors are operationalized as follows. For information behavior, we asked participants to name the sources they consulted when preparing their most recent written assignment. Usage of Google or Wikipedia was classified as "internet sources" while going to the library or asking the lecturer for additional material was termed "traditional sources." In addition, we constructed an index "evaluation of information," which was made up of three items measuring self-reported behavior with regard to the evaluation of URLs or sources, and the date of information found online as well as the differentiation between facts and opinions. A factor "online activity" was constructed using questions about twitter usage, blog authorship, total hours spent online per day, and number of memberships in social networks. For "media competence," we asked respondents to rate their own skills in doing research with library catalogs, using online databases for literature, working with text-processing tools, and using online search engines. A factor "critical attitude towards media" was constructed to capture negative attitudes towards digital technologies in general. High values on this scale mean that respondents think digital communication is too impersonal, internet content is mostly expendable, or that they sometimes feel "swamped" by all the information found online or scared of technology in general. All indices were constructed using factor scores from principal components analysis and were scaled to have mean of zero and standard deviation of one. The two measures of justice were each measured on a three-item scale. "Control-related procedural justice" refers to the perception that professors give the opportunity to voice opinions on grades, and influence or object to them. By contrast, factor scores for "validity-related procedural justice" are high if respondents think their professors are unbiased, consistently use the same standards for grading, and give grades that best reflect students' state of knowledge. We furthermore included an index "satisfaction with performance" consisting of three items measuring overall satisfaction with one's own performance as well as comparisons with previous expectations. For "social integration," we asked whether participants felt they had successfully made contacts with other students, still retain those contacts, and whether they know classmates who they can discuss study-related questions with. As measures of "pre-enrollment information," students were asked how well they felt informed about study contents, requirements for examinations, and workload prior to enrolling in their present study program. Finally, we asked whether participants felt graduating was an important step to reach their goals in life (dubbed "educational aspirations").

A number of socio-demographic and other control variables were included in the analysis. The most recent grade obtained for a written assignment was used as a proxy for academic performance. Since many studies find that high school grades predict success in university, we also asked for the university entrance diploma (German *Abitur*) grade average. Gender, migration background, and parental education were included as socio-demographic variables. Migration background is operationalized through parents' country of birth and is coded 1 if one or both parents (or the respondents themselves) had migrated to Germany and 0 for all others. Parental education is a measure of whether or not respondents come from an academic family background and counts the number of parents with a degree in tertiary education. An important aspect is the number of semesters the respondent has already studied at the time of answering the questionnaire, since it is well known that dropouts occur more often at an earlier stage during studies. We also asked whether participants had completed an apprenticeship before going to college, i.e. whether they were previously enrolled in Germany's dual vocational training system. Finally, respondents were asked whether they already had an internship during their current study course. Having had work experience (either before or while studying) might affect dropout intentions because these students might have a better idea of what they could do instead if they give up on their studies.

We used ordinary least squares (OLS) regression models to estimate the impact of our predictors on dropout intentions and binary logit models to analyze actual dropouts. Missing data (with the exception of actual dropouts where the available sample was much smaller) were multiply imputed using the software MICE (Van Buuren & Groothuis-Oudshoorn, 2011) available for R (R Core Team, 2013). Each model was separately estimated in ten multiply imputed datasets and the results were averaged using Rubin's (1987) rules. In addition, we carried out split-sample analyses with regard to academic performance and subject group. For performance, a median-split was executed based on grades. Three subject groups are evaluated separately to explore whether the impact of our predictors differs by field of study. These are "language and cultural studies," "mathematics and natural sciences," and "law, economy, and social sciences," as defined by the German Statistical Office. While each of these aggregations still encompasses a heterogeneous group of study majors, the differences between these groups are arguably large enough to be able to uncover mediating effects by the learning environment.

Empirical Results

We start by giving results for *dropout intentions* and then compare the findings to predictors of actual dropout. Table 2.1 displays the regression results where groups of three of our main variables of interest are sequentially added to the models, while socio-demographic and other control variables are included in all models.

The first column (Model 1) introduces indices of information behavior as correlates of dropout intention. As expected, evaluation of information is negatively associated with dropout intentions. Thus, if students critically reflect on sources of information found while doing research on the internet, they are less likely to consider leaving university. This can be interpreted as showing an interrelation between the internalization of standards of academic practice and academic integration. On the other hand, it does not seem to make a difference whether students predominantly search online or offline for study-related information. Relying primarily on traditional (offline) sources has a small positive effect on dropout intentions, but this effect vanishes in the full model. Among the control

Table 2.1 Predictors of dropout intention—linear regression models

	Dependent variable: dropout intentions				
	(1)	(2)	(3)	(4)	(5)
Intercept	−.320***	−.259**	−.054	−.157	−.088
	(.091)	(.092)	(.088)	(.090)	(.080)
Information behavior					
Evaluation of information	−.204***				−.070***
	(.019)				(.018)
Internet sources	.002				−.013
	(.017)				(.015)
Traditional sources	.039*				.002
	(.019)				(.017)
Media competencies and use					
Media competence		−.092***			−.003
		(.016)			(.015)
Online activity		.210***			.125***
		(.017)			(.015)
Critical attitude towards media		.118***			.087***
		(.016)			(.015)
Justice perceptions					
Control-related procedural justice			−.019		−.026
			(.017)		(.016)
Validity-related procedural justice			−.156***		−.136***
			(.016)		(.015)
Satisfaction with performance			−.354***		−.273***
			(.017)		(.017)
Social integration and information					
Social integration				−.190***	−.106***
				(.016)	(.015)
Pre-enrollment information				−.145***	−.064***
				(.016)	(.015)
Educational aspirations				−.183***	−.134***
				(.016)	(.015)
Control variables					
Grades	−.110***	−.133***	−.028	−.120***	−.041
	(.024)	(.023)	(.027)	(.022)	(.024)
University entrance grade average	−.056**	−.046*	−.064***	−.023	−.027
	(.020)	(.019)	(.019)	(.020)	(.018)
Gender (female)	−.051	−.026	−.098*	−.020	−.088**
	(.040)	(.041)	(.040)	(.037)	(.034)
Migration background	.037	−.006	.007	.054	−.019
	(.047)	(.047)	(.048)	(.045)	(.042)

(continued)

Table 2.1 (continued)

	Dependent variable: dropout intentions				
	(1)	(2)	(3)	(4)	(5)
Parental education	.017	.0003	.003	.014	.020
	(.023)	(.023)	(.021)	(.023)	(.020)
Semesters studied	.031***	.031***	.018***	.014**	.018***
	(.005)	(.005)	(.005)	(.005)	(.004)
Apprenticeship before college	−.208***	−.220***	−.254***	−.175**	−.136**
	(.059)	(.059)	(.055)	(.057)	(.052)
Internship during studies	.159***	.158***	.165***	.131***	.078*
	(.040)	(.040)	(.039)	(.038)	(.034)
Observations	3816	3816	3816	3816	3816
R^2	.078	.104	.208	.164	.286

Note: Ordinary least squares regression coefficients (standard errors in parentheses) are presented. Subject group dummies are not shown. For coding of variables and data sources see Appendix
*$p < .05$, **$p < .01$, ***$p < .001$

variables, grades show negative correlations with dropout intentions in all models. High school grade average shows a negative effect as well in most models, albeit with a smaller effect size. Both variables are coded such that high values mean better grades; hence, the results suggest that better academic performance is associated with lower dropout intentions, as expected. Gender, migration background, and parental education are not significantly correlated with intentions to leave college. These intentions tend to become more pronounced the more time a respondent has already spent in his study program. Students who completed vocational training before coming to university appear to be more motivated to stay in their study program. By contrast, work experience through an internship during their studies increases the likelihood that participants consider dropping out of college.

In Model 2, predictors related to media competence and media use were introduced into the analyses while retaining the control variables of the first model. As expected, a positive self-assessment regarding ICT skills was associated with lower dropout intentions. By contrast, extensive use of online applications such as social media, twitter, and blogs was positively related to plans of leaving college. This finding surely requires further exploration, but it might show that usage of digital media is not per se associated with better academic integration, if it does not come with specific study-related ICT skills. Being opposed to media altogether, on the other hand, tends to be accompanied by higher dropout intentions as well.

Regarding the impact of justice perceptions, only validity-related procedural justice positively affects intentions to stay in university (see Model 3). That is, students who perceive the grading process to be fair and their professors to be unbiased are less likely to consider leaving college prematurely. Control-related procedural justice – the perception of being able to influence grades and voice objections—is apparently not important for the decision to drop out of a study program. Satisfaction with their own performance turns out to be the best predictor of students' dropout intentions. This factor mediates the effect of grades, which is reduced once satisfaction with performance is entered into the models. This suggests that even if grades are below average, but a student is still satisfied with his or her performance, dropout intentions are usually low. Generally, however, satisfaction with performance is greater if grades are better.

Finally, Model 4 shows a considerable effect of social integration on dropout intentions of the expected sign. This effect is independent of grades, i.e. contacts with classmates do not primarily lower dropout intentions through peer effects on academic performance, but can apparently rather be attributed to the non-academic benefits of social networks on campus. Dropout intentions are also lower if students felt well informed about their study course regarding contents, workload, and terms of examinations before enrolling. In addition, educational aspirations, i.e. the ascription of importance to obtaining a degree, tend to go hand in hand with fewer thoughts about leaving college.

Do these patterns differ by field of study? Figure 2.2 replicates the analyses shown in Table 2.1, full model (5), separately for three subject groups. Note that around 14% of students belong to neither of these groups (among them students of medicine) and were dropped for these analyses. The plots show effect size and 95% confidence intervals for each predictor on dropout intentions within the respective field of study. The same stepwise regressions were estimated as in Table 2.1 (i.e., only three factors of interest entered the analysis at a time together with all control variables; the plots show the combined results). Overall, the findings are surprisingly stable across subject groups. A few notable differences can be highlighted. For instance, media use and ICT competencies seem to be less strongly related to dropout intentions in law, economics, and social sciences compared with the other study majors. In particular, online activity has no effect at all for this group while it does for the other two. Grades also seem to matter less for students of law, economics, and social sciences (although high school grade average still exerts a significant effect). Most other factors, in particular information behavior, justice perceptions, and social integration are virtually identical in their effects on dropout intentions across fields of study.

We also explored whether our findings differ for high- versus low-performing students. For instance, one might expect justice perceptions with regard to the grading process to be less of a concern the better the individual's grades are. Similarly, a high level of social integration could be more important as a motivation to stay in university for students with lower academic success rates. However, as Fig. 2.3 shows, this is not the case. We did a median-split with regard to both grades and high-school grade average such that the upper panel of Fig. 2.3 only comprises above-average students with regard to both measures of performance, while the lower panel shows results for below-average performers only. As the results suggest, most of the effects found in the full sample equally hold for students with stronger or weaker performances. This is true for most of our factors of interest such as media competences, justice perceptions, social integration, and satisfaction with performance. With regard to evaluation of information, the coefficient is higher for low-performing students, suggesting they can benefit more from adopting strategies

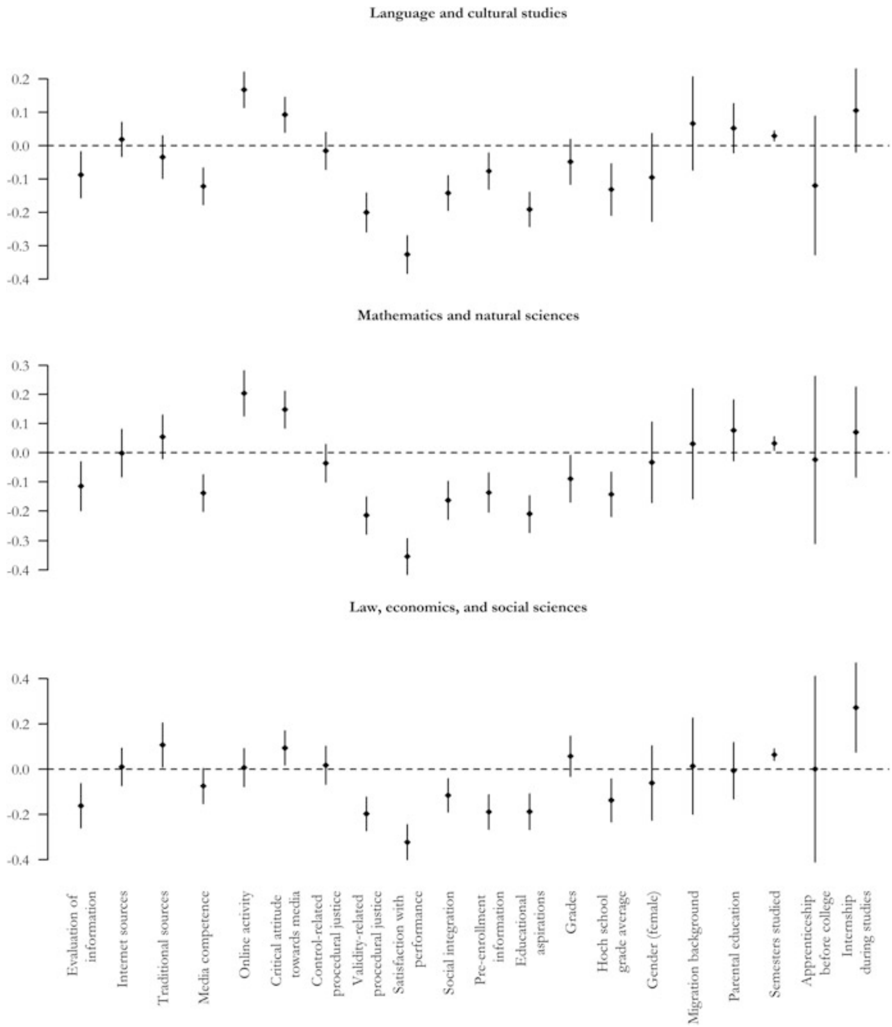

Fig. 2.2 Determinants of dropout intention by subject group. *Note*: Plot shows ordinary least squares estimates and 95% confidence intervals. For coding of variables, see Appendix

of critical reflection on sources of information. Higher performers, by contrast, can apparently be lured out of university with internships during studies, while this is not the case for other students.

Finally, we want to know whether our results hold not only for stated dropout intentions, but also for the prediction of actual dropouts. Table 2.2 replicates the analyses of Table 2.1 but with a binary dependent variable coded 1 for survey participants who had prematurely left college around 1.5 years after our survey. All in all, as the logistic regression results suggest, the prediction of actual dropouts

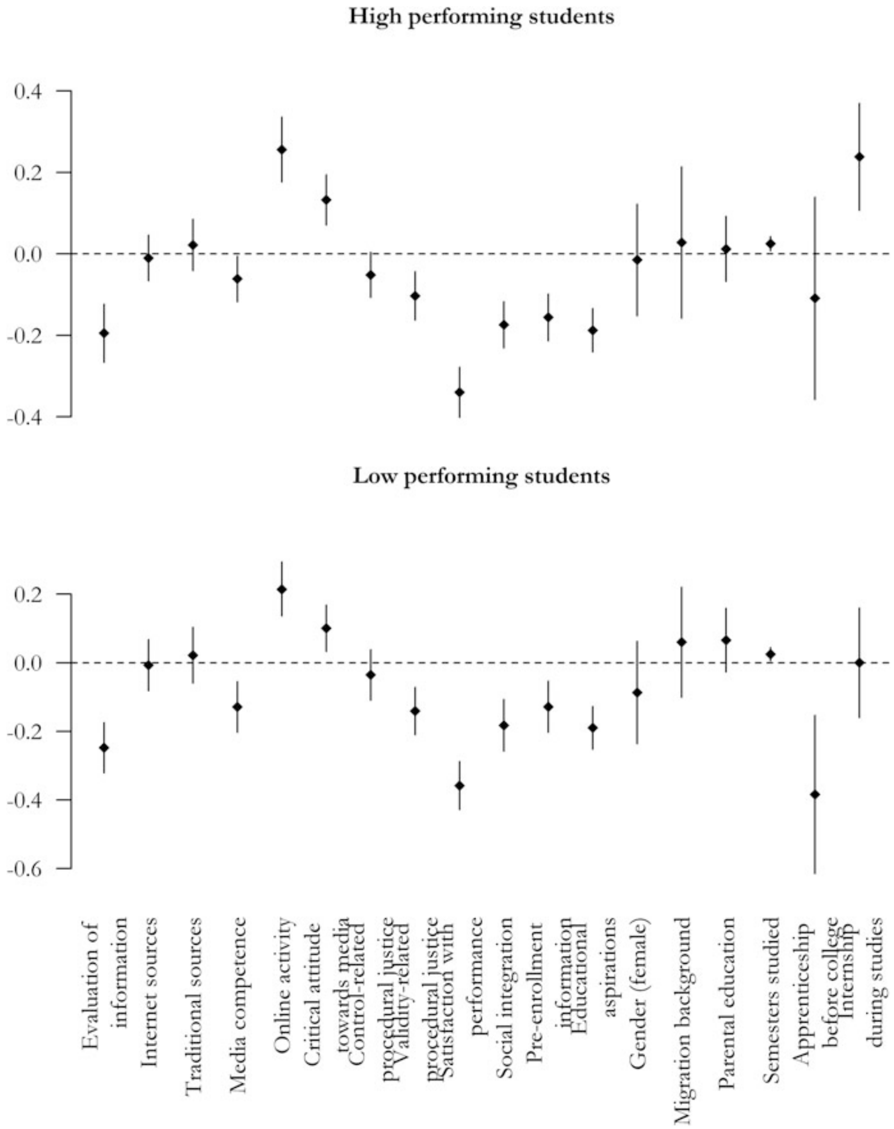

Fig. 2.3 Determinants of dropout intention by performance level. *Note*: Plot shows ordinary least squares estimates and 95% confidence intervals. For coding of variables, see Appendix

is much more difficult than the explanation of dropout intentions. Among the factors that were previously identified as important correlates of intentions to leave university, only a few are also significant predictors of dropouts. Most notably, perceptions of a fair grading process (validity-related procedural justice) and a high level of social integration are negatively associated with later dropout. We

Table 2.2 Predictors of actual dropout—logistic regression models

	Dependent variable: dropout				
	(1)	(2)	(3)	(4)	(5)
Intercept	−2.365***	−2.238***	−2.264***	−2.134***	−2.040***
	(.604)	(.604)	(.616)	(.606)	(.463)
Information behavior					
Evaluation of information	.118				.114
	(.122)				(.099)
Internet sources	.224*				.157*
	(.106)				(.079)
Traditional sources	.136				.073
	(.110)				(.084)
Media competence and use					
Media competence		−.193		.018	−.193
		(.105)		(.078)	(.105)
Online activity		−.296**		−.280***	−.296**
		(.096)		(.077)	(.096)
Critical attitude towards media		−.062		.059	−.062
		(.107)		(.086)	(.107)
Justice perceptions					
Control-related procedural justice			−.193		.018
			(.105)		(.078)
Validity-related procedural justice			−.296**		−.280***
			(.096)		(.077)
Satisfaction with performance			−.062		.059
			(.107)		(.086)
Social integration and information					
Social integration				−.186*	−.191**
				(.090)	(.073)
Pre-enrollment information				−.018	.014
				(.095)	(.076)
Educational aspirations				−.023	.026
				(.093)	(.077)
Control variables					
Grades	−.069	−.058	−.011	−.051	−.015
	(.132)	(.130)	(.144)	(.129)	(.106)
University entrance grade average	−.127	−.150	−.194	−.137	−.043
	(.110)	(.108)	(.112)	(.109)	(.086)
Gender (female)	.004	−.017	−.176	−.036	.121
	(.209)	(.208)	(.209)	(.205)	(.169)
Migration background	.082	.011	.055	.092	.105
	(.245)	(.251)	(.249)	(.246)	(.196)

(continued)

Table 2.2 (continued)

	Dependent variable: dropout				
	(1)	(2)	(3)	(4)	(5)
Parental education	.113	.123	.125	.117	.086
	(.129)	(.129)	(.129)	(.129)	(.101)
Semesters studied	−.162***	−.153***	−.175***	−.156***	−.117***
	(.043)	(.043)	(.045)	(.043)	(.030)
Apprenticeship before college	−.166	−.129	−.158	−.191	−.392
	(.398)	(.398)	(.399)	(.399)	(.330)
Internship during studies	−.174	−.222	−.170	−.287	−.010
	(.223)	(.221)	(.228)	(.221)	(.177)
Observations	1478	1478	1478	1478	1478
Log Likelihood	−413.75	−414.82	−406.10	−414.69	−601.54
Akaike Inf. Crit.	857.50	859.65	842.20	859.39	1251.09

Note: Logit models with binary dependent variable. Predictors are lagged 1 year. Subject group dummies are not shown. For coding of variables and data sources see Appendix *$p<.05$, **$p<.01$, ***$p<.001$

also find an effect from the usage of internet sources for the preparation of written assignments here, suggesting that relying (solely) on Google and Wikipedia for research might negatively affect academic integration. In addition, it is well known that many students quit university in an early phase of their studies, so it is not surprising that dropouts become less frequent the longer the respondents studied.

These findings point to the multi-faceted nature of dropouts. While low perceived fairness and poor social integration do predict later dropout behavior, there are probably many types of dropout that do not fit into the patterns we could cover with our survey questions. For instance, some students leave after having been accepted at another university, others are forcibly dropped from their course after failing a crucial test. In such cases, early warnings for dropout intentions need not necessarily have been visible in the initial survey. It should be noted that our reduced dataset for these analyses might not be representative for all dropouts, since half of our sample consists of panel participants. In particular, among those who could not be traced with administrative data and refused to be contacted again by our survey, former students who dropped out because of dissatisfaction or insufficient performances might be overrepresented. Furthermore, at the time of the first survey wave, many of those who failed very early in their studies might already have gone or been less motivated to take part in the survey.

Summary and Conclusions

Interpretation of Our Results and Recommendations

Informational environments can play an important role in the intention to leave college prematurely beyond academic performance and related factors, which have long been studied as predictors of student attrition. Our analyses indicate that even among students with good grades, poor social integration and perceptions of injustice regarding grading processes can induce dropout intentions and indeed lead to later dropout. Being well informed on the contents and requirements is also important for later motivation to continue in the chosen study course. These correlations can be found equally among students of the humanities, law, or social sciences as well as in the natural sciences.

Our findings also highlight the role of digital media use. Extensive online activities (i.e., the frequent use of Twitter, social networks, or blogs) do not contribute do higher academic integration per se, as our data suggest. Spending a lot of time online can in fact be associated with higher dropout intentions, a finding that calls for further investigation of the mechanisms involved. However, rejecting digital media altogether is also accompanied by poorer academic integration. Rather, possessing study-related media competencies and critically evaluating resources found on the internet are shown to be important in reducing dropout intentions.

Compared with the dropout intentions stated in our online survey, we find that actual dropouts as shown in the administrative records are harder to predict. This finding points to the diverse reasons for leaving university prematurely, many of which do not necessarily fit into the patterns of poor academic performance or low social integration. For instance, we saw that especially among well-performing students, work experience outside university as part of an internship can trigger motivation to leave college despite otherwise favorable views on the study course and life on campus in general. Many students also change their subject or continue their studies at another university for a variety of reasons. Informational environments can also play a role here, for instance through online research or information from social networks about specific contents of study courses elsewhere. Finally, many individual reasons that have their origins outside of university's reach (such as family or health-related issues) can influence motivation to stay in or leave college.

For faculty administrations and study programs that seek to understand and address dropout rates, we recommend paying attention to the multitude of facets the dropout phenomenon can have. This means that the specific situation of a study program must be taken into consideration. For instance, study programs with strict admission procedures are different from courses without entrance restrictions. Some programs are known to attract many students that ultimately seek to "bridge time" until, e.g., application processes for similar and more prestigious courses are decided upon. This situation obviously requires different measures compared with programs where most of the dropouts can be attributed to high performance requirements in examinations.

However, as we have demonstrated in our analyses, several factors are relevant for most fields of study. For instance, a poor level of information before enrollment is often followed by dropout intentions. This seems especially salient for the subject groups encompassing law, economics, and social sciences, but also for other major subjects. Quite often, students enter these programs without a full understanding of the requirements (e.g., with regard to mathematical skills) and contents of the course in question. Program directors may want to increase their efforts to communicate these contents to applicants or potentially interested high-school pupils. Putting great emphasis on fair and transparent standards of performance evaluations is also recommended since perceptions of unfair grading processes are a rather strong predictor of dropout intentions as well as actual dropouts.

Outlook: The Potential of (Digital) Media

Finally, as our findings show, media competencies can play a role in reducing dropout intentions. This seems to be the case for the natural sciences, but also for the humanities. To date, ICT skills are predominantly conveyed outside university. However, informational environments are not fixed. They can be actively changed, and electronic media play an increasing role in informational environments. Hence, there is considerable potential for the use of media in general and electronic media in particular in tackling problems of information in higher education contexts.

For example, the ever-increasing availability of books, journals, and other scholarly works on the internet should greatly facilitate research activities related to papers or examinations for students. With remote access to university library subscriptions from computers at home as well as from mobile devices, many offers are currently available anywhere and anytime. However, platforms such as JSTOR, SpringerLink, Google Scholar, or university library catalogs need to be better understood by students and they need to acquire the necessary skills to find and retrieve scientific content with these tools. The availability of sophisticated search engines and platforms specifically designed for academic purposes stands in contrast to prevailing research practices found among many freshmen (but also more advanced students) who often use rather simple search strategies. Moreover, with the increasing amount of information available online, being able to critically evaluate sources of information becomes more important. This is directly linked to learning about conventions prevalent in academia (e.g., citing scientific journals in essays rather than Wikipedia or random blogs), which can increase academic integration and hence prevent dropouts.

Appendix

Table 2.3 Definitions and descriptive statistics of the variables used in the analyses

Variable	Definition/item(s) used	Valid *N*	Mean	SD	Min	Max
Dropout	Administrative data: Participants who had left university without a degree at time of second wave	1,478	0.155	0.362	0	1
Grades	Grade received for last written assignment (reverse-coded: higher is better)	1,600	1.884	0.853	1.000	6.000
High school grade average	University entrance diploma grade average (“Abiturnote”, reverse-coded: higher is better)	2,904	2.056	0.609	1.000	3.800
Gender	0 = Male, 1 = Female	3,083	0.640	0.480	0	1
Parental education	Number of parents having tertiary education	2,143	0.859	0.826	0	2
Apprenticeship before college	Participant has completed an apprenticeship/ vocational training (“Ausbildung”) before going to university	3,062	0.092	0.289	0	1
Migration background	0 = Both parents were born in Germany, 1 = At least one parent was born abroad	2,542	0.191	0.393	0	1
Language and cultural studies	Major subject belongs into this group as defined by the German Statistical Office	3,816	0.343	0.475	0	1
Mathematics and natural sciences	Major subject belongs into this group as defined by the German Statistical Office	3,816	0.251	0.434	0	1
Law, economics, and social sciences	Major subject belongs into this group as defined by the German Statistical Office	3,816	0.168	0.374	0	1
Medicine	Major subject belongs into this group as defined by the German Statistical Office	3,816	0.100	0.301	0	1
Other subject group		3,816	0.036	0.186	0	1
Semesters studied	Number of semesters participant has been enrolled in current degree course at time of survey	3,304	4.235	3.758	0	50
Internship during studies	Already completed an internship during current course of studies (1 = No, 2 = Yes).	3,107	1.599	0.490	1	2

(continued)

Table 2.3 (continued)

Variable	Definition/item(s) used	Valid *N*	Mean	SD	Min	Max
Factor: evaluation of information	"While preparing my last written assignment I have: – paid attention to the URL/ source of an information to evaluate it – paid attention to the date when information I found on the internet was last edited – tried to differentiate between facts and opinions"	3,816	0.000	1.000	−2.958	1.881
Factor: internet sources	"For my last written assignment I: – used Google – used Wikipedia"	3,816	0.000	1.000	−2.755	1.202
Factor: traditional sources	"For my last written assignment I: – went to the library and did research there – consulted my lecturer"	3,816	0.000	1.000	−2.169	1.775
Factor: control-related procedural justice	"My professors give me the opportunity to express my views on the grading" "My professors give me the opportunity to influence the grades I am given" "My professors give me the opportunity to object to the grade"	3,816	−0.000	1.000	−2.239	3.093
Factor: validity-related procedural justice	"My professors consistently use the same standards for grading" "My professors are unbiased when grading" "My professors ensure that my grade best reflects the state of my knowledge"	3,816	0.000	1.000	−3.721	1.775
Factor: satisfaction with performance	"I am satisfied with my study performance" "I have fully met my expectations regarding my academic performance" "My achievements in university are better than initially expected"	3,816	0.000	1.000	−2.863	2.092
Factor: social integration	"I have a lot of contact with other students from my class" "I know many classmates whom I can discuss study-related questions with" "I managed to make good contacts with other students so far"	3,816	0.000	1.000	−2.837	1.497

Factor: pre-enrollment information	"How was your level of information before enrolling in your present degree course regarding: – Study contents – Requirements for exams – Study-related workload"	3,816	−0.000	1.000	−1.940	2.477
Educational aspirations	"Graduating is an important step to reach my goals in life"	3,108	5.856	1.334	1	7
Factor: online activity	"Have you ever used Twitter?" (1 = Never, 5 = several times a day) "Have you ever written in a blog?" (1 = Never, 5 = Several times a day" "How many hours do you spend online on an average day?" "How many social networks are you a member of?"	3,816	0.000	1.000	−1.094	4.928
Factor: media competence	"How do you rate your skills in the following domains: – Doing research with library catalogs – Using online databases for literature – Text-processing tools (Word, OpenOffice, LateX, etc.) – Using search engines on the internet"	3,816	−0.000	1.000	−4.660	1.984
Factor: critical attitude towards media	"Technology sometimes scares me" "Communication via computer and internet is too impersonal" "I sometimes feel swamped by the information flood" "Most offers on the internet are expendable"	3,816	−0.000	1.000	−2.144	3.433
Factor: dropout intention (log)	"I am seriously considering changing to another university" "I am seriously considering abandoning my course of studies" "I frequently thought about dropping out of university"	3,816	0.545	0.441	0.065	1.888

References

Baacke, D. (1996). Medienkompetenz–Begrifflichkeit und sozialer Wandel. In A. v. Rein (Ed.), *Medienkompetenz als Schlüsselbegriff* (pp. 122–124). Klinkhardt: Bad Heilbrunn.

Baacke, D. (1997). *Medienpädagogik*. Tübingen: Niemeyer.

Baeten, M., Kyndt, E., Struyven, K., & Dochy, F. (2010). Using student-centred learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness. *Educational Research Review*, *5*(3), 243–260.

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education*, *12*(2), 155–187.

Bourdieu, P. (1996). *The state nobility: Elite schools in the field of power*. Stanford, CA: Stanford University Press.

Burger, R. (2015). *CampusPanel user handbook: Documentation for the Student Panel of the ScienceCampus Tübingen (wave 'b')*. Tübingen: University of Tübingen.

Burger, R., & Groß, M. (2016). Gerechtigkeit und Studienabbruch. Die Rolle der wahrgenommenen Fairness von Benotungsverfahren bei der Entstehung von Abbruchsintentionen. *Zeitschrift für Erziehungswissenschaft*. doi:10.1007/s11618-016-0672-8

Calvó-Armengol, A., Patacchini, E., & Zenou, Y. (2009). Peer effects and social networks in education. *The Review of Economic Studies*, *76*(4), 1239–1267.

Castells, M. (2011). *The rise of the network society: The information age: Economy, society, and culture* (Vol. 1). Chichester: Wiley.

Coleman, J. S. (1961). *The adolescent society*. Glencoe: The Free Press.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*, S95–S120.

Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, *86*(3), 386–400.

Cox, M. J., & Marshall, G. (2007). Effects of ICT: Do we know what we should know? *Education and Information Technologies*, *12*(2), 59–70.

Delaney, L., Harmon, C., & Redmond, C. (2011). Parental education, grade attainment and earnings expectations among university students. *Economics of Education Review*, *30*(6), 1136–1152.

Engels, T. C., Ossenblok, T. L., & Spruyt, E. H. (2012). Changing publication patterns in the social sciences and humanities, 2000–2009. *Scientometrics*, *93*(2), 373–390.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist*, *53*(1), 5.

Grosch, M., & Gidion, G. (2011). *Mediennutzungsgewohnheiten im Wandel: Ergebnisse einer Befragung zur studiumsbezogenen Mediennutzung*. Karlsruhe: KIT Scientific Publishing.

Gross, M., & Latham, D. (2012). What's skill got to do with it? Information literacy skills and self-views of ability among first-year college students. *Journal of the American Society for Information Science and Technology*, *63*(3), 574–583.

Gutiérrez, A., & Tyner, K. (2011). Media education, media literacy and digital competence. *Comunicar*. doi:10.3916/C38-2011-02-03

Hasan, S., & Bagde, S. (2013). The mechanics of social capital and academic performance in an Indian college. *American Sociological Review*, *78*(6), 1009–1032.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, *18*(5), 527–544.

Hesse, F. W., Gaiser, B., & Reinhardt, J. (2006). e-teaching.org: Das Lehren mit digitalen Medien lernen. In K. Solbach & W. Spiegel (Eds.), *Entwicklung von Medienkompetenz im Hochschulbereich: Perspektiven, Kompetenzen und Anwendungsbeispiele* (pp. 55–70). Düsseldorf: kopaed.

Heublein, U., & Wolter, A. (2011). Studienabbruch in Deutschland. Definition, Häufigkeit, Ursachen, Maßnahmen. *Zeitschrift für Pädagogik*, *57*(2), 214–236.

Hillmert, S., & Jacob, M. (2010). Selections and social selectivity on the academic track: A life-course analysis of educational attainment in Germany. *Research in Social Stratification and Mobility*, *28*(1), 59–76.

HRK. (2012). *Hochschule im digitalen Zeitalter: Informationskompetenz neu begreifen—Prozesse anders steuern*. Bonn: Hochschulrektorenkonferenz. Retrieved from https://www.hrk.de/uploads/media/Entschliessung_Informationskompetenz_20112012.pdf

Kerres, M., & Voß, B. (2006). Kompetenzentwicklung für E-Learning: Support-Dienstleistungen lernförderlich gestalten. In K. Solbach & W. Spiegel (Eds.), *Entwicklung von Medienkompetenz im Hochschulbereich: Perspektiven, Kompetenzen und Anwendungsbeispiele* (pp. 35–54). Düsseldorf: kopaed.

Kleimann, B., Göcks, M., & Özkilic, M. (2008). *Studieren im Web 2.0. Studienbezogene Web-und E-Learning-Dienste*. Hannover: HIS GmbH. Retrieved from https://hisbus.his.de/hisbus/docs/hisbus21.pdf

Kolb, M., Kraus, M., Pixner, J., & Schüpbach, H. (2006). Analyse von Studienverlaufsdaten zur Identifikation von studienabbruchgefährdeten Studierenden. *Das Hochschulwesen*, *54*(6), 196–201.

Lang, V., & Hillmert, S. (2014). *CampusPanel user handbook: Documentation for the Student Panel of the ScienceCampus Tübingen (wave 'a')*. Tübingen: University of Tübingen.

Lavy, V., Paserman, M. D., & Schlosser, A. (2012). Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal*, *122*(559), 208–237.

Leventhal, G. S. (1980). What should be done with equity theory? In K. Gergen, M. Greenberg, & R. Willis (Eds.), *Social exchange: Advances in theory and research* (pp. 27–55). New York, NY: Plenum Press.

Lomi, A., Snijders, T. A., Steglich, C. E., & Torló, V. J. (2011). Why are some more peer than others? Evidence from a longitudinal study of social networks and individual academic performance. *Social Science Research*, *40*(6), 1506–1520.

Madge, C., Meek, J., Wellens, J., & Hooley, T. (2009). Facebook, social integration and informal learning at university: 'It is more for socialising and talking to friends about work than for actually doing work'. *Learning, Media and Technology*, *34*(2), 141–155.

Margaryan, A., Littlejohn, A., & Vojt, G. (2011). Are digital natives a myth or reality? University students' use of digital technologies. *Computers & Education*, *56*(2), 429–440.

Meister, D. M., & Meise, B. (2010). Emergenz neuer Lernkulturen: Bildungsaneignungsperspektiven im Web 2.0. In B. Herzig, D. M. Meister, H. Moser, & H. Niesyto (Eds.), *Jahrbuch Medienpädagogik 8* (pp. 183–199). Wiesbaden: VS Verlag.

Pascarella, E. T., & Terenzini, P. T. (1983). Predicting voluntary freshman year persistence/withdrawal behavior in a residential university: A path analytic validation of Tinto's model. *Journal of Educational Psychology*, *75*(2), 215.

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261–288.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *The Quarterly Journal of Economics*, *116*(2), 681–704.

Sarcletti, A., & Müller, S. (2011). Zum Stand der Studienabbruchforschung. Theoretische Perspektiven, zentrale Ergebnisse und methodische Anforderungen an künftige Studien. *Zeitschrift für Bildungsforschung*, *1*(3), 235–248.

Schäffer, D. (2015). *E-Learning als Teil des persönlichen, intentionalen Lernraumes im Studium. Eine explorative Studie an Studierenden an der Fakultät für Erziehungswissenschaft an der Universität Bielefeld*. Berlin: epubli GmbH.

Scherfer, M., & Weber, H. (2014). Methoden zur Analyse von Studienabbruch und-wechsel am Beispiel der Abbrecherstudie der Universität Stuttgart. *Qualität in der Wissenschaft (QiW)*, *1*(2014), 17–22.

Schmidt-Hertha, B., & Strobel-Dümer, C. (2014). Computer literacy among the generations: How can older adults participate in digital society? In G. K. Zafiris & M. N. Gravani (Eds.), *Challenging the 'European area of lifelong learning'* (pp. 31–40). Netherlands: Springer.

Smith, S. D., Salaway, G., & Caruso, J. B. (2009). The ECAR study of undergraduate students and information technology 2009. ECAR. Retrieved from http://www.educause.edu/Resources/TheECARStudyofUndergraduateStu/187215.

Stauder, A. (2013). 2012 survey of the preservation, management, and use of audiovisual media in European higher education institutions. *OCLC Systems & Services*, *29*(4), 218–234.

Stinebrickner, R., & Stinebrickner, T. (2014). Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics*, *32*(3), 601–644.

Tien, F. F., & Fu, T. T. (2008). The correlates of the digital divide and their impact on college student learning. *Computers & Education*, *50*(1), 421–436.

Timmers, C. F., & Glas, C. A. (2010). Developing scales for information-seeking behaviour. *Journal of Documentation*, *66*(1), 46–69.

Thibaut, J. W., & Walker, L. (1975). *Procedural justice: A psychological analysis*. Hillsdale, NJ: Erlbaum.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition*. Chicago: University of Chicago Press.

Tippelt, R., & Schmidt, B. (2006). Zur beruflichen Weiterbildungs-und Erwachsenenbildungsforschung: Forschungsthemen und Trends. In *Datenreport Erziehungswissenschaft 2006* (pp. 81–100). VS Verlag für Sozialwissenschaften.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3).

Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation*, *55*(3), 249–270.

Wolter, S. C., Diem, A., & Messer, D. (2014). Drop-outs from Swiss universities: An empirical analysis of data on all students between 1975 and 2008. *European Journal of Education*, *49*(4), 471–483.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, *85*(1), 9–23.

Zylka, J., Müller, W., & Martins, S.W. (2011). Media literacy worldwide. Similarities and differences of theoretical approaches. 2011 IEEE Global Engineering Education Conference (EDUCON). doi:10.1109/EDUCON.2011.5773219.

Chapter 3
The Role of Cognitive Conflicts in Informational Environments: Conflicting Evidence from the Learning Sciences and Social Psychology?

Jürgen Buder, Brett Buttliere, and Ulrike Cress

Introduction

How people navigate through their informational environments, what information they seek, and how they relate to what they find in their informational environment depends on external and internal factors. External factors such as institutionalized rules and regulations have an influence on what information can be accessed, and how external information is framed has an influence on the behavior of individuals (Hillmert, Groß, Schmidt-Hertha, & Weber, 2017; Zurstiege et al., 2017). However, at the same time the use of informational resources is dependent on internal factors such as cognitions, attitudes, motivations, and emotions (Kimmerle et al., 2017). The present chapter focuses on an internal factor that we believe can explain to some extent how people navigate their informational environments. In particular, we suggest that the processing of information from an environment depends on the amount of cognitive conflict that a person experiences. A cognitive conflict arises when people experience a discrepancy or mismatch between their own cognitive structures and pieces of information they encounter in their informational environment—in short, "when the world tells you that you are wrong."

The chapter is structured as follows: we begin by providing a definition of what cognitive conflicts are (and what they aren't). Then we summarize research and theory on cognitive conflicts from two academic fields where they are frequently investigated: (1) in the learning sciences, and (2) in social psychology. Thereafter, we report on some of our own conceptual and empirical studies that investigated the role of cognitive conflict within social media environments (in Wikipedia, and

J. Buder (✉) • B. Buttliere • U. Cress
Leibniz-Institut für Wissensmedien (IWM), Schleichstr. 6, 72076, Tübingen, Germany
e-mail: j.buder@iwm-tuebingen.de

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_3

in online discussion forums). In particular, the link between cognitive conflict and production of content will be examined. Finally, we integrate findings from the learning sciences and from social psychology, and attempt to explain some of the different patterns of results that can be found, and how they might be reconciled.

Cognitive Conflicts

In this chapter we cover two different research fields (learning sciences and social psychology). Unfortunately, terms like "conflict" or "cognitive conflict" are used quite differently across these fields. Therefore, a definition for the purposes of this chapter is required. Our working definition of cognitive conflict is as follows: Cognitive conflict occurs when an individual perceives a discrepancy or mismatch between one's own knowledge or attitudes and the knowledge or attitudes that the individual encounters in the environment; this discrepancy or mismatch may involve quantitative or qualitative differences. This definition is admittedly quite abstract, thus requiring further elaboration. Therefore we will flesh out the definition by discussing four of its constituents.

The first constituent of our working definition is that cognitive conflicts take place in the heads of individuals, an assumption that can also be found in similar conceptualizations (e.g., disequilibrium, Piaget, 1970; cognitive dissonance, Festinger, 1957). Cognitive conflicts are perceived discrepancies that do not necessarily reflect objective discrepancies. There can be cases where discrepancies are objectively present, but will not be perceived. For instance, people have a general tendency to see members of an outgroup as highly similar and homogeneous (Mullen & Hu, 1989). Conversely, people might exaggerate discrepancies between themselves and other groups (e.g., Robinson, Keltner, Ward, & Ross, 1995). Thus, cognitive conflicts are subjectively constructed and colored.

The second constituent of our working definition refers to the parties involved in a conflict. In our conceptualization of informational environments we have two important stakeholders, an individual and a physical or digital environment that surrounds this individual (Buder & Hesse, 2016, 2017). Therefore, in principle one could conceptualize cognitive conflicts as perceived discrepancies within an individual, as perceived discrepancies in the environment, or as perceived discrepancies between individual and environment. While psychologists also investigate discrepancies within an individual (e.g., Kleiman & Hassin, 2013) or discrepancies within the environment (e.g., McGurk & MacDonald, 1976), this chapter focuses on discrepancies between person and environment (Piaget, 1970), with the emphasis that other persons are part of the environment (Zimmerman & Blom, 1983). This focus is most closely connected to the scope of informational environments.

The third constituent of our working definition refers to the object of a conflict. When defining a conflict as a discrepancy or mismatch, one may ask what this discrepancy or mismatch is about. In principle, a person and the environment can differ with regard to a myriad things. For instance, two persons might differ with regard to age, gender, ethnicity, personality, or mood. However, our definition explicitly focuses on conflicts that are cognitive in nature. In this regard, the current chapter is about conflicts that refer to mental representations of individuals. In particular, the two objects of conflict under scrutiny are knowledge discrepancies and attitudinal discrepancies. We have focused on these two objects of conflict because they represent very common conceptualizations in the respective research fields that this chapter covers. In the learning sciences, cognitive conflicts typically are conceptualized with regard to knowledge. For instance, both the Piagetian (Piaget, 1970) and the Vygotskyan (Vygotsky, 1978) research traditions suggest that learning may occur when a discrepancy between the knowledge of an individual and the knowledge in the environment (which could be another person) arises. In social psychology, cognitive conflicts often arise when people encounter attitudes and opinions in their environment that differ from their own attitudes (Hart et al., 2009). The distinction between knowledge and attitudes is not always clear-cut. By and large, it can be said that knowledge refers to pieces of information for which truth or falsity can be demonstrably shown (e.g., two different solutions to a math problem). In contrast, some issues can be characterized by very low demonstrability (e.g., the question of whether the Beatles or the Rolling Stones were the "better" band in the 1960s), and when different attitudes and viewpoints clash, the ensuing conflicts often lead to hot debate. However, there are cases where the line between knowledge and attitudes blur. For instance, scientific discourse is often based on demonstrable facts (e.g., experimental results), but people can get into heated debate on how to interpret and contextualize these facts.

The fourth constituent of our definition refers to quantitative versus qualitative differences. Quantitative discrepancies are about amounts and sizes. For instance, two students might differ in the amount of learning material that they have mastered (quantitative knowledge discrepancy), or they might have similar political attitudes, but differ in extremity (quantitative attitude discrepancy). In contrast, qualitative discrepancies are more fundamental in nature, as they refer to different kinds of knowledge or attitudes. For instance, two students might arrive at different solutions for a math problem (qualitative knowledge discrepancy), or they might have diametrically opposed political attitudes (qualitative attitude discrepancy).

In sum, our working definition is constructed from four constituents. First, cognitive conflicts take place in individual heads and involve the subjective perception of a discrepancy. Second, the perceived discrepancy arises between a person and the environment (which might include other persons). Third, cognitive conflicts could arise with regard to discrepancies about knowledge, or with regard to different attitudes and opinions. And fourth, cognitive conflicts can occur with regard to quantity (where the amount of knowledge or the strength of an attitude matters) or with regard to quality (where different pieces of knowledge or different attitudes clash). In the following, we sketch research on cognitive conflicts from the vantage point of two different disciplines, beginning with the learning sciences.

Cognitive Conflicts: Findings from the Learning Sciences

As learning can be conceived of as the acquisition and appropriation of knowledge, it comes as little surprise that cognitive conflicts in the learning sciences are primarily about discrepancies in knowledge. In order to discuss several types of conflict that learning scientists are interested in, consider the following four examples: (1) two people meet who have very a different level or amount of knowledge: this could be a parent and a child, a teacher and a student, an expert and a novice, or two learners with different ability levels. (2) An individual encounters new information in the environment (e.g., reading that Canberra is the capital of Australia); (3) two students arriving at different solutions for a given math task; (4) a philosopher and a psychologist discussing how the human mind operates. Each of these four examples involves some kind of discrepancy. Moreover, each of these examples can be used as a proxy for different concepts and theories in the learning sciences. Therefore, here we discuss the four different conceptualizations in turn.

The first example (e.g., teacher-student interaction) involves different levels of knowledge, thus it refers to the quantitative aspect of knowledge that was part of our definition of cognitive conflicts. This example can best be associated with the cultural-historical perspective of learning proposed by Vygotsky (1978). According to Vygotsky, all learning can be conceived of as an appropriation of cultural values and traditions, and culture is transmitted by means of social interaction with other people. From this basic idea, it follows that all learning takes place on a social level first and foremost before it becomes internalized by a person. Learning on a social level often unfolds in interaction with people who have a higher level of culturally acquired knowledge (e.g., mother and child) and/or who play a more central role in a community (e.g., employer and employee). This discrepancy in knowledge is a driving force behind learning, and at the same time the discrepancy perpetuates cultural values: learners acquire the knowledge or the explanations received from interaction partners with higher knowledge, and they grow into a cultural community. This growth begins by "legitimate peripheral participation" (Lave & Wenger, 1991), the performance of simple tasks within the community (e.g., a boy who helps his mother to wash the dishes). Over time, the person grows into a more central role within the community. One important addition to Vygotsky's cultural-historical approach is the notion of a "zone of proximal development," defined as the difference between the things that a learner can do without external guidance and the things that can be accomplished when scaffolded by a more advanced member of the community. A corollary of this assumption is that not all group compositions of learners are equally well suited. This has led to an investigation of group compositions, most notably by Webb (1991). She analyzed the performance of students at different knowledge levels in terms of communication behavior and learning outcomes and reported that, overall, groups with heterogeneous ability levels outperformed groups with homogeneous ability levels. Homogeneous groups of learners with low ability often lacked the skills to provide adequate explanations to their peers. In contrast, homogeneous groups of high ability learners also did not provide many explanations, presumably because they thought that their peers already knew how to solve a problem. When groups were heterogeneous with respect to ability (e.g., high and medium; low and medium), the less advanced group members asked more questions, and the more advanced group members provided more explanations, taking up some kind of teacher-student relationship. Thus, there is some empirical evidence suggesting that cognitive conflicts with regard to quantitative differences in knowledge give rise to better learning than an absence of cognitive conflicts.

The second example of a knowledge discrepancy involves an individual who encounters new information in the environment. Let us suppose a person who does not know what the capital of Australia is (quantitative discrepancy), or a person who believes that it must be Sydney (qualitative discrepancy), and who subsequently is informed that the capital is Canberra. These types of discrepancies can best be linked to Piagetian accounts of learning (Piaget, 1970). Piaget suggested that learning involves adaptive changes to cognitive schemas. Moreover, he distinguished between two types of adaptive forces. One of these forces, termed assimilation, entails a conservative stance of trying to apply one's existing schemas to new information in the environment and changing the schemas as little as possible. Assimilation provides continuity to knowledge structures, thus enabling learners to derive meaning from an informational environment. Both the learner who did not know the capital of Australia and the learner who erroneously thought that it was Sydney would experience the new information about Canberra as a cognitive conflict. In both cases, however, the information could be assimilated by making slight adjustments to one's knowledge, like adding the element "Canberra" to a schema of "Australia," or by updating a schema (replacing "Sydney" with "Canberra"). While there is some debate about the question of what actually counts as assimilation in a Piagetian sense (Block, 1982), a majority of scholars in the learning sciences would probably agree that adding new elements to an existing cognitive schema can be regarded as learning through assimilation, although this type of learning is not particularly "deep," elaborated, or fundamental.

The third example above (about two learners arriving at different solutions for a math task) addresses these "deeper" forms of learning. As a math task is reasonably demonstrable, it can be concluded that when two learners arrive at different outcomes, at least one of them must have been clearly wrong. Provided that the wrong solution of a learner was based on a misconception (e.g., about how to manually multiply two numbers), such a qualitative discrepancy between person and environment is difficult to reduce by ignoring the environment. Rather, the learner needs to fundamentally transform his or her cognitive schemas. Piaget (1970) refers to these transformative changes and re-organizations of knowledge structures as accommodation. Accommodation is the counterpart to assimilation, and while assimilation is about fitting the environment to one's knowledge, accommodation involves re-shaping one's understanding of the world by adapting one's knowledge to the external world. These qualitative cognitive conflicts between person and environment ultimately give rise to learning and development, as without it a person would be caught in "perseverative solipsism" (Block, 1982). Accommodation is an effortful process, but it is the hallmark of what learning scientists refer to as "learning" proper: an opportunity for cognitive growth. Consequently, large parts of educational psychology and the learning sciences are inspired by Piagetian thought about accommodation. For instance, conceptual change was a widely studied phenomenon in the 1980s (Posner, Strike, Hewson, & Gertzog, 1982), and this research led to insights into the learner-related conditions leading to accommodation (e.g., dissatisfaction with an existing conception, intelligibility and plausibility of a new concept) as well as into the environmental conditions leading to accommodation (e.g., anomalies, analogies, and metaphors).

The fourth and final example for knowledge discrepancies was about people with different scientific viewpoints on an issue, such as a philosopher discussing the human mind with a psychologist. Like the previous three examples, this is a discrepancy of knowledge, as both researchers can probably back up their view based on accrued evidence and knowledge within their respective fields. Moreover, this is certainly an example of a qualitative discrepancy, yet it is clearly different from the third example where at least one learner must have been demonstrably wrong. Rather, the philosopher and the psychologist both bring very different perspectives into a problem. This multiperspectivity also is addressed by the learning sciences, particularly in the socio-cognitivist view of the so-called "Geneva school" (Doise & Mugny, 1984). While Piaget mainly focused on the individual, scholars from the Geneva School combined Vygotskyan and Piagetian views by suggesting that one of the best ways to facilitate learning is to engender interpersonal (socio-cognitive) conflict between peers who have different viewpoints. As different viewpoints might just rely on different types of knowledge and expertise, the ensuing socio-cognitive conflict creates a disequilibrium in all participants. And since participants might be motivated to learn from each other, or even to co-construct a new, joint understanding that integrates multiple perspectives, opportunities for assimilation and accommodation for all members arise. Consequently, creating different viewpoints and having students work on the ensuing conflicts has become a central feature in many approaches of the learning sciences. For example, it was empirically shown that cognitive conflict among different viewpoints leads to better perspective taking (Tjosvold & Johnson, 1977), and to better learning results (Smith, Johnson, & Johnson, 1981). Consequently, several cooperative learning methods (O'Donnell & O'Kelly, 1994) explicitly build on learners who differ in their knowledge and expertise (jigsaw classroom; Aronson, Blaney, Stephin, Sikes, & Snapp, 1978), who will disagree in their viewpoints (constructive controversy; Johnson, Johnson, & Tjosvold, 2000), or who take up different roles during interaction (reciprocal teaching by Palincsar & Brown, 1984; scripted cooperation by O'Donnell, 1999). All these approaches make use of knowledge discrepancies between viewpoints, largely regarding them as a source for efficient learning (see also Graesser, Lippert, & Hampton, 2017).

In the previous paragraphs we have outlined four different theoretical strands in the learning sciences that all refer to cognitive conflict as implied by our working definition: the commonality among all four approaches is that conflict is conducive to learning and development, thus making it something to be embraced by teachers and learners alike. Consequently, many practical approaches in the learning sciences emphasize the creation or exploitation of cognitive conflicts, whether this is accomplished through group compositions, through creating conditions for conceptual change, or through the implementation of various viewpoints or various roles in interpersonal interaction. As we will see, such a positive overall evaluation of cognitive conflicts is in stark contrast to the viewpoints of social psychology.

Cognitive Conflicts: Findings from Social Psychology

According to our definition, cognitive conflicts are discrepancies between the "stakeholders" of an informational environment, in other words between person and environment. Equipped with this notion, the way that social psychologists investigate these types of discrepancies is predominantly about attitudes. An attitude can be defined as a subjective evaluation of an "attitude object" (Olson & Zanna, 1993), with the latter typically being a physical object ("ice cream"), a person ("John F. Kennedy"), a group ("New York Yankees"), or an abstract concept ("individual freedom"). Attitudes can either have a positive or a negative valence, depending on the type of evaluation. Moreover, many social psychologists agree that attitudes have affective components ("I like ice cream"), cognitive components ("I agree that John F. Kennedy was a good President"), and behavioral components ("I would never go to a Yankees game"), the so-called tripartite model (Rosenberg & Hovland, 1960). Although we already mentioned that the differentiation between knowledge and attitudes is a little blurry, there is reason to believe that attitudes have a lower demonstrability than knowledge—as attitudes always involve a subjective interpretation and evaluation, they are difficult to "prove" or "disprove." Therefore, discrepancies between person and environment cannot be easily resolved by some kind of objective "proof." When people have not established an attitude about an object, a quantitative discrepancy between person and environment may arise, and it is not unlikely that the person will build an attitude that is in line with the information from the environment (similar to Piagetian assimilation). However, the far more interesting case occurs when there is a qualitative discrepancy between person and environment. Many social psychological paradigms investigate how people process congenial information (information that is in line with a pre-existing attitude) versus uncongenial information (information that runs counter to the pre-existing attitude). In the following, we highlight some typical results from research about congenial and uncongenial information by looking at three areas: (1) selection of information, (2) interpretation and evaluation of information, and (3) memory for information. Relevant theoretical constructs for each of these three areas are discussed.

The first area refers to the type of information that people select from their environment. This is the main question in research on selective exposure (Frey, 1986; Hart et al., 2009). In a typical selective exposure study, the attitude of a person is captured. Subsequently, participants are given a free choice to expose themselves to material (texts, audio recordings) that is either congenial or uncongenial, and which type of material is preferred is measured. Hart et al. (2009) conducted a meta-analysis of 91 selective exposure studies containing 300 experimental conditions that largely followed the experimental procedure described above. The results clearly demonstrate a congeniality bias, indicating that people have a marked tendency to select and expose themselves to congenial information and to stay away from uncongenial information. This is very much in line with the everyday observation that most people prefer to read or listen to news that is compatible with one's worldview (such as NY Times for liberally-minded, and Fox News for conservatively inclined people). Although there are countless moderators of this effect, the general consensus is that people exhibit a preference for congenial over uncongenial information.

How can this congeniality bias in selective exposure be explained, and what are moderators of the bias? Selective exposure research has its roots in Festinger's (1957) account of cognitive dissonance. Festinger postulated that people are averse to conflicting (dissonant) pieces of information and therefore will strive to reduce this dissonance. In fact, one of the predictions of dissonance theory is that people will seek congenial (consonant) information and avoid uncongenial (dissonant) information, which paved the way for research on selective exposure (Frey, 1986; Hart et al., 2009). While there is strong support for an overall uncongeniality bias (Hart et al., 2009), research on selective exposure has also uncovered a number of important moderators. For instance, Hart et al. (2009) referred to a distinction between defense motivation and accuracy motivation. When people are defense-motivated, they strive to validate their worldview, and exhibit a clear congeniality bias, particularly when the contentious topic touches on important personal values, when the quality of congenial information is high, when people feel committed, or when they are close-minded. However, people can also be accuracy-motivated, for instance when they are rewarded for finding a "correct" answer. Under these circumstances, the congeniality bias disappears or may even turn into an uncongeniality bias, as the exposure to information will be determined by the perceived utility of information (both congenial and uncongenial) to achieve one's goals.

A second area of social psychological research on cognitive conflicts refers to the question of how people interpret and evaluate congenial and uncongenial information (see also Kimmerle et al., 2017). Typical studies on this issue are quite similar to studies on selective exposure. However, rather than requesting participants to select one or several pieces of information out of many others, participants are asked to assess the quality of both congenial and uncongenial pieces of information. Experiments like these typically show that congenial information is rated as higher in quality than uncongenial information, an effect that was termed biased assimilation (Lord, Ross, & Lepper, 1979), or evaluation bias (Greitemeyer & Schulz-Hardt, 2003; Mojzisch, Grouneva, & Schulz-Hardt, 2010). A somewhat similar effect can be found in studies where participants are asked to test hypotheses, showing that the questions that people ask in order to find out about the personality of a target person will be chosen to confirm rather than to reject their hypotheses

(Snyder & Swann, 1978). This style of hypothesis testing that is driven by positive (confirmatory) questions is referred to as confirmation bias (Klayman & Ha, 1987). Similar to the congeniality bias in selective exposure, biased assimilation, evaluation bias, and confirmation bias all indicate that congenial information is treated preferentially.

How can these biases be explained from theory? Answers to these questions can be derived from so-called dual-process theories of persuasive communication (Chaiken, 1980; Petty & Cacioppo, 1979). These theories are typically concerned about the way that mildly uncongenial information influences attitudes. The term "dual-process theories" was coined as both the Elaboration Likelihood Model (ELM; Petty & Cacioppo, 1986) and the Heuristic-Systematic Model of Persuasion (Chaiken, 1980) suggested that there are two types of attitudes change. When people do not care very much about an issue, or if their attention is distracted, they will only superficially process a persuasive message. In this case, information processing is typically influenced more by contextual cues and by heuristics than by the content of the message itself. For instance, one and the same (uncongenial) message leads to stronger attitude change if it purportedly comes from an expert than from a layperson. However, if people are sufficiently motivated and undistracted, they process a persuasive message more deeply, and they attend to the quality of the message content rather than to heuristic cues. The ELM suggests that this scrutiny will then lead to the generation of favorable or unfavorable thoughts about the message. Messages of high quality will lead to more favorable thoughts and more subsequent attitude change than messages of low quality. Combining theoretical constructs from dual-process theories with the findings that people tend to rate the quality of congenial messages to be higher than the quality of uncongenial messages (evaluation bias), it can be concluded that congenial messages will generally lead to more favorable thoughts (and more attitude change) than uncongenial messages. In other words, the evaluation bias might be due to the occurrence of favorable thoughts while processing a message. In fact, there is even empirical evidence showing that uncongenial information is scrutinized longer and more deeply than congenial information (Ditto & Lopez, 1992; Edwards & Smith, 1996). However, this bias does not occur because people are open to read and receive uncongenial information, but rather as an attempt to thoroughly disconfirm and debunk the uncongenial view! Like with selective exposure, interpretation and evaluation of conflicting information show clear patterns of a congeniality bias; the seemingly inconsistent finding that uncongenial information is scrutinized longer than congenial information is generally held to be an indicator that people will go to some lengths to defend their existing worldview.

The third area of social psychological research on cognitive conflicts concerns the question of how people remember congenial and uncongenial information. A typical study investigating memory for attitudinal information displays congenial and uncongenial arguments to participants and later asks for a memory test (free recall or recognition of arguments). For a long time, social psychologists believed that congenial information is better remembered than uncongenial information. To test for this contention, Alice Eagly and colleagues conducted a meta-analysis comprising 65 studies on memory for attitudinal information (Eagly, Chen, Chaiken, & Shaw-Barnes, 1999). They reported an overall congeniality bias indicating that congenial information is better remembered than uncongenial information. However, this effect was much smaller than initially assumed, and it became even weaker in more recent and rigorous studies on the effects on memory. Apparently, factors other than congeniality might play a more important role for memory effects (e.g., overall extremity of a position; Judd & Kulik, 1980).

While empirical results on memory for congenial and uncongenial information are apparently more complex than results for selective exposure and selective interpretation, they are examples of attempts to marry social psychological research on attitudes with basic research from cognitive psychology. A prime example of this theoretical trend is provided by the work of Russell Fazio and colleagues. Fazio revolutionized attitude research in the 1980s by suggesting a distinction between explicitly stated attitudes (until then the predominant type of attitudes that were measured) and implicit attitudes. From that time on, it became more common to assess attitudes by way of priming experiments, measuring reaction times rather than explicitly asking participants to assess their attitude. Fazio and colleagues (e.g., Fazio, Powell, & Herr, 1983) predicted that the strength of an attitude can best be measured by accessibility: the stronger an attitude is, the faster a person will react to an attitude object, and this faster reaction time will also be a strong predictor of how an attitude will relate to actual behavior. In addition to introducing the notion of accessibility of attitudes, Fazio's team also made predictions about how attitudes are formed, maintained, and strengthened over time. One prediction is that direct experience with an attitude object will increase attitude strength (Fazio & Zanna, 1981). To reframe this prediction in the context of this chapter, one can assume that frequent exposure to congenial information is likely to increase attitude strength, a finding that was supported by empirical investigations (mere exposure effect; Zajonc, 1968). A second prediction arriving from the accessibility theory is that frequent expression of an attitude will also strengthen the attitude (Fazio, Chen, McDonel, & Sherman, 1982). Put differently, when a person repeatedly expresses his or her (congenial) view, the corresponding attitude will become stronger (Brauer, Judd, & Gliner, 1995). While the accessibility theory does not directly address issues like processing of uncongenial information or attitude change, some predictions can be derived. If attitude strength is resting on the frequency of exposure to congenial information and on the frequency of expression of congenial information, attitude change might arise from the frequency of exposure and the frequency of expression of uncongenial information. Actually, there is some evidence that this might the case. Allport (1954) already suggested that frequent encounter with conflicting viewpoints might lead to more moderate attitudes (contact hypothesis), a prediction that was later qualified by the finding that the contact with another group must be positively connoted (Desforges et al., 1991). Similarly, the case can be made that when participants are "nudged" to make uncongenial statements, their attitudes will shift towards the uncongenial position. Indeed, one of the most effective ways to induce attitude shifts in the dissonance literature was to request participants to write counterattitudinal (uncongenial) essays. In sum, the accessibility theory relates attitudes to more basal cognitive reactions, and it suggests that attitude strength is a function of the frequency with which congenial information was encountered and expressed. By the same token, it can be assumed that attitude change must be a function of the frequency with which uncongenial information is encountered and expressed.

How can these empirical findings and theoretical accounts from social psychology be summarized and integrated? First of all, lots of different empirical investigations have shown that people have a tendency to prefer congenial information over uncongenial information. The reasons for this phenomenon are multi-faceted, but related to each other: dissonance theory suggests that cognitive conflicts are regarded as unpleasant, thus leading to a preference for congeniality. Similarly, it can be concluded that unless specifically prompted, people have a tendency to exhibit defense motivation, which would also explain conflict avoidance. Moreover, congenial information is likely to be more accessible in memory because it was encountered and expressed more often in the past. As a consequence, it might be that a systematic evaluation of congenial information leads to more favorable thoughts and stronger persuasion than the systematic evaluation of uncongenial information. While theoretical accounts in social psychology have also outlined some ways of how these biases can be overcome (e.g., creating accuracy motivation, frequent encounter and expression of uncongenial views), the general thrust is that people are not easily swayed to change their attitudes, and that they generally do not use conflicting information in a very open-minded manner.

Learning Sciences and Social Psychology: Contradictory of Complementary?

At first glance, one might get the feeling that the findings from the learning sciences and from social psychology about cognitive conflicts are diametrically opposed to each other. In the learning sciences, discrepancies between person and environment are clearly positively connoted, as conflicts are regarded as a harbinger of cognitive development and growth. Only through the acknowledgment of discrepancies and the subsequent re-organization of our cognitive schemata, only through guidance from more sophisticated learners, and only through encountering different viewpoints do we learn. In contrast, the picture that social psychology paints on cognitive conflicts appears to be rather bleak and negative: people tend to avoid conflicting, uncongenial information; they rate uncongenial information lower, and remember conflicting information slightly worse. How can these apparently contradicting views across disciplines be reconciled? One way to talk about the differences would be to assume that conflicts in the learning sciences and in social psychology are different. Looking back at our working definition of cognitive conflicts, one could argue that cognitive conflicts in the learning sciences are mainly about knowledge, whereas cognitive conflicts in social psychology are mainly about attitudes. One could further argue that knowledge-related conflicts are embraced, whereas attitude-related conflicts are avoided. However, we believe that there is not enough evidence for such an assumption: cognitive conflicts in the learning sciences are also regarded as highly conducive to learning when the boundaries between knowledge and attitudes blur (such as in research on cognitive controversies, or on interdisciplinary debates).

However, we would suggest that the apparent differences in the findings of the learning sciences and social psychology have to do with different types of investigations, and that, in fact, the diverse views might be reconcilable. When learning scientists talk about the positive side of cognitive conflicts, they typically investigate situations where conflicts are actually "played out" among stakeholders. For instance, they pair up learners with different levels or different types of expertise in a real, communicative setting that involves mutual influence among stakeholders. In contrast, many social psychological studies care more about subjective construals of cognitive conflicts. Rather than having participants actually partake in a mutual, communicative setting they typically investigate participants as relatively passive recipients of (uncongenial) information. These different foci will lead to different research results, but the underlying mechanisms might actually be quite similar. Put simply, it might be the case that cognitive conflicts are slightly averse for many individuals, but ultimately might be beneficial for them. In order to better compare learning sciences and social psychology, one might look at subjective construals of cognitive conflicts in learning situations, and at actual communicative mutuality in situations where different attitudes are the main concern. Indeed, there is evidence in the learning sciences that subjective construals of cognitive conflicts might be slightly averse (just as in a selective exposure study)—learners are not always highly motivated to revise their cognitive schemata, as low ability beliefs, or low willingness to expend cognitive effort are a frequent occurrence in classroom settings (Legault, Green-Demers, & Pelletier, 2006). Conversely, some social and organizational psychologists have also emphasized the potentially beneficial role that certain types of conflicts may have (Schulz-Hardt, Jochims, & Frey, 2002; Schwenk, 1990). However, most social psychological studies on cognitive conflicts do not investigate communicative scenarios that are the hallmark of the to-and-fro of real-life, interpersonal conflicts. If a person disagrees with a text (like in selective exposure studies), the only option is to ignore it. However, if a person were afforded an opportunity to talk back or otherwise communicate with dissenting others, ignoring others is only one option—another is to actively argue against others. In other words, behavioral outcomes might change whether one looks at the "consumption" versus the production of information. This thought is the starting point of our own work on cognitive conflicts, which we discuss below.

Cognitive Conflicts in Social Media

As alluded to in the previous section, it might be the case that a different picture about cognitive conflicts emerges when the focus of research does not only look at the consumption of information, but also looks at how cognitive conflicts impact the production of information. A natural test-bed to investigate such conflicts between person and environment is provided by social media, such as Facebook, Twitter, online discussion forums, or wikis. A commonality among these very different outlets is that they do not merely regard participants as "consumers" of information, but fundamentally rest on the notion that participants will provide information that becomes public to others. With their combination of consumption and production ("prosumption"; Ritzer, Dean, & Jurgenson, 2012), social media have introduced a major shift in how informational environments are constituted (Buder & Hesse, 2017).

The dual role of prosumption is captured and expanded into the co-evolution model of social media and knowledge construction (Cress & Kimmerle, 2008; Kimmerle, Moskaliuk, Oeberst, & Cress, 2015). This model combines Piagetian thought with a systems-theory perspective. The co-evolution model describes the interaction between two self-organized, coupled, and interdependent systems: a cognitive system of an individual person, and a social system (e.g., the Wikipedia as a collectively built artifact, with its own sets of norms and rules). This distinction maps nicely with the distinction between person and environment made in the current chapter. For each of these two systems, the other system constitutes its environment. The flow of information among these two systems can be further described in a way that the externalization of one system might pave the way for an internalization of the other system. For instance, a person might read something on Wikipedia, and learn something along the way, which may either result in adding new concepts (internal assimilation) or fundamentally re-structuring cognitive schemata (internal accommodation). Conversely, a person might change the content of Wikipedia by externalizing his/her knowledge, either making small additions (external assimilation) or by re-structuring the content of a Wikipedia entry (external accommodation). This mutual adaptation of person and environment leads to a co-evolution of both systems. Finally, the co-evolution model makes quite precise predictions about cognitive conflicts. A conflict between a system and its environment leads to an incongruity that calls for actions to bring both systems back into balance with their respective environments. If there is no incongruity between person and environment (e.g., because a person already knows the content of a Wikipedia entry), this will not be associated with any productive activity. Conversely, if the discrepancy between the two systems is too high, the systems will potentially ignore or reject the input made to them. Therefore, medium-sized incongruities between person and environment should result in the highest amount of activity and will lead to the highest likelihood that a system will produce an externalization.

Predictions of the co-evolution model were tested in a series of laboratory experiments. A study by Moskaliuk, Kimmerle, and Cress (2009) varied quantitative knowledge discrepancy between a person (cognitive system) and environment (wiki as a social system). Sixty-one participants individually received ten pieces of information about the causes of schizophrenia, with some pieces arguing for biological causes and others arguing for social causes. Depending on the experimental condition, participants were confronted with either an empty wiki (high quantitative conflict), a wiki that already contained all ten arguments (low quantitative conflict), or a wiki that contained four arguments (medium incongruity). How many words were added to the wiki (external assimilation), and how many structural edits were made in the wiki (external accommodation) were then measured. Partially in line with the co-evolution model, the conditions with medium conflict and high conflict yielded more added words (assimilation). More importantly, and completely in line with the co-evolution model, medium-sized conflicts yielded a larger number of edits than low conflicts or high conflicts (accommodation). This indicates that cognitive conflicts in social media can, at least to some extent, be associated with higher productivity in terms of generated content. These results were also confirmed by a study in which the wiki content was kept constant, but the amount of knowledge that the individual held was varied across three conditions (Kimmerle, Moskaliuk, & Cress, 2011). The wiki contained eight information pieces about the causes of schizophrenia (four in favor of biological causes, and four in favor of social causes). Participants either received no information (high quantitative conflict), information pertaining to all eight arguments (low quantitative conflict), or information pertaining to only the four biological or the four social causes (medium incongruity). External assimilation was higher in the medium-sized conflict condition and the low-conflict condition than in the high-conflict condition. Mirroring the effect of the previous study, medium-sized conflicts generated the highest number of structural edits (external accommodation), significantly more than in the other conditions. While these studies provided some support for the co-evolution model, they also were prone to alternative explanations, as medium-sized quantitative conflicts were perfectly confounded with large-sized qualitative conflicts. In the 2009 study, participants with a balanced view were confronted with a lop-sided wiki in the medium conflict conditions; in the 2011 study, participants with a lop-sided view were confronted with a balanced wiki in the medium conflict conditions. Two further experiments tried to address this issue (Moskaliuk, Kimmerle, & Cress, 2012).

In the first of these experiments participants initially received information that gave them a lop-sided view on the causes of schizophrenia. In a low-redundancy condition, the wiki contained only counter-arguments. In a high-redundancy condition, the wiki contained the counter-arguments and the arguments that the participant held. There was also a medium-redundancy condition where the wiki contained all counter-arguments, and half of the arguments that the participant held. External assimilation was strongest in the low-redundancy condition (low quantitative, but high qualitative conflict). With regard to external accommodation, the medium-redundancy condition (medium-sized quantitative, and medium-sized qualitative conflict) outperformed the other two conditions.

The second experiment compared qualitative conflict among two conditions. In both conditions, the wiki contained four arguments, and the person was in possession of four arguments, hence quantitative discrepancies were held constant. However, qualitative differences were varied: in a low qualitative conflict condition, both the wiki view and the person's view were balanced with respect to the causes of schizophrenia. However, in a high qualitative conflict condition, the wiki view was diametrically opposed to the view of the person. The high-conflict condition yielded much higher activity than the low-conflict condition, both with respect to external assimilation and accommodation.

In sum, the four studies presented here showed that when providing participants with an opportunity to produce information, experiencing some amount of cognitive conflict led to higher productivity than experiencing no amount of conflict. The set of studies also addressed the distinction between quantitative and qualitative conflicts used in this chapter. With regard to quantitative conflicts, medium-sized incongruities led to the highest amount of externalizations (particularly accommodation activities). In other words, people are most likely to transform the structure of a wiki if the wiki is neither empty nor complete. With regard to qualitative conflicts, high incongruities led to the highest amount of externalizations (both assimilation and accommodation). In other words, the larger the discrepancy between person and environment, the more likely it is that the person will change wiki content. In any case, a general tendency for conflict aversion (either quantitative or qualitative) that seems so pervasive in social psychology could not be found. It should be noted that the difference between, say, a selective exposure study and the wiki studies is not so much about attitudes versus knowledge, as the wiki studies with their focus on biological and social causes of schizophrenia were exploring the "grey area" where attitudinal differences and knowledge difference converge. Rather, it appears as if the difference between a selective exposure study and a wiki study is that the latter provide an opportunity for the person to act on the environment. Once an opportunity for information generation arises, neither quantitative nor qualitative cognitive conflicts were shied away from.

Similar results occurred for a series of experimental studies that investigated the production of contributions in online discussion forums depending on the amount of qualitative cognitive conflict (Buder, Buttliere, & Ballmann, 2015; Buder & Rudat, 2014; Buttliere & Buder, 2017). All these experiments began with measuring participants' individual attitudes on a conflicting issue (for vs. against alternative medicine). Subsequently, participants were confronted with a carefully controlled set of fictitious online forum posts, half of which were congenial to the participants, and half of which were uncongenial. Participants' reactions to these posts were subsequently captured.

In a first study, the think-aloud method was employed (Buder & Rudat, 2014). While participants were reading through 24 discussion posts on alternative medicine, they were requested to verbalize their thoughts, in particular whether and why they intended to reply or not to reply to a message. After this was accomplished, the 24 posts were again displayed, and participants were requested to write actual replies to those posts they wanted to respond to. Results indicated that participants had a higher intention to reply to uncongenial than to congenial posts, suggesting that larger qualitative cognitive conflicts spurred more productivity than smaller cognitive conflicts. There was no significant difference among congenial and uncongenial posts in explicit verbalizations *not* to reply to a particular post. Further analyses of the verbalizations indicated that the main reasons associated with an intention to reply were experiences of conflict (particularly with uncongenial posts), perceived weakness of arguments (slightly more with congenial posts), and experiences of consensus (particularly with congenial posts). Main reasons for not replying were experience of consensus (mostly with congenial posts), perceived lack of own knowledge, and perceived weakness of arguments (mostly with congenial posts). Actual replies that participants later wrote mimicked the conflict pattern: participants replied more to uncongenial than to congenial posts. Further analyses revealed that replies to congenial posts exhibited agreement and disagreement in equal parts; conversely, replies to uncongenial posts exhibited a large amount of disagreement, and showed no trace of agreement. In sum, the think-aloud study suggests that participants are more likely to respond to uncongenial content than to congenial content, and that these responses appear to predominantly express disagreement. In other words, when given an opportunity to respond to uncongenial information, participants are far from being conflict-averse. Rather, it seems as if they explicitly seek out conflicting information to make their case. This, of course, fits perfectly with the everyday observation of controversies in social media where discussions on controversial issues can garner lots of responses and heated debate.

The main results of the think-aloud study were replicated in a controlled online experiment with 96 participants (Buder et al., 2015). In this study, the size of qualitative cognitive conflicts was captured with much higher precision. In a pre-study, independent raters assessed all 24 discussion posts from the think-aloud study on their relative position on a scale ranging from “totally for alternative medicine” to “totally against alternative medicine.” This allowed measuring a moment-to-moment conflict between a person and each of the 24 posts on a fine-grained level. Results again showed that the likelihood of responding to a post was higher the larger the conflict between person and post was. Moreover, the amount of cognitive conflict was positively associated to the length of replies: not only did participants seek out highly conflicting posts to respond to, they also used more words in their replies.

A set of further studies, only one of which has been published so far (Buttliere & Buder, 2017), then directly pitted these prior finding against a classical selective exposure condition. That is, participants were either given forum posts and asked to indicate how much they would like to read more about the topic (read condition; mimicking selective exposure procedures), or they were asked to indicate how much they would like to respond to the posts (write condition; mimicking our previous studies). In the study by Buttliere and Buder (2017), a repeated-measures design was employed, with half of the participants going through the read condition first, and the other half going through the write condition first. Collapsed over both measurement times, the expected three-way interaction between person attitude, post attitude, and condition emerged. This means that in the read condition participants clearly exhibited a congeniality bias that hundreds of studies in the selective exposure literature have found. However, once participants were given an opportunity to talk back, the congeniality bias was reversed into an uncongeniality bias: the more qualitative conflict there was between person and post, the more likely the person was to respond to a post.

Modern informational environments have introduced social media not only as a means for information consumption, but also for information production. However, both the learning sciences and social psychology strongly focus on the information consumption part (knowledge acquisition and learning; information exposure and attitude change). In order to take a closer look at the information production part, the conceptual and empirical work presented in this section focused on the way how people actively, visibly, and interpersonally respond to cognitive conflict. Both the studies in the wiki context and those in the online forum context suggest that cognitive conflicts lead to productivity.

Conclusions

This chapter began with a working definition of cognitive conflict that emphasized that conflicts appear between the central "stakeholders" in an informational environment: the individual person and the environment. Moreover, conflicts can be defined with regard to different objects (knowledge, attitude), and they can be quantitative or qualitative in nature. Based on this definition, major conceptual and empirical findings in two research fields were discussed that address cognitive conflicts (learning sciences and social psychology). This discussion resulted in the seemingly paradox observation that cognitive conflicts are overwhelmingly positively connoted in the learning sciences, and overwhelmingly negatively connoted in social psychology. However, we argued that the comparison between learning sciences and social psychology might be like comparing apples and oranges, as the learning sciences typically investigate objective outcomes of cognitive conflicts, whereas social psychology typically investigates subjective construals of cognitive conflicts. In order to address this issue, we presented conceptual and empirical findings from our own research laboratories where objective outcomes of cognitive conflicts are investigated with regard to content production in social media. These analyses provided very little support for the common social psychological tenet that people tend to be conflict-averse. Once given an opportunity to talk back, people might actively seek and engage in cognitive conflicts.

In fact, tying together the different research strands described so far, we would argue that it might be possible to arrive at a conceptualization of cognitive conflicts that reconciles the findings from both research areas. Below, we briefly sketch the constituents of such an integrative account.

First, an integrative view on cognitive conflicts suggests that it might be highly useful to distinguish between the subjective construal of cognitive conflicts and the outcomes of cognitive conflicts that actually play out through a communicative exchange of information. Disregarding this distinction might create the impression that findings from the learning sciences (mostly on communicative scenarios) and findings from social psychology (mostly on subjective construals) were incommensurable.

Second, with regard to subjective construals, a case could be made that cognitive conflicts are experienced as slightly aversive. The social psychological literature on dissonance or on selective exposure strongly suggests such a notion, and it would be interesting to have systematic research into potentially similar effects in learning contexts. One could argue that the resolution of cognitive conflicts requires mental effort and cognitive elaboration, and hence can be perceived as strenuous. As humans are often regarded as "cognitive misers" (Fiske & Taylor, 1984) who will only expend as much effort as is needed in a situation (bounded rationality; Simon, 1959), it is not unlikely that this will result in a slightly negative, personal evaluation of cognitive conflicts.

Third, with regard to the possibility to engage in a mutual dialogue with a dissenting source, quite different outcomes might materialize. If an individual is not afforded an opportunity to act on the environment, social psychology suggests that people will not only perceive uncongenial information as aversive, but might engage in active behavior to prevent encountering uncongenial information. However, when an affordance to talk back is present, uncongenial information is no longer avoided, but might be actively sought out. One might even speculate that the notion of "defense motivation" that was used to explain the congeniality bias in selective exposure is a slight misnomer: our data from social media contexts suggests that "defense motivation" has an "offense" component inasmuch as people try to exert influence on others.

Fourth, when the interaction between person and environment affords mutual influence (the environment may impact the person and vice versa), cognitive conflict will be associated with change. As implied by the co-evolution model (Kimmerle et al., 2015), this change can affect both stakeholders. On the one hand, the change may affect the person, as is the case for learning and knowledge acquisition, and as is the case for attitude change. However, change may also affect the environment, as is the case of a learner who shares knowledge, or the person who tries to engage with others in a debate.

In essence, cognitive conflicts (both in the learning sciences and in social psychology) are conducive to change. The learning sciences have traditionally focused on changing the self through assimilation and accommodation. However, with the advent of collaborative learning (O'Donnell & O'Kelly, 1994) the notion of changing the environment (particularly, changing other minds) has begun to be explored. The co-evolution model of knowledge construction (Cress & Kimmerle, 2008; Kimmerle et al., 2015) nicely captures this mutuality of change. Social psychology also has a traditional focus on changing the self. Beginning with Festinger's (1957) seminal work on cognitive dissonance, hundreds of studies have investigated how people deal with uncongenial information. In fact, Festinger already suggested several strategies of how people can resolve conflict and dissonance. One of these strategies, involving an avoidance of uncongenial information, paved the way for research on selective exposure. A second strategy, changing one's attitude in the direction of the uncongenial message, gave rise to research on persuasion. However, Festinger also mentioned a third strategy, and this involves trying to change dissenters via persuasion, which is just a form of changing the environment rather than the self. Ironically, this almost forgotten mode of dissonance reduction was not taken up much by other researchers since the 1950s. However, our own recent work presented in this chapter suggests that changing the environment is a viable strategy to deal with cognitive conflicts. Indeed, we believe that in the informational environments created by today's social media, this third mode of how people deal with cognitive conflict is much more prevalent than experimental psychology suggests.

References

Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Perseus Books.

Aronson, E., Blaney, N., Stephin, C., Sikes, J., & Snapp, M. (1978). *The jigsaw classroom*. Beverly Hills, CA: Sage.

Block, J. (1982). Assimilation, accommodation, and the dynamics of personality development. *Child Development*, *53*, 281–295.

Brauer, M., Judd, C. M., & Gliner, M. D. (1995). The effects of repeated expressions on attitude polarization during group discussions. *Journal of Personality and Social Psychology*, *68*, 1014–1029.

Buder, J., Buttliere, B., & Ballmann, A. (2015). Cognitive conflict in forum discussions on scientific topics. In H. Ogata et al. (Eds.), *Proceedings of the 23rd international conference on computers in education* (pp. 4–6). Jhongli City, Taiwan: Asia-Pacific Society for Computers in Education.

Buder, J., & Hesse, F. W. (2016). Designing digital technologies for deeper learning. In M. Spector, B. B. Lockee, & M. Childress (Eds.), *Learning, design, and technology: An international compendium of theory, research, practice, and policy*. New York, NY: Springer.

Buder, J., & Hesse, F. W. (2017). Informational environments: Cognitive, motivational-affective, and social-affective forays into the digital transformation. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 1–25). New York: Springer.

Buder, J., & Rudat, A. (2014). Antecedents of replies and non-replies in online discussion forums: Evidence from a think-aloud study. In C.-C. Liu et al. (Eds.), *Proceedings of the 22nd international conference on computers in education* (pp. 19–21). Jhongli City, Taiwan: Asia-Pacific Society for Computers in Education.

Buttliere, B., & Buder, J. (2017). Reading more vs. writing back: Situation affordances drive reactions to conflicting information on the Internet. *Computers in Human Behavior*, *74*, 330–336.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*, 752–766.

Cress, U., & Kimmerle, J. (2008). A systemic and cognitive view on collaborative knowledge building with wikis. *International Journal of Computer-Supported Collaborative Learning*, *3*, 105–122.

Desforges, D. M., Lord, C. G., Ramsey, S. L., Mason, J. A., Van Leeuwen, M. D., West, S. C., & Lepper, M. R. (1991). Effects of structured cooperative contact on changing negative attitudes toward stigmatized social groups. *Journal of Personality and Social Psychology*, *60*, 531–544.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, *63*, 568–584.

Doise, W., & Mugny, G. (1984). *The social development of the intellect*. Oxford: Pergamon Press.

Eagly, A. H., Chen, S., Chaiken, S., & Shaw-Barnes, K. (1999). The impact of attitudes on memory: An affair to remember. *Psychological Bulletin*, *125*, 64–89.

Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology*, *71*, 5–24.

Fazio, R. H., Chen, J. M., McDonel, E. C., & Sherman, S. J. (1982). Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association. *Journal of Experimental Social Psychology*, *18*, 339–357.

Fazio, R. H., Powell, M. C., & Herr, P. M. (1983). Toward a process model of the attitude–behavior relation: Accessing one's attitude upon mere observation of the attitude object. *Journal of Personality and Social Psychology*, *44*, 723–735.

Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency. *Advances in Experimental Social Psychology*, *14*, 161–202.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition*. New York, NY: McGraw-Hill.

Frey, D. (1986). Recent research on selective exposure to information. *Advances in Experimental Social Psychology*, *19*, 41–80.

Graesser, A. C., Lippert, A. M., & Hampton, A. J. (2017). Successes and failures in building learning environments to promote deep learning: The value of conversational agents. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 273–298). New York, NY: Springer.

Greitemeyer, T., & Schulz-Hardt, S. (2003). Preference-consistent evaluation of information in the hidden profile paradigm: Beyond group-level explanations for the dominance of shared information in group decisions. *Journal of Personality and Social Psychology*, *84*, 322–339.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*, 555–588.

Hillmert, S., Groß, M., Schmidt-Hertha, B., & Weber, H. (2017). Informational environments and college student dropout. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 27–52). New York, NY: Springer.

Johnson, D. W., Johnson, R. T., & Tjosvold, D. (2000). Constructive controversy: The value of intellectual opposition. In M. Deutsch & P. T. Coleman (Eds.), *The handbook of conflict resolution: Theory and practice* (pp. 65–85). San Francisco, CA: Jossey-Bas.

Judd, C. M., & Kulik, J. A. (1980). Schematic effects of social attitudes on information processing and recall. *Journal of Personality and Social Psychology*, *38*, 569–578.

Kimmerle, J., Bientzle, M., Cress, U., Flemming, D., Greving, H., Grapendorf, J., … Sassenberg, K. (2017). Motivated processing of health-related information in online environments. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 75–96). New York, NY, Springer.

Kimmerle, J., Moskaliuk, J., & Cress, U. (2011). Using wikis for learning and knowledge building: Results of an experimental study. *Educational Technology & Society*, *14*, 138–148.

Kimmerle, J., Moskaliuk, J., Oeberst, A., & Cress, U. (2015). Learning and collective knowledge construction with social media: A process-oriented perspective. *Educational Psychologist*, *50*, 120–137.

Klayman, J., & Ha, Y.-w. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228.

Kleiman, T., & Hassin, R. R. (2013). When conflicts are good: Nonconscious goal conflicts reduce confirmatory thinking. *Journal of Personality and Social Psychology*, *105*, 374–387.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, MA: Cambridge University Press.

Legault, L., Green-Demers, I., & Pelletier, L. (2006). Why do high school students lack motivation in the classroom? Toward an understanding of academic amotivation and the role of social support. *Journal of Educational Psychology*, *98*, 567–582.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098–2109.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

Mojzisch, A., Grouneva, L., & Schulz-Hardt, S. (2010). Biased evaluation of information during discussion: Disentangling the effects of preference consistency, social validation, and ownership of information. *European Journal of Social Psychology*, *40*, 946–956.

Moskaliuk, J., Kimmerle, J., & Cress, U. (2009). Wiki-Supported learning and knowledge building: Effects of incongruity between knowledge and information. *Journal of Computer Assisted Learning*, *25*, 549–561.

Moskaliuk, J., Kimmerle, J., & Cress, U. (2012). Collaborative knowledge building with wikis: The impact of redundancy and polarity. *Computers and Education*, *58*, 1049–1057.

Mullen, B., & Hu, L.-T. (1989). Perceptions of ingroup and outgroup variability: A meta-analytic integration. *Basic and Applied Social Psychology*, *10*, 233–252.

O'Donnell, A. M. (1999). Structuring dyadic interaction through scripted cooperation. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 179–196). Mahwah, NJ: Erlbaum.

O'Donnell, A. M., & O'Kelly, J. (1994). Learning from peers: Beyond the rhetoric of positive results. *Educational Psychology Review*, *6*, 321–349.

Olson, J. M., & Zanna, M. P. (1993). Attitudes and attitude change. *Annual Review of Psychology*, *44*, 117–154.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*, 117–175.

Petty, R. E., & Cacioppo, J. T. (1979). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology*, *37*, 1915–1926.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). New York, NY: Academic Press.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. 1, pp. 703–732). New York, NY: Wiley.

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*, 211–227.

Ritzer, G., Dean, P., & Jurgenson, N. (2012). The coming age of the prosumer. *American Behavioral Scientist*, *56*, 379–398.

Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naïve realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology*, *68*, 404–417.

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1–14). New Haven, CT: Yale University Press.

Schulz-Hardt, S., Jochims, M., & Frey, D. (2002). Productive conflict in group decision making: Genuine and contrived dissent as strategies to counteract biased information seeking. *Organizational Behavior and Human Decision Processes*, *88*, 563–586.

Schwenk, C. R. (1990). Effects of devil's advocacy and dialectical inquiry on decision making: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *47*, 161–176.

Simon, H. A. (1959). Theories of decision making in economics and behavioral science. *American Economic Review*, *49*, 253–283.

Smith, K., Johnson, D. W., & Johnson, R. T. (1981). Can conflict be constructive? Controversy versus concurrence seeking in learning groups. *Journal of Educational Psychology*, *73*, 651–663.

Snyder, M., & Swann, W. B. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, *36*, 1202–1212.

Tjosvold, D., & Johnson, D. W. (1977). Effects of controversy on cognitive perspective taking. *Journal of Educational Psychology*, *69*, 679–685.

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, *22*, 366–389.

Zajonc, R. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*, 1–27.

Zimmerman, B. J., & Blom, D. E. (1983). Toward an empirical test of the role of cognitive conflict in learning. *Developmental Review*, *3*, 18–38.

Zurstiege, G., Zipfel, S., Ort, A., Mack, I., Meitz, T. G. K., & Schäffeler, N. (2017). Managing obesity prevention using digital media—A double-sided approach. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 97–123). New York, NY: Springer.

Chapter 4
Motivated Processing of Health-Related Information in Online Environments

Joachim Kimmerle, Martina Bientzle, Ulrike Cress, Danny Flemming, Hannah Greving, Johannes Grapendorf, Claudia Sassenrath, and Kai Sassenberg

Introduction

The Internet provides users with a virtually unlimited selection of information from which they can acquire health information in a self-regulated manner. Users can decide without any professional guidance which link from a search engine result list they would like to follow, which comment in an online social support group they prefer to read, and at which point they want to leave one online platform and switch to another. The self-regulated nature of this information processing and acquisition in online informational environments suggests that motivation asserts a particularly strong impact on the informational outcomes of individuals' online behavior. This is the case, since self-regulation in essence means that the guidelines for behavior and information selection come from the users themselves.

Researchers at the intersection of motivation and cognition have pointed out that motivational processes often have a major impact on people's search for information, their evaluations of information, their memory and recall, as well as on their judgments and decision-making (Hart et al., 2009; Kruglanski, 1999; Kunda, 1999; Molden & Higgins, 2005; see also Kruglanski et al., 2012). Even though people's information processing is sometimes driven by a requirement for accuracy, that is, by a need to come to a preferably correct assessment, accurate assessment is not always the end result. Striving for accuracy implies acquiring as much knowledge and understanding as possible about a particular topic. This approach

J. Kimmerle (✉) • U. Cress • K. Sassenberg
Leibniz-Institut für Wissensmedien (IWM), Schleichstr. 6, 72076, Tübingen, Germany

Eberhard Karls University Tübingen, Tübingen, Germany
e-mail: j.kimmerle@iwm-tuebingen.de

M. Bientzle • D. Flemming • H. Greving • J. Grapendorf • C. Sassenrath
Leibniz-Institut für Wissensmedien (IWM), Schleichstr. 6, 72076, Tübingen, Germany

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_4

may lead to largely unbiased information processing, but it requires assessable information along with people's willingness and ability to engage in elaborate and unprejudiced handling of that information. Very often, however, information processing is influenced by people's *directed motivation*, which is the endeavor to validate a certain idea or an existing judgment during information processing. This type of information processing is usually accompanied by the acquisition of information that confirms an individual's own attitude, opinion, or impression, and that would presumably lead to achieving one's own goals. Directed motivation thus leads to biased information processing, which we will refer to in this chapter as *motivated information processing*.

The self-regulated nature of dealing with information on the Internet suggests that online informational environments strongly promote motivated information processing. This is the case, in particular, when the domain that users deal with is personally relevant to them (Petty, Cacioppo, & Goldman, 1981). Personal relevance and, as a consequence, motivation are typically high when people search for and deal with *health-related information*. The Internet has developed into a huge and broad information source, above all in the health domain, as people increasingly use it for health-related concerns. Therefore, it is highly plausible that a person's motivation has an impact on how he or she conducts health-related Internet searches and how they deal with the information they find online. At the same time, it is particularly important that laypeople are able to seek, understand, and use health information adequately (Jordan et al., 2013). In this chapter, we summarize research on the effects of motivated processing of health-related information. We address these processes in two separate sections. First, we present research on the motivated processing of health-related information that is influenced by the particular *conceptions* that individuals hold about health in general, about the nature and reliability of health-related knowledge, and about specific health-related topics, such as particular medical treatments. We thereby analyze how people's *individual health concepts* influence their information processing, examine the role of their health-related *epistemological beliefs*, and describe how their previous *opinions* on a health-related topic influence the way they handle information. Second, we address the role of *affect* in motivated processing of health-related information. Here we analyze how searching for information on the Internet can cause fear in laypeople and how patients who have received a medical diagnosis handle online information due to the threat that results from their illness.

Throughout the entire presentation of research findings on motivated processing of health-related information, we also take into account and emphasize the particular characteristics of the information that Internet users encounter when searching for information in online environments. Our presentation makes clear that the processing of health-related information on the Internet is affected not only by the individual factors that people bring with them, but also by the characteristics of the presented information that frame the information as positive or negative, expressing a particular perspective toward the topic (see Fig. 4.1). In concluding, we provide an integrative discussion of the issues analyzed and address implications of the research findings presented in this chapter.

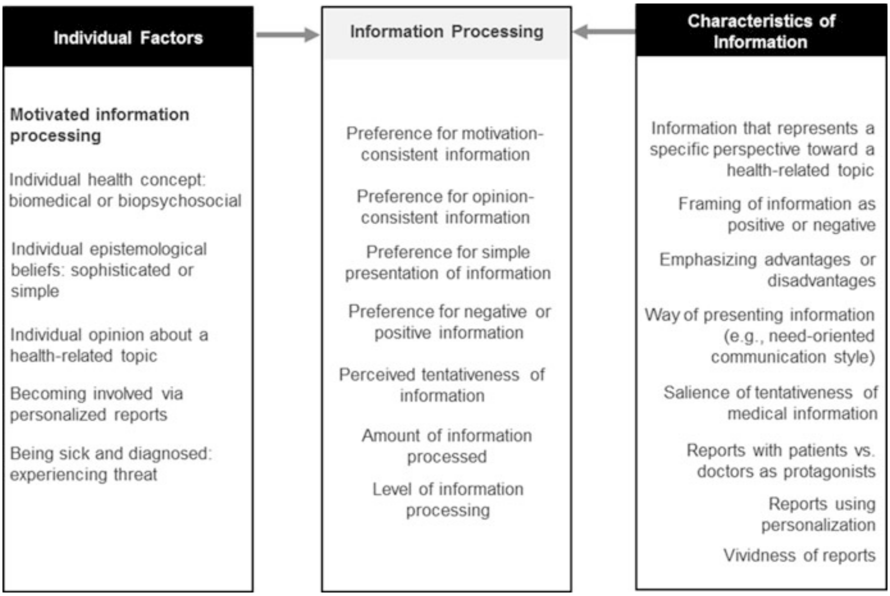

Fig. 4.1 The impact of individual factors and the characteristics of information on the processing of health-related information

The Role of Individual Conceptions in Processing Health-Related Information

When individuals' conceptions and understandings of the world come into play, their processing of information is not a purely rational process. Personal opinions and beliefs have a large influence on how information is dealt with. Information that corresponds to a person's own opinions or beliefs are usually preferred over information that does not (Fischer, Greitemeyer, & Frey, 2008; Greitemeyer, 2014; Sassenberg, Landkammer, & Jacoby, 2014). Such non-conflicting information is more likely to be selected, considered to be more important, and more easily remembered than inconsistent information (Nickerson, 1998; Taber & Lodge, 2006). This phenomenon has been described explicitly for both laypeople and for medical experts (Bornstein & Emler, 2001; Hróbjartsson et al., 2012; Nickerson, 1998). Therefore, it is not surprising that the preference for information that is consistent with one's own concepts also applies to how health-related information is handled on the Internet (for the role of cognitive conflicts in informational environments see also Buder, Buttliere, & Cress, 2017; Graesser, Lippert, & Hampton, 2017).

How people use the Internet to search for and further process health-related information is very often affected by their motivation to validate and maintain their beliefs, their knowledge, and their opinions. Such motivation has an impact on how individuals handle health-related information. In this context, the influence on information processing is mostly driven by a consistency principle (Greitemeyer, 2014; Nickerson, 1998). When individuals' concepts are consistent with (online) information, individuals are more willing to process the information, evaluate the information better, and have a better attitude toward the information or the topic that the information deals with (Schweiger, Oeberst, & Cress, 2014). In this part of the chapter, we examine several factors that we have structured from general to more specific. First, these concepts include people's general understanding of what they regard as "health" and what they think is important about health, that is, their *individual health concepts* (section "Individual Health Concepts in the Reception and Production of Information"). How individuals handle information also depends on their ideas about what health-related "knowledge" is. So information processing is influenced by people's understanding of knowledge in the field of medicine and by their appraisal of the suitability of medical knowledge, that is, by their health-related *epistemological beliefs* (section "Epistemological Beliefs in Processing Health-Related Information"). Third, how people handle health-related information is influenced by their previous *opinions* on a particular health-related topic (section "Opinions in Processing Health-Related Information").

Individual Health Concepts in the Reception and Production of Information

In the health domain, several actors (e.g., patients, physicians, or therapists) are involved who have to share information in different social contexts. A medical doctor, for instance, may talk to a patient to obtain information about the symptoms of a disease or to provide information about a medical intervention such as a particular surgery. To ensure a smooth process of diagnosis and therapeutic intervention, different health professionals have to combine their knowledge. This is very challenging, because different actors have diverse expertise, opinions, and beliefs. Even if they all have the same aim (i.e., maintaining and enhancing the health of a patient; Miles, 2005), this aim itself is open to interpretation. This is the case because different health concepts, that is, general concepts of health and how health is understood, coexist. Individual health concepts vary among different health professionals, between medical experts and patients as well as among different patients (Bensing, 2000; Domenech, Sánchez-Zuriaga, Segura-Ortí, Espejo-Tort, & Lisón, 2011; Larson, 1999; Patel, Arocha, & Kushniruk, 2002). Health concepts are influenced by the individual's socialization and experience of illness and are assumed to be socially constructed (Makoul, 2011; Piko & Bak, 2006).

In the current Western health system, two different health concepts are prominent—the *biomedical* health concept and the *biopsychosocial* health concept. The biomedical health concept is a disease-oriented concept of health that is based on the *International statistical classification of diseases and related health problems* (ICD; World Health Organization, 1992). The biopsychosocial health concept, in contrast, can be defined as an understanding of health in which the patient's personal functioning and participation in daily life are central, a concept based on the *International classification of functioning, disability and health* (ICF; World Health Organization, 2001). The classification systems that underlie these health concepts are both established in the health sector. Even though it is possible to use both classification systems equally (Simeonsson, Scarborough, & Hebbeler, 2006), the ICF is used primarily by health professionals who work in the field of rehabilitative medicine, like physiotherapists. Research has shown that the health concept can influence how patients and medical experts behave in health-related situations (Domenech et al., 2011).

In our own research, we investigated which of the two therapeutic health concepts physiotherapists have and how it develops over time (Bientzle, Cress, & Kimmerle, 2014a). We found that physiotherapists already had a strong biopsychosocial health concept in the early phases of vocational training, while a more pronounced biomedical concept only occurred in the course of their professional life. Moreover, we examined in several studies how the individual health concept influences the reception and production of health-related information. Regarding the *reception* of information, we found that information was rated as more relevant for the decision to participate in a medical intervention if the individual health concept and the health concept presented in the information were consistent with each other (Bientzle, Cress, & Kimmerle, 2015). Participants also acquired more knowledge when information corresponded to their individual health concept (Bientzle, Cress, & Kimmerle, 2013). These results are in line with research findings in the area of health communication that show that the reception of health-related information can be effectively supported by adapting information to the characteristics of the target group (Albada, Ausems, Bensing, & van Dulmen, 2009; Hawkins, Kreuter, Resnicow, Fishbein, & Dijkstra, 2008; Kreuter & Wray, 2003). Therefore, our research contributes to such health communication research by suggesting that patients and other laypeople are more willing to accept particular health-related information when their health concepts are taken into account in the presentation of this information.

For the *production* of information in online settings, we also found an important impact of people's individual health concepts. Our research indicates that the individual health concept influences the way people communicate in online forums: Medical experts with a strongly pronounced biomedical health concept used more scientific words and fewer emotional words in their online communication than experts with a weak biomedical concept (Bientzle, Griewatz et al., 2015). When medical experts had the task of revising a text in a collaboratively developed wiki, which was either consistent or inconsistent with their own individual health concept, we found that they were more tolerant toward false conclusions in the text that was compatible with their own concept. In contrast, in the inconsistent wiki-text, medical experts tried to resolve false conclusions by revising scientific content more often (Bientzle, Cress, & Kimmerle, 2014b).

Epistemological Beliefs in Processing Health-Related Information

Assumptions regarding the structure and stability of knowledge also have an impact on the evaluation of information and the acquisition of knowledge (Hofer, 2001; Muis, 2004). These assumptions are called epistemological beliefs (EBs). EBs are defined as "the cognitions (i.e., understandings) individuals have on knowledge and knowing and determine how (new) knowledge is perceived and processed" (Roex & Degryse, 2007, p. 616). In that sense, EBs can be either more sophisticated or more simple. People with sophisticated EBs consider knowledge to be subjective and dynamic, whereas people with simple EBs regard knowledge as objective and static. Despite this seemingly clear dichotomy, EBs reflect a complex and multidimensional construct that is difficult to measure in its entirety.

The ability to handle information in a sophisticated and reflective way in the health domain is quite important, because the half-life of medical knowledge is steadily decreasing (Arbesman, 2012). Medical knowledge that is currently considered to be sound and valid can be tomorrow's history, assessed as obsolete or even wrong. Thus, EBs play a major role in adequately handling information in the health context. Individuals with sophisticated health-related EBs are better able to process health-related information accurately than people with simple EBs. EBs are also dependent upon people's educational level (Muis, Bendixen, & Haerle, 2006). In our own research we found that physiotherapy students and physiotherapy professionals have EBs of differing sophistication regarding knowledge in physiotherapy and regarding knowledge in medicine: Professionals assessed health-related knowledge as more subjective, constructed, and relative than did first-year students (Bientzle et al., 2014a).

EBs also play an important role in Internet search and the assessment of scientific and health-related information online (for an overview see Bråten, Britt, Strømsø, & Rouet, 2011). During an Internet search, various pieces of information from multiple sources with inconsistent or contradicting perspectives must be integrated and weighed against each other (Rouet, 2006). This may be particularly difficult for people with simple EBs. The necessity of grappling with contradictions might appear to be unreasonable for people who tend to underestimate the complexity of scientific knowledge. Research has shown that students with simple EBs were not able to process the information in a hypertext system as deeply as students with sophisticated EBs (Jacobson & Spiro, 1995). They also learned less and had poorer self-regulation abilities when they read hypertexts (Pieschl, Stahl, & Bromme, 2008). In our own research, we investigated how individuals deal with scientific medical information and found overall that individuals with sophisticated EBs recognized the tentativeness of medical information to a greater extent than people with simple EBs (Feinkohl, Flemming, Cress, & Kimmerle, 2016a). In particular, we illustrated that a lack of reliability in health-related information was more likely identified by participants with more sophisticated EBs than by participants with rather simple EBs (Kimmerle, Flemming, Feinkohl, & Cress, 2015). When Internet users deal with health-related information online, that is, when they read health news articles on the Internet and write user comments in online forums, people with sophisticated EBs address the tentativeness of medical information more often than Internet users with simple EBs (Feinkohl, Flemming, Cress, & Kimmerle, 2016b).

Furthermore, we could show that people with more sophisticated EBs were better in recalling health-related information (Feinkohl et al., 2016a). Overall, individuals with simple EBs were less critical in evaluating health-related information as presented in online media. These findings are in line with study results in other knowledge domains, which found that sophisticated EBs are also positively correlated with learning and academic performance (Hofer, 2001; Muis, 2004).

Health-related and medical knowledge and information are already in themselves unreliable constructs. When it comes to health-related information on the Internet, the reliability may even further decrease. This is the case when individuals (e.g., medical experts and laypersons) have to deal with knowledge from different sources and of different quality on the Internet. Moreover, alternative health communities on the Internet even have their own rules and quality standards to assess information (Kimmerle et al., 2013). Against this background, we might assume that sophisticated EBs may to some degree counteract people's tendency to engage in motivated information processing. This becomes all the more important when people have to deal with health-related information on the Internet and need to evaluate the available information adequately.

Opinions in Processing Health-Related Information

When processing health-related information, individuals easily form an opinion about the topic they are dealing with (Burakgazi & Yildirim, 2014; Kimmerle & Cress, 2013). Newspaper articles or other media contributions dealing with health-related topics focus especially on the formation of opinion. Examples of such reports are media reports on new medical treatments such as deep brain stimulation (DBS). DBS involves a surgical procedure in which electrodes are implanted in the patient's brain. It reduces tremors and other negative symptoms of patients with Parkinson's disease. Researchers have also started to apply DBS as a treatment for drug addiction and other diseases, but its effectiveness is highly questionable and currently far from being entirely clear (Gilbert & Ovadia, 2011). Authors of media stories on DBS tend to over-simplify their reports in order to make them more easily understandable and free of conflicting information for readers without medical background knowledge (Kimmerle et al., 2015). As a consequence, these reports are often either framed overly positively, for example, by over-emphasizing the beneficial effects of DBS, or overly negatively through emphasis on risks or side effects (Gilbert & Ovadia, 2011; Racine, Waldman, Palmour, Risse, & Illes, 2007; see also Moynihan et al., 2000). This one-sided reporting is usually inadequate because health-related information, especially about novel medical findings, is often provisional, unreliable, inconsistent, contradictory, or controversial: That is to say, health-related research results are often “tentative” (Bromme & Goldman, 2014; Flemming, Feinkohl, Cress, & Kimmerle, 2015). This is a problem especially for laypeople, as they are often not aware of this tentativeness. Laypeople rather tend to perceive medical findings as clear, stable, reliable, or ever-lasting when they form their opinions about medical treatments from media content, which can lead to misconceptions and a false sense of clarity (Scharrer, Bromme, Britt, & Stadtler, 2012; Sinatra, Kienhues, & Hofer, 2014; van Deemter, 2010).

In our research, we investigated how prior opinions regarding DBS affected how individuals perceived the tentativeness of information on DBS. We found that, when participants had a negative opinion about DBS, they detected the tentativeness in a media report on DBS more often compared to those media consumers who read the report with a positive opinion about DBS (Kimmerle et al., 2015). Moreover, the amount of detected tentativeness declined if the information in the article was framed in a positive way. The two effects were independent of each other. In other words, participants with a negative opinion were more skeptical. They were less vulnerable to over-simplification and one-sided positive framing of health-related information than participants with a positive opinion. So, personal opinions have an influence on how information is processed when it comes to being critical toward health-related information as reported in the media. In particular, positive opinions seem to emphasize people’s bias of not recognizing the tentativeness of medical research findings.

We also found that prior opinions have an impact on how laypeople evaluate the advantages and disadvantages of medical interventions (Bientzle, Griewatz et al., 2015; Kimmerle, Bientzle, & Cress, 2017): Studies on the assessment of information about a mammography screening program showed that women who had a positive opinion of mammography screening in advance (and were thus willing to participate in the program anyway), also rated particular pieces of information representing the advantages of screening as more important than information that addressed disadvantages. Women who did not hold such a positive prior opinion on mammography screening did not differ in their assessment of advantages and disadvantages of mammography screening. So, the mere presentation of balanced information (both advantages and disadvantages) about a medical intervention does not guarantee that the information will be processed in a balanced way. Thus, health communicators, including medical professionals such as physicians as well as science journalists responsible for presenting health-related information, have to be highly aware of the role of people’s prior opinions in the processing of that information.

Our own findings address information processing of laypeople and (potential) patients, but experts are not immune to the influence of their own opinions either. For instance, Moreno and Johnston (2013) showed the occurrence of confirmation bias in the decision-making of healthcare providers. This is important, because opinion leaders within the healthcare system have an exceptional influence on the promotion of particular medical practices (Doumit, Gattellari, Grimshaw, & O’Brien, 2007). This is relevant, since the opinions of health professionals and those of patients may also interact in a complex way. In an online-consultation study, for instance, we found that after a consultation with a physician who provided information about risks and benefits of mammography screening, people had a more positive opinion about mammography if their needs had been made salient beforehand and when the physician used a need-oriented communication style (Fissler, Bientzle, Cress, & Kimmerle, 2015). If the physician used a purely fact-orientated communication style, in contrast, this led to an attitude change to a negative direction. It seems that not only the actual information itself but also the fit to a person's opinions, needs, and expectations is of importance when laypeople process health-related information online (Bientzle, Griewatz et al., 2015). We have also shown that people adapt their opinion about a medical treatment more strongly to the opinion of the physician if the physician communicates in a patient-centered compared to a physician-centered way (Bientzle, Fissler, Cress, & Kimmerle, in press).

Summary

The research presented so far clearly indicates that people's concepts and beliefs have a strong impact on how they process health-related information that they search for and find online. They prefer to process information that is consistent with their own concepts. This applies to their very general understanding of health (i.e., their individual health concepts), to their assumptions about the nature of health-related knowledge (i.e., their individual health-related EBs), and to their specific opinions toward particular health-related topics.

We have shown that the health concept is a factor that varies among patients and among different health professionals. In a series of studies, the fit between the individual health concept and the health concept in which information was presented was identified as a factor that influenced the reception and production of health-related information on the Internet. The results seem to be crucially relevant for inter-professional information exchange and health communication issues in the health system. Our studies indicate that it is very important to consider people's health concepts very thoroughly, regardless of whether we are examining situations of mere information processing or, further, the production of information. Both processes are strongly influenced by what people regard as relevant when they talk about health. When patients and other medical laypeople search the Internet for health-related information, they should be aware of their own understanding of health and of their ultimate goal in this search process.

Individual EBs vary between students and health professionals and among different health domains. This influences the processing of health-related information, which is often tentative and ill-structured. A certain level of sophistication of EBs seems to counteract the motivated processing of health-related information. Moreover, dealing with information on the Internet entails further challenges, like the integration of inconsistent information sources and perspectives. Opinions also have a strong impact on how people handle health-related information. They tend to process information in a way that is consistent with their prior opinions. If individuals have a positive opinion about a specific health-related topic, they tend to underestimate the tentativeness of the pertinent medical knowledge on this topic. Our research shows that this distortion of the information processing may in part be compensated for in those Internet users who possess more sophisticated EBs concerning medical knowledge.

How individuals process health-related information is not only influenced by people's concepts of the world; how people search for and perceive health-related information can also be influenced by their emotions and affective reactions. The following section addresses the role of affect in the motivated processing of health-related information on the Internet.

The Role of Affect in Processing Health-Related Information

Health-related Internet searches may have different implications for individuals with respect to their feelings. In this section, we address two different roles that emotions may play in processing health information in online environments. Due to a motivated processing of information, some individuals may feel worse than before finding the information, while other individuals may find the online information helpful and feel better. In two sub-sections, we present empirical evidence for both possibilities and describe under which conditions individuals feel worse or better when searching the Internet for health-related information. Feeling worse can result from a person's focus on negative information during a health-related Internet search. This focus on negative information may arise from a sort of *cyberchondria* (section "Health-Related Internet Search and Cyberchondria"). Individuals who have been *diagnosed* with a particular disease, in contrast, are motivated to process information on this disease in a goal-oriented way and tend to focus on positive information that they deem helpful for their situation (section "Information Processing of Diagnosed Patients on the Internet").

Health-Related Internet Search and Cyberchondria

Having access to health-related information online can sometimes evoke fear in laypeople. It seems to be very common that individuals conduct Internet searches when experiencing unfamiliar states of their bodies or when there is information in the media warning about a potential epidemic (e.g., Ginsberg et al., 2009). Such searches for information often lead to the fear that one is sick, and individuals find evidence on the Internet that the slightest symptoms are an indication of severe illness. Based on this anecdotal evidence, White and Horvitz (2009) coined the term *cyberchondria* to label the phenomenon that individuals frequently encounter negative information during health-related Internet searches—in particular when they are searching for the sources of physical symptoms. Even though the empirical research on the causes and antecedents of this phenomenon is still in its very early stages, (weak forms of) cyberchondria seem to occur quite frequently (Muse, McManus, Leung, Meghreblia, & Williams, 2012).

In our own research, we aimed to study whether the affective content of online information and how it is presented are facilitating factors in this fear of being sick. To be more precise, we wanted to gain insight into the characteristics of online information that contribute to fear and information processing of negative information among users. The results could help to understand and, ultimately, to prevent, biased information processing. We conducted experiments in which laypeople read reports about a specific, health-related treatment, namely DBS (see section "Opinions in Processing Health-Related Information"). DBS is applied to patients with disabling neurological symptoms, such as those of Parkinson's disease. It does not cure them but, if applied successfully, it delivers substantial relief from major symptoms. In reports about DBS, the mere description of the surgery can elicit uneasy feelings in laypeople and supply them with information they experience negatively. In particular, that the brain surgery is partly undertaken when the patient is conscious, and that it may have the side effect of changing a patient's personality, produce emotional responses in laypeople.

Besides the content of information sources on DBS, the way such reports on DBS are written can influence laypeople's information processing as well. In particular, journalists reporting on DBS frequently use *personalization* as a technique to increase the attractiveness of their reports. They introduce a fictitious or a real protagonist at the beginning of the text to make the storyline more vivid and interesting. In the case of medical treatments, patients and doctors are prototypical protagonists. Most likely, reports with patients as protagonists are more personally dramatic to laypeople than reports of doctors applying DBS, as most laypeople have at some time experienced being a patient. As a consequence, reports with patients as protagonists dramatize the patient's perspective for laypeople and cause them to identify with the patient protagonists, taking on their perspective (Batson & Shaw, 1991; Batson et al., 1988). Such patient reports often simultaneously convey negative information, for example, about the irksome procedure of the brain surgery. As a result, such negative information elicits negative emotions in laypeople. Given the fact that people preferentially process affect-congruent information (e.g., Blaney, 1986; Bower, 1981), feeling negative about a medical treatment would foster memory with a preference for negative information about this treatment.

Thus, we hypothesized that reports on DBS with patients as protagonists would elicit a stronger negative affect and, in turn, better memory of negative information, than reports with doctors as protagonists. Following this reasoning, reports with patients as protagonists reporting on more positive experiences with DBS should then certainly also elicit a more positive affect. However, given that this study focused on a treatment that can at best reduce—not cure—symptoms of a severe illness, we mainly expected effects with regard to negative emotions in recipients. This hypothesis was tested in a study (Sassenrath, Greving, & Sassenberg, 2017) in which participants first listened to an audio report about DBS. We applied this procedure because we had found in a previous study that an audio presentation increased the vividness of reports as well as the affective responses (Sassenrath, Greving, & Sassenberg, 2016b). These audio reports about DBS either involved a *patient or a doctor as the protagonist*. Afterwards, we assessed affect. Next, participants read additional information about DBS. After working for 5 min on a different, unrelated task, participants were unexpectedly asked to recall as much information about DBS as possible. In line with the prediction, participants who had heard reports with patients as protagonists demonstrated more negative and less positive affect than those who had heard reports with doctors as protagonists. Moreover, after listening to reports with patients as protagonists, the negative affect further led to a focus on negative information (i.e., more recall of negative information and less recall of positive information). In sum, these findings indicate that reports on medical treatments of severe illnesses using patients as protagonists elicit more negative affect and a preferential processing of negative information compared to reports using doctors as protagonists.

In another study (Sassenrath, Greving, & Sassenberg, 2017), we aimed to rule out the possibility that the effects reported above resulted from the fact that the text focused on one patient as a protagonist rather than on patients in general. To this end, we compared participants' affective responses and memory after listening to a report with one patient as the protagonist (reporting from a first-person perspective), versus a report with the same content but using wording of patients' experiences in general (reporting form a third-person perspective). Otherwise, the same paradigm was applied as in the study described above. Again, patients' reports on their experiences with DBS from a first-person perspective elicited stronger emotions and—indirectly via negative affect—recall of more negative information about DBS than patients' reports from a third-person perspective.

Taken together, these studies suggest that the personalization features of patients' reports contribute to the development of negative affect and memory of negative information in laypeople. For Internet users searching for health-related information, these findings imply that reports using individuals as protagonists who are similar to users cause users to develop fear as a result of their preferential processing of negative information. In the same way, personal online reports (e.g., in online social support groups) are very likely to elicit the same kind of identification and, thus, similar fearful effects to news reports. More generally, the *vividness of reports* is more likely to evoke affective responses. In other words, patient protagonists and other features that increase identification and perspective taking, or vividness in news reports about diseases, are likely to fuel fear and vigilance toward negative information in Internet users, and should, thus, be applied with caution.

Information Processing of Diagnosed Patients on the Internet

Motivated information processing may also come into play for patients who have been diagnosed with a certain illness. An illness is an undesirable state, and individuals are usually highly motivated to overcome this state. Reduced resources as a result of the illness paired with high motivation to overcome the illness usually lead to threat appraisals (Blascovich & Tomaka, 1996). Threat is not necessarily understood in the sense of an existential threat, but in the sense of a threat to achieving goals, or simply a threat to being in a desirable state (i.e., to being healthy). Threatened individuals search for information that may help them to overcome the illness, because they feel they lack the resources to do so on their own. Threat may, therefore, lead to the preferential processing of positive information. According to the counter-regulation principle (Rothermund, 2011), individuals allocate their attention and their processing capacity to stimuli with a valence opposing their current motivational state: In a positive motivational state, negative information is preferably processed, whereas in a negative motivational state, positive information is preferably processed. Threat is a negative motivational state that individuals clearly aim to change. There is plenty of evidence for the counter-regulation principle. Individuals in a negative motivational state, such as threat, preferentially process positive information, and (less relevant in the current context) individuals in a positive motivational state preferentially process negative information. These effects occur for a range of different psychological processes, such as attention and complex judgments and decisions (e.g., Koranyi & Rothermund, 2012; Rothermund, Gast, & Wentura, 2011; Sassenberg, Sassenrath, & Fetterman, 2015; Schwager & Rothermund, 2014; for the role of emotion regulation in informational environments see also Azevedo et al., 2017).

Counter-regulation is, due to its unconscious nature, likely to apply to conditions characterized by allowing self-regulation that makes it possible for individuals to deal with information in a manner that suits their needs. This is especially the case for the unguided and iterative multi-step process of Internet searches (e.g., Brand-Gruwel, Wopereis, & Walraven, 2009). Therefore, it seems likely that an information search on the Internet would also be affected by threat. To be more precise, counter-regulation should affect all of the steps of an Internet search process, that is, (a) the generation of search terms, (b) the selection of links, (c) the time spent on webpages, (d) the encoded information, (e) the judgment about a topic made after the information search, and (f) the recall of information. Overall, we predicted a bias toward positive information in all these steps when individuals felt threatened, compared to when they did not. These predictions were tested in a series of five experiments (Greving & Sassenberg, 2015; Greving, Sassenberg, & Fetterman, 2015) using different manipulations of threat and different materials. For ethical reasons, we primarily manipulated the threat and only conducted one longitudinal field study with real patients (Sassenberg & Greving, 2016).

In a first study, we tested the prediction that threat would lead to the generation of more positive search terms (Greving & Sassenberg, 2015, Experiment 1). In order to induce threat, participants were first asked to think about a current demand in their life they had trouble coping with. In a control condition, they were asked to think about a typical situation in their life. Afterwards, participants were asked to generate search terms for an Internet search on living organ donation. In line with the prediction, more positive search terms were generated by individuals who had thought about a current threat than by participants in the control condition.

In a second study, we induced a health threat to test the impact of threat on the selection of links from a search engine result page (SERP; Greving et al., 2015, Experiment 2). To this end, a well-established manipulation by Ditto and Lopez (1992; Ditto, Scepansky, Munro, Apanovitch, & Lockhart, 1998) was applied. Participants took an ostensible saliva test for a newly discovered enzyme deficit. In the threat condition, they learnt that they had this deficit, whereas in the control condition participants were informed that the test had not worked out. Afterwards, participants received a SERP regarding the enzyme deficit. This SERP was presented in the format of a 4×4 table to avoid the possibility that participants would merely select the links at the top (because they assumed that they were more relevant as is the case with Google result pages). Participants had to choose 8 out of 16 links. As assumed, participants in the threat condition chose a higher proportion of positive links than participants in the control condition.

A third experiment studied the whole Internet search process (Greving et al., 2015, Experiment 1). After recalling a threatening or, in the control condition, an ordinary situation related to an upcoming examination, participants conducted a free Internet search for 10 min. The Internet search process was logged and the pages on the computer screens were recorded. Participants in the threat condition looked longer at positive pages and also recalled more positive information than participants in the control condition. This effect was also replicated in a controlled laboratory experiment in which a SERP was completely pre-programmed (Greving & Sassenberg, 2016).

A fourth follow-up study tested whether the recall of positive information was also more likely when the reading material was held constant across conditions (Greving et al., 2015, Experiment 3), which had not been the case in the previous study. Therefore, a threat (vs. control) was induced with a similar recall procedure as in the first search term generation study. Afterwards, individuals read a fixed set of texts taken from an Internet search on living organ donation that did not differ between conditions. Participants in the threat condition in fact recalled more positive information from these texts than participants in the control condition. This demonstrated that the impact of threat on the recall of positive information did not depend on the amount and valence of the material that the threatened individuals were assessing. In this study, participants also judged the risks and benefits associated with donating living organs. In line with our expectation that threat leads to a positive bias, participants in the threat condition also judged living organ donation to be more beneficial (but not less risky).

Finally, threat not only affects the search process that is mostly characterized by encoding information, but also the retrieval of information at the end of the search process. Therefore, a final experimental study (Greving & Sassenberg, 2015, Experiment 2) used the materials of the study summarized above, but reversed the order of the elements: Participants first read a text on living organ donation, then a threat was manipulated, and afterwards participants were asked to recall information from the text. This procedure allowed us to study the mere retrieval effects of a threat. Again, supporting the idea that threat leads to preferential processing of positive information, more positive information was in fact recalled from memory when a threat was induced after reading the text than when no threat was induced.

In sum, these studies provided evidence that threat creates a tendency toward positive bias in all of the steps of the multistep-process of information search on the Internet. Since it is desirable to complement study sets of internally valid experiments with externally valid field studies, we conducted a longitudinal field study on the outcomes of Internet search behavior resulting from how actual patients perceived their own health state (Sassenberg & Greving, 2016). Patients with chronic inflammatory bowel disease were recruited for a two-wave longitudinal study with a 7-month lag. At both points in time, the frequency of their health-related Internet search, the severity of the illness as an indicator of threat, and health-related optimism as a relevant outcome measure were assessed via self-reports. It was predicted that among patients using the Internet frequently to search for health-related information, the findings from the experimental studies would be replicated—more threat would lead to more health-related optimism. No such effect, in contrast, was expected for patients using the Internet rarely for this purpose. Indeed, the interaction of threat and frequency of health-related Internet use at the first measurement point predicted health-related optimism at the second measurement point after 7 months (controlling for optimism at the first measurement point). The pattern underlying the interaction effect fit the prediction. We were thus also able to show in a field setting that a health threat led to a stronger positivity bias (more health-related optimism) among individuals using the Internet frequently to acquire health information. Stated differently, among patients with a greater health threat, more frequent Internet searches regarding their health led to more optimism. In contrast, among patients with a smaller health threat, the opposite effect occurred, indicating that they were less optimistic about their health. Additional analyses provided no evidence for such an effect for other external sources of information.

We can conclude that patients searching for information regarding their (diagnosed) disease are likely to be subject to a positivity bias—in particular if the disease is threatening to them. On the one hand, the positivity bias is a means of emotionally coping with the negative psychological side effects of disease, because the sense of threat will most likely be held under control by focusing on positive information. It follows that patients experiencing a threat might be particularly prone to information or advertisements for products promising a cure, even if the promises are invalid or inaccurate.

Summary

The research described in this section has shown that a motivated processing of health-related information on the Internet may have negative effects, in particular regarding people's emotions and their preferential processing of negative information. Already the term cyberchondria itself points to the fact that strong forms of health-related anxiety induced by an Internet search are categorized as a type of disorder (a specific type of obsessive-compulsive disorder; Norr, Oglesby et al., 2015). Thus, anxiety as an outcome of a health-related Internet search is a dysfunctional and undesirable response, whereas the positivity bias in counter-regulation demonstrated by the information processing of diagnosed patients may be a functional and healthy response to negative motivational states (Rothermund, 2011). We have shown that this positivity bias, however, only occurs when individuals are currently in a negative motivational state and have a strong desire to change their situation. This conclusion fits the results from the longitudinal study on patients with inflammatory bowel disease: Those with a serious illness and frequent episodes, who are likely to see their illness as a threat, showed a positive bias as an outcome of frequent Internet searches, whereas those with less sense of threat and infrequent episodes showed the opposite effect of Internet searches. To mention one implication of these processes, it could be that patients gleaning information from the Internet might confront doctors with their unrealistically positive views of their own situation. This is a possibility that requires attention and doctors need to be prepared. Otherwise, patients' health might suffer because they might, as an example, stop a treatment too early (for a detailed discussion of the ethical implications of these findings see Sassenberg & Wiesing, 2016).

Conclusion

Online informational environments promote the impact of motivated information processing. This is particularly true when personal relevance is high, as is the case when people search for and deal with health-related information. The empirical studies presented in this chapter illustrate very clearly that how people handle health-related information on the Internet is strongly influenced by their particular motives. When people process health-related information in online situations, they often have the desire to validate their prior knowledge, opinions, and judgments. As a consequence, they deal with online information in a way that allows them to confirm their existing understanding of or attitude toward a health topic. We have presented much empirical evidence for how this may result in biased information processing. In addition, our presentation of research findings regarding cyberchondria and the behavior of diagnosed patients makes it clear that experiencing anxiety (as a consequence of an information search on the Internet) and threat (as a consequence of being sick) also frequently results in biased processing and representation of information.

All things considered, comprehending people's processing of health-related information on the Internet is even more complicated, since it is not only influenced by their motivation but also by particular *characteristics of the information* (see Fig. 4.1). As our research has shown, these characteristics comprise several aspects. How people deal with health-related information also depends on the specific perspective on a health topic that is represented in the information in an online setting. For example, it makes a difference for information processing when people with a biomedical or a biopsychosocial individual health concept encounter information that also expresses a biomedical or a biopsychosocial view on health, which is either compatible with or contrary to their own point of view. People with a particular motivation also process information differently when it is framed in a positive way than when it is framed negatively. The same applies to health information on the Internet that either emphasizes the advantages or the disadvantages of a particular medical treatment. The particular focus of the health information interacts strongly with people's prior opinions on the topic. Moreover, the way in which the information is presented is relevant. For example, it makes a difference whether health professionals in an online consultation take the personal needs of a patient into account or are merely focused on their medical issues.

In medical news reports, the presentation of the information is also essential for the processing of this information. For example, a report that explicitly addresses the tentativeness of medical research findings may support the readers in recognizing and understanding this tentativeness. This is particularly important for those readers who possess rather simple epistemological beliefs. Finally, it makes a difference to what extent a report uses the personalization of health information. A report may use either patients or doctors as protagonists in depicting a health story, or it may apply a first- or a third-person perspective to make the story more or less personal. How Internet users who search for health information will react to medical news depends on whether or not such reports aim to convey a vivid and dramatic picture of a health-related topic to draw more attention to it.

Due to this interplay between individual factors on the one hand and the characteristics of information on the other, future research should make an attempt to take both aspects into consideration at the same time. This chapter is a first step in providing an integrative view of both aspects. This interplay is an issue that should also be taken into account in applied situations. Health communicators, physicians, and health journalists should be aware of these interactions between individual factors and ways of communicating. Only then can laypeople be adequately supported in their self-regulated processing of health-related information in online environments.

References

Albada, A., Ausems, M. G., Bensing, J. M., & van Dulmen, S. (2009). Tailored information about cancer risk and screening: A systematic review. *Patient Education and Counseling*, *77*, 155–171. doi:10.1016/j.pec.2009.03.005

Arbesman, S. (2012). *The half-life of facts: Why everything we know has an expiration date*. New York, NY: Penguin.

Azevedo, R., Millar, G. C., Taub, M., Mudrick, N. V., Bradbury, A. E., & Price, M. J. (2017). Using data visualizations to foster emotion regulation during self-regulated learning with advanced learning technologies. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 225–247). New York: Springer.

Batson, C. D., Dyck, J. L., Brandt, J. R., Batson, J. G., Powell, A. L., McMaster, M. R., & Griffitt, C. (1988). Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis. *Journal of Personality and Social Psychology*, *55*, 52–77. doi:10.1037/0022-3514.55.1.52

Batson, C. D., & Shaw, L. L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry*, *2*, 107–122. doi:10.1207/s15327965pli0202_1

Bensing, J. (2000). Bridging the gap: The separate worlds of evidence-based medicine and patient-centered medicine. *Patient Education and Counseling*, *39*, 17–25. doi:10.1016/S0738-3991(99)00087-7

Bientzle, M., Cress, U., & Kimmerle, J. (2013). How students deal with inconsistencies in health knowledge. *Medical Education*, *47*, 683–690. doi:10.1111/medu.12198

Bientzle, M., Cress, U., & Kimmerle, J. (2014a). Epistemological beliefs and therapeutic health concepts of physiotherapy students and professionals. *BMC Medical Education*, *14*, 208. doi:10.1186/1472-6920-14-208

Bientzle, M., Cress, U., & Kimmerle, J. (2014b). The role of inconsistencies in collaborative knowledge construction. In *Proceedings of the 11th international conference of the learning sciences* (Vol. I, pp. 102–109). Boulder, CO: International Society of the Learning Sciences.

Bientzle, M., Cress, U., & Kimmerle, J. (2015). The role of tentative decisions and health concepts in assessing information about mammography screening. *Psychology, Health & Medicine*, *20*, 670–679. doi:10.1080/13548506.2015.1005017

Bientzle, M., Fissler, T., Cress, U., & Kimmerle, J. (in press). The impact of physicians' communication styles on evaluation of physicians and information processing: A randomized study with simulated video consultations on contraception with an intrauterine device. *Health Expectations*. It's still in press. doi:10.1111/hex.12521

Bientzle, M., Griewatz, J., Kimmerle, J., Küppers, J., Cress, U., & Lammerding-Koeppel, M. (2015). Impact of scientific versus emotional wording of patient questions on doctor-patient communication in an Internet forum: A randomized controlled experiment with medical students. *Journal of Medical Internet Research*, *17*, e268. doi:10.2196/jmir.4597

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin*, *99*, 229–246. doi:10.1037/0033-2909.99.2.229

Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. *Advances in Experimental Social Psychology*, *28*, 1–51.

Bornstein, B. H., & Emler, A. C. (2001). Rationality in medical decision making: A review of the literature on doctors' decision-making biases. *Journal of Evaluation in Clinical Practice*, *7*, 97–107. doi:10.1046/j.1365-2753.2001.00284.x

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148. doi:10.1037/0003-066X.36.2.129

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J. F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist*, *46*, 48–70. doi:10.1080/00461520.2011.538647

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education*, *53*, 1207–1217. doi:10.1016/j.compedu.2009.06.004

Bromme, R., & Goldman, S. R. (2014). The public's bounded understanding of science. *Educational Psychologist*, *49*, 59–69.

Buder, J., Buttliere, B., & Cress, U. (2017). The role of cognitive conflicts in informational environments: Conflicting evidence from the learning sciences and social psychology? In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 53–74). New York: Springer.

Burakgazi, S. G., & Yildirim, A. (2014). Accessing science through media: Uses and gratifications among fourth and fifth graders for science learning. *Science Communication*, *36*, 168–193.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, *63*, 568–584. doi:10.1037/0022-3514.63.4.568

Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology*, *75*, 53–69. doi:10.1037/0022-3514.75.1.53

Domenech, J., Sánchez-Zuriaga, D., Segura-Ortí, E., Espejo-Tort, B., & Lisón, J. F. (2011). Impact of biomedical and biopsychosocial training sessions on the attitudes, beliefs, and recommendations of health care providers about low back pain: A randomised clinical trial. *Pain*, *152*, 2557–2563. doi:10.1016/j.pain.2011.07.023

Doumit, G., Gattellari, M., Grimshaw, J., & O'Brien, M. (2007). Local opinion leaders: Effects on professional practice and health care outcomes. *Cochrane Database System Review*, *1*.

Feinkohl, I., Flemming, D., Cress, U., & Kimmerle, J. (2016a). The impact of epistemological beliefs and cognitive ability on recall and critical evaluation of scientific information. *Cognitive Processing*, *17*, 213–223. doi:10.1007/s10339-015-0748-z

Feinkohl, I., Flemming, D., Cress, U., & Kimmerle, J. (2016b). The impact of personality factors and preceding user comments on the processing of research findings on deep brain stimulation: A randomized controlled experiment in a simulated online forum. *Journal of Medical Internet Research*, *18*, e59.

Fischer, P., Greitemeyer, T., & Frey, D. (2008). Self-regulation and selective exposure: The impact of depleted self-regulation resources on confirmatory information processing. *Journal of Personality and Social Psychology*, *94*, 382–395. doi:10.1037/0022-3514.94.3.382

Fissler, T., Bientzle, M., Cress, U., & Kimmerle, J. (2015). The impact of advice seekers' need salience and doctors' communication style on attitude and decision making: A web-based mammography consultation role play. *JMIR Cancer*, *1*, e10. doi:10.2196/cancer.4279

Flemming, D., Feinkohl, I., Cress, U., & Kimmerle, J. (2015). Individual uncertainty and the uncertainty of science: The impact of perceived conflict and general self-efficacy on the perception of tentativeness and credibility of scientific information. *Frontiers in Psychology*, 6. doi:10.3389/fpsyg.2015.01859

Gilbert, F., & Ovadia, D. (2011). Deep brain stimulation in the media: Over-optimistic media portrayals call for a new strategy involving journalists and scientists in the ethical debate. *Frontiers in Integrative Neuroscience*, *5*(16). doi:10.3389/fnint.2011.00016

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*, 1012–1014. doi:10.1038/nature07634

Graesser, A. C., Lippert, A., & Hampton, D. (2017). Successes and failures in building learning environments to promote deep learning: The value of conversational agents. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 273–298). New York: Springer.

Greitemeyer, T. (2014). I am right, you are wrong: How biased assimilation increases the perceived gap between believers and skeptics of violent video game effects. *PLoS One*, *9*, e93440. doi:10.1371/journal.pone.0093440

Greving, H., & Sassenberg, K. (2015). Counter-regulation online: Threat biases retrieval of information during Internet search. *Computers in Human Behavior*, *50*, 291–298.

Greving, H., & Sassenberg, K. (2016). When positive information is preferred: Counter-regulating threat during Internet search. Unpublished manuscript, Leibniz-Institut für Wissensmedien, Tübingen, Germany.

Greving, H., Sassenberg, K., & Fetterman, A. (2015). Counter-regulating on the internet: Threat elicits preferential processing of positive information. *Journal of Experimental Psychology: Applied*, *21*, 287–299.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*, 555–588.

Hawkins, R. P., Kreuter, M., Resnicow, K., Fishbein, M., & Dijkstra, A. (2008). Understanding tailoring in communicating about health. *Health Education Research*, *23*, 454–466. doi:10.1093/her/cyn004

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review*, *13*, 353–383. doi:10.1023/A:1011965830686

Hróbjartsson, A., Thomsen, A. S. S., Emanuelsson, F., Tendal, B., Hilden, J., Boutron, I., … Brorson, S. (2012). Observer bias in randomised clinical trials with binary outcomes: Systematic review of trials with both blinded and non-blinded outcome assessors. *British Medical Journal*, *344*, e1119. doi:10.1136/bmj.e1119

Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research*, *12*, 301–333.

Jordan, J. E., Buchbinder, R., Briggs, A. M., Elsworth, G. R., Busija, L., Batterham, R., & Osborne, R. H. (2013). The health literacy management scale (HeLMS): A measure of an individual's capacity to seek, understand, and use health information within the healthcare setting. *Patient Education and Counseling*, *91*, 228–235.

Kimmerle, J., Bientzle, M., & Cress, U. (2017). "Scientific evidence is very important for me": The impact of behavioral intention and the wording of user inquiries on replies and recommendations in a health-related online forum. *Computers in Human Behavior*, *73*, 320–327.

Kimmerle, J., & Cress, U. (2013). The effects of TV and film exposure on knowledge about and attitudes toward mental disorders. *Journal of Community Psychology*, *41*, 931–943.

Kimmerle, J., Flemming, D., Feinkohl, I., & Cress, U. (2015). How laypeople understand the tentativeness of medical research news in the media: An experimental study on the perception of information about deep brain stimulation. *Science Communication*, *37*, 173–189. doi:10.1177/1075547014556541

Kimmerle, J., Thiel, A., Gerbing, K. K., Bientzle, M., Halatchliyski, I., & Cress, U. (2013). Knowledge construction in an outsider community: Extending the communities of practice concept. *Computers in Human Behavior*, *29*, 1078–1090. doi:10.1016/j.chb.2012.09.010

Koranyi, N., & Rothermund, K. (2012). Automatic coping mechanisms in committed relationships: Increased interpersonal trust as a response to stress. *Journal of Experimental Social Psychology*, *48*, 180–185.

Kreuter, M. W., & Wray, R. J. (2003). Tailored and targeted health communication: Strategies for enhancing information relevance. *American Journal of Health Behavior*, *27*, 227–232.

Kruglanski, A. W. (1999). Motivation, cognition, and reality: Three memos for the next generation of research. *Psychological Inquiry*, *10*, 54–58.

Kruglanski, A. W., Bélanger, J. J., Chen, X., Köpetz, C., Pierro, A., & Mannetti, L. (2012). The energetics of motivated cognition: A force-field analysis. *Psychological Review*, *119*, 1–20.

Kunda, Z. (1999). *Social cognition: Making sense of people*. Cambridge, MA: MIT Press.

Larson, J. S. (1999). The conceptualization of health. *Medical Care Research and Review*, *56*, 123–136. doi:10.1177/107755879905600201

Makoul, G. (2011). Physician understanding of patient health beliefs. *Journal of General Internal Medicine*, *26*, 574–574. doi:10.1007/s11606-011-1691-z

Miles, S. H. (2005). *The Hippocratic oath and the ethics of medicine*. New York, NY: Oxford University Press.

Molden, D. C., & Higgins, E. T. (2005). Motivated thinking. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 295–320). New York, NY: Cambridge University Press.

Moreno, J. P., & Johnston, C. A. (2013). The role of confirmation bias in the treatment of diverse patients. *American Journal of Lifestyle Medicine*, *7*, 20–22.

Moynihan, R., Bero, L., Ross-Degnan, D., Henry, D., Lee, K., Watkins, J., … Soumerai, S. B. (2000). Coverage by the news media of the benefits and risks of medications. *New England Journal of Medicine*, *342*, 1645–1650.

Muis, K. R. (2004). Personal epistemology and mathematics: A critical review and synthesis of research. *Review of Educational Research*, *74*, 317–377. doi:10.3102/00346543074003317

Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-generality and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review*, *18*, 3–54. doi:10.1007/s10648-006-9003-6

Muse, K., McManus, F., Leung, C., Meghreblia, B., & Williams, J. M. G. (2012). Cyberchondriasis: Fact or fiction? A preliminary examination of the relationship between health anxiety and searching for health information on the Internet. *Journal of Anxiety Disorders*, *26*, 189–196.
Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220. doi:10.1037/1089-2680.2.2.175
Norr, A. M., Oglesby, M. E., Raines, A. M., Macatee, R. J., Allan, N. P., & Schmid, N. B. (2015). Relationships between cyberchondria and obsessive-compulsive symptom dimensions. *Psychiatry Research*, *230*, 441–446.
Patel, V. L., Arocha, J. F., & Kushniruk, A. W. (2002). Patients' and physicians' understanding of health and biomedical concepts: Relationship to the design of EMR systems. *Journal of Biomedical Informatics*, *35*, 8–16. doi:10.1016/S1532-0464(02)00002-3
Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, *41*, 847–855.
Pieschl, S., Stahl, E., & Bromme, R. (2008). Epistemological beliefs and self-regulated learning with hypertext. *Metacognition and Learning*, *3*, 17–37.
Piko, B. F., & Bak, J. (2006). Children's perceptions of health and illness: Images and lay concepts in preadolescence. *Health Education Research*, *21*, 643–653. doi:10.1093/her/cyl034
Racine, E., Waldman, S., Palmour, N., Risse, D., & Illes, J. (2007). 'Currents of hope': Neurostimulation techniques in U.S. and U.K. print media. *Cambridge Quarterly of Healthcare Ethics*, *16*, 312–316.
Roex, A., & Degryse, J. (2007). Introducing the concept of epistemological beliefs into medical education: The hot-air-balloon metaphor. *Academic Medicine*, *82*, 616–620. doi:10.1097/ACM.0b013e3180556abd
Rothermund, K. (2011). Counter-regulation and control-dependency. *Social Psychology*, *42*, 56–66.
Rothermund, K., Gast, A., & Wentura, D. (2011). Incongruency effects in affective processing: Automatic motivational counter-regulation or mismatch-induced salience? *Cognition and Emotion*, *25*, 413–425.
Rouet, J. F. (2006). *The skills of document use: From text comprehension to Web-based learning*. Mahwah, NJ: Erlbaum.
Sassenberg, K., & Greving, H. (2016). Internet searching about disease elicits a positive perception of own health when severity of illness is high: A longitudinal questionnaire study. *Journal of Medical Internet Research*, *18*(3), e56.
Sassenberg, K., Landkammer, F., & Jacoby, J. (2014). The influence of regulatory focus and group vs. individual goals on the evaluation bias in the context of group decision making. *Journal of Experimental Social Psychology*, *54*, 153–164.
Sassenberg, K., Sassenrath, C., & Fetterman, A. K. (2015). Threat$\neq$ prevention, challenge$\neq$ promotion: The impact of threat, challenge and regulatory focus on attention to negative stimuli. *Cognition and Emotion*, *29*, 188–195.
Sassenberg, K., & Wiesing, U. (2016). Internet-informierte Patienten–Empirische Evidenz für einseitige Informationsverarbeitung und ihre medizinethischen Implikationen. *Zeitschrift für Medizinische Ethik*, *62*, 299–311.
Sassenrath, C., Greving, H., & Sassenberg, K. (2016a). Are you concerned? Using patients as protagonists in reports on medical treatments affects recipients' affective experiences and memory performance stronger than using doctors as protagonists. Unpublished manuscript, Leibniz-Institut für Wissensmedien, Tübingen, Germany.
Sassenrath, C., Greving, H., & Sassenberg, K. (2016b). The impact of communication channel on affect and memory. Unpublished data, Leibniz-Institut für Wissensmedien, Tübingen, Germany.
Sassenrath, C., Sassenberg, K., & Greving, H. (2017). It has to be first-hand: The effect of first-person testimonials in medical communication on recipients' emotions and memory. *Cogent Medicine*, *4*(1), 1354492.
Bromme, R., & Goldman, S. R. (2014). The public's bounded understanding of science. *Educational Psychologist*, *49*, 59–69.

Schwager, S., & Rothermund, K. (2014). On the dynamics of implicit emotion regulation: Counter-regulation after remembering events of high but not of low emotional intensity. *Cognition and Emotion*, *28*, 971–992.

Schweiger, S., Oeberst, A., & Cress, U. (2014). Confirmation bias in web-based search: A randomized online study on the effects of expert information and social tags on information search and evaluation. *Journal of Medical Internet Research*, *16*, e94. doi:10.2196/jmir.3044

Simeonsson, R. J., Scarborough, A. A., & Hebbeler, K. M. (2006). ICF and ICD codes provide a standard language of disability in young children. *Journal of Clinical Epidemiology*, *59*, 365–373. doi:10.1016/j.jclinepi.2005.09.009

Sinatra, G. M., Kienhues, D., & Hofer, B. K. (2014). Addressing challenges to public understanding of science: Epistemic cognition, motivated reasoning, and conceptual change. *Educational Psychologist*, *49*, 123–138.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*, 755–769. doi:10.1111/j.1540-5907.2006.00214.x

van Deemter, K. (2010). *Not exactly: In praise of vagueness*. Oxford: Oxford University Press.

White, R. W., & Horvitz, E. (2009). Cyberchondria: Studies of the escalation of medical concerns in web search. *ACM Transactions on Information Systems*, *27*(4), 23.

World Health Organization. (1992). *International statistical classification of diseases and related health problems, tenth revision (ICD-10)*. Geneva: World Health Organization.

World Health Organization. (2001). *The international classification of functioning, disability and health (ICF)*. Geneva: World Health Organization.

Chapter 5
Managing Obesity Prevention Using Digital Media: A Double-Sided Approach

Guido Zurstiege, Stephan Zipfel, Alexander Ort, Isabelle Mack, Tino G.K. Meitz, and Norbert Schäffeler

Introduction

Since the nineteenth century, Western industrialized societies have been under the strong influence of a process that has been referred to as "mediatization" (Krotz, 2001). This means that the media and even more so "media logics" (Altheide & Snow, 1979) have socially gained more and more ground in the past 200 years. Among other social "metaprocesses" (Lundby, 2009)—such as individualization, commercialization or globalization—mediatization has had a crucial influence on the structure of modern societies and the dynamics of their social differentiation until today. Thus, the media have shaped modern societies, and they have also revolutionized our individual lives. Today, the media serve many functions: they provide information and entertainment and are thus fundamental to our cognitive and motivational orientation in the world we now live in. Essentially, the media constitute an Arendtian "space of appearance" (Silverstone, 1999) and thus help us to connect with others. It is hardly surprising, therefore, that in all areas of social life, the media have large and multifaceted, although in many cases contradictory, effects

G. Zurstiege (✉)
Division for Empirical Media Research, Institute for Media Studies, University of Tübingen, Wilhelmstr. 50, 72074, Tübingen, Germany
e-mail: guido.zurstiege@uni-tuebingen.de

S. Zipfel • I. Mack • N. Schäffeler
Department of Psychosomatic Medicine and Psychotherapy, Tübingen University Hospital, Tübingen, Germany

A. Ort
Department of Communication and Media Research (DCM), Empirical Communication Research, University of Fribourg, Fribourg, Switzerland

T.G.K. Meitz
Institute of Communication Research, Friedrich-Schiller-University Jena, Jena, Germany

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_5

today. In our paper, we give an overview of some of these multifaceted effects. We do so with reference to one of the most perilous problems of Western societies: obesity and overweight in childhood and adolescence. We will show that the media are an important instrument for combating this global epidemic and at the same time one of its main causes. The following sections of the paper are organized as follows: In the next short section, we will describe the social and individual challenges associated with overweight and obesity in childhood and adolescence. Against this background, we will then delineate how the media might be seen as a major instrument of health promotion on the one hand and as a major health risk factor on the other hand. In the subsequent sections we will present results of our empirical research on the effects of digital health campaigns and serious games.

Obesity Is Not Only a Growing Individual Problem But Also a Social Problem

Obesity is one of the most prevalent and challenging public health issues among industrialized countries. The World Health Organization has described it as a "global epidemic." Obesity is defined as having a body mass index (BMI—relation of body weight to body height) above 30 kg/m^2 for adults and as a weight above the 90th percentile (overweight) or above the 97th percentile (obesity) of the age-related average for children. It is mainly the result of an imbalance between energy intake and energy expenditure with intake far outstripping expenditure (Kahn, Hull, & Utzschneider, 2006). Obesity is associated with significant somatic and psychosomatic co-morbidity and has been demonstrated to shorten life expectancy (De Niet & Naiman, 2011; Haslam & James, 2005). Whereas obesity becomes more common with increasing age, it also affects children. In the USA, 31.8% of children and adolescents between 2 and 19 years old are obese (Ogden, Carroll, Kit, & Flegal, 2014). Up to now obesity is less frequent in European countries but also alarming. In Germany about 8% of female and 10% of male children in similar age groups are overweight or obese (HBSC-Studienverbund Deutschland, 2015). Without intervention, most overweight children will stay overweight or become obese into adulthood (Reilly et al., 2003). Hence obesity prevention and treatment is becoming a very important topic and programs should be started early on to prevent overweight and obesity in adolescence and adulthood (Serdula, 1993). Therefore, obesity not only affects the individual but puts stress on the healthcare system, as it increases direct as well as indirect costs substantially.

In addition to the somatic and economic burdens, there are also social barriers for the affected subjects. Obesity puts stress on society itself because obesity is a stigma (prejudice towards a group). Obese people are thought to be lazy, silly, not determined, unattractive, and unhealthy. They are discriminated against in their working life and seem to have less prospect of promotion (Giel et al., 2012). Unfortunately, not only does obesity occur early, so does stigmatization. These convictions can be shown even in children: 10-year-old children stigmatize obese children and consider them more unlikeable and lazy, less intelligent and less attractive. Obese children are rejected as playmates by other children (Thiel et al., 2008). Media most often support and (re-)produce these beliefs. Stigmatizing ideas are often used in a discriminating manner to joke about this growing part of the population (Thiel, Giel, Thedinga, & Zipfel, 2016), which seems to work very well in comedy—at least in Germany. And media exposure increases the strength of these stigmatizing attitudes (Latner, Rosewall, & Simmonds, 2007). Stigmatization not only makes the people discriminated against feel bad but also builds up a barrier to seeking help and lowers motivation to change. One important predictor of overweight is a sedentary lifestyle. In this regard, a further and more direct role of the media in influencing obesity can be identified because of its association with an increase in sedentary lifestyle (De Craemer et al., 2012; van Stralen et al., 2012). In Germany, about two-thirds of the children between 11 and 17 years old report a screen time of 1–5 h in a survey done from 2009 to 2012 (Manz et al., 2014; Robert Koch-Institut, 2015).

Whereas digital media contributes to the problem, it can at the same time be used to educate affected individuals about obesity, stimulate new and more healthy behavior in a protective environment, contribute to more physical activity using motion-controlled games (exergames), and work as a persuasive technology to assist in changing behavior and keeping up the changed behavior. The preventive approach described in this chapter focuses on obesity in childhood as a critical time in its development. In particular, the research is concerned with major barriers like nutritional habits, physical activity, and psychosocial aspects as key factors for the development and maintenance of childhood obesity. To support life-style changes digital media is used, because it is an important part of children's lives, is available, and arouses interest. As these brief remarks show, the media have multifaceted effects on overweight and obesity. The following section provides a brief overview of research dealing with this relationship.

A Complex Relationship: How the Media Are Related to Overweight and Obesity

In the context of sports events, the media spectacularly stage healthy bodies of professional athletes. In the context of consumer magazines, the media report on current health trends and healthy lifestyles, giving advice and providing information on how to eat what, get sufficient sleep, exercise adequately, keep fit, age successfully, etc. In many cases, the media themselves may even become a crucial instrument of health promotion. Mass media like newspapers, television or radio programs disseminate health-related messages continuously, thus offering important motivation and identifying opportunities. Even more, new wearable digital media devices enable their wearers to monitor and record their physical activity. They also enable them to share and compare their fitness-data with peers who thus become involved in an open health competition, as it were. Last but not least, from the perspective of public healthcare, digital media in particular may help to promote healthy lifestyles. As compared to mass media like newspapers, television, or radio, digital media are "affordable" media. Therefore, one of the greatest revolutionary contributions of digital media is the fact that they have significantly empowered social actors on the "long tail" of society, those with less financial means. In former times those actors had no publicly perceptible voice due to their insufficient media budgets. Public healthcare is exactly such an actor in that, compared to industry standards, it traditionally has a low budget to spend on media and communication and is now digitally empowered to raise its voice perceptibly in its "own" media outlets on Twitter, Facebook, and YouTube.

Despite all of these remarkable opportunities the media in general and new digital media in particular offer for public healthcare, the media themselves can in many ways be seen as a major risk factor threatening public health. A focus on obesity and overweight in childhood and adolescence brings to light at least four interrelated risk factors associated with the media: (a) Marketing for energy-dense and nutrient poor (EDNP) food affects children's food knowledge, preferences, and eating behavior. (b) In general, screen media exposure is related to obesity and overweight. (c) Food intake accompanying media consumption transforms eating behavior. (d) As an integral part of the material home environment, the media are related to childhood obesity and overweight.

Research on food marketing to children and adolescents has concentrated on two questions. First, to what extent are children targeted by food marketing and what are the characteristics of this food marketing? Second, what are the effects of EDNP food marketing and how do children respond to it? Cairns, Angus, Hastings, and Caraher (2013) have reviewed 99 primary studies and 16 review articles that have contributed empirical answers to these two questions. Cairns et al. report, according to previous research, that most of the food advertisements for children are for foods high in fat, salt, or sugar, such as pre-sugared breakfast cereals, fast foods, and soft drinks as well as non-carbonated beverages or savory snacks. Moreover, children are proportionately more often targeted by food marketing than adults. Food marketing takes advantage of specific creative strategies while addressing children and adolescents. One of the core slogans of food marketing is "the taste of it." Most research concentrates on TV-advertising, although food marketing is increasingly deploying digital media in order to address children (Federal Trade Commission, 2012). Concerning the effects of such marketing, Cairns et al. (2013) consistently report, as Livingstone and Helsper (2004) had done about 10 years before, that research has repeatedly and consistently shown to what extent advertising affects children's food knowledge, preferences, and eating behavior: The more exposure to EDNP food marketing, the lower the nutrition knowledge. The more exposure to EDNP food marketing, the more likely food preferences will change to food high in salt, fat, or sugar. And the more changes toward these food preferences take place, the more likely corresponding purchase intentions occur.

Specific promotional strategies within EDNP food marketing have been particularly criticized. One of these strategies in food marketing is consistent with one of the major semantic tactics of advertising in general—the "appropriation of the opposing motive" (Luhmann, 2000, p. 46). According to advertising's logic, one can save money by spending it on the right brand, or one can live a sustainable life by driving the right car. The same logic regularly applies in food marketing, especially in EDNP food marketing. Here, the appropriation of the opposing motive reads thus: one can become fit and athletic by consuming EDNP food and beverages. Against this background it is surely not by accident that the vast majority of the food products advertised with athlete endorsers are energy dense and nutrient poor (Bragg, Yanamadala, Roberto, Harris, & Brownell, 2013).

Research on the relation between the amount of exposure to screen media and obesity has also concentrated on television. TV is still one of the most important sources for food marketing and, furthermore, accounts for the most screen time for children and adolescents. Today, though, TV screen time is declining while other forms of screen time (e.g., video games, computers, and smartphones) are rising. Focusing on various types of screen time, Falbe et al. (2013) found that among 7,792 adolescents throughout the USA, more screen time exposure predicts greater 2-year gains in BMI. Although these 2-year gains in BMI are relatively small, the authors indicate that cumulative effects could become clinically important (Falbe et al., 2013, pp. e1501–e1502).

Due to the fact that screen time is positively associated with BMI, but statistically independent from physical activity, the authors argue that screen time is likely further to affect dietary preferences and energy intake of adolescents. As Cairns et al. (2013) have also argued, the promotion of EDNP food often blurs the line between TV programs and commercials. According to Falbe et al. (2013), the same may happen in video games. Although video games do not contain advertising, they increasingly contain product placements or might even be completely designed as advergames, i.e., a hidden form of marketing. "With the emergence of digital media and changes in the way publishers monetize content, online advertising known as "native advertising" or "sponsored content," which is often indistinguishable from news, feature articles, product reviews, editorial, entertainment, and other regular content, has become more prevalent." (Federal Trade Commission, 2015, p. 2). It is clear from this statement by the US Federal Trade Commission that the prevalence of integrated forms of advertising has increased significantly. One of the main drivers of this development is the EDNP food industry.

Research has shown consistently that screen media exposure is positively, yet modestly, associated with children's BMI. As indicated by Falbe et al. (2013), this effect may be due to the fact that screen time affects children's and adolescents' energy intake by stimulating snacking behavior. The relationship of screen media exposure, snacking behavior, and BMI "appear[s] complicated and remain[s] poorly understood," as Carson and Janssen (2012) have argued in the context of their large-scale survey, with 15,973 children and adolescents in grades 6–10 participating in the Canadian Health Behaviour in School-Aged Children Survey (HBSC). Yet Murer, Saarsalu, Zimmermann, and Herter-Aeberli (2016) provide empirical evidence that children's food consumption while watching television is definitely associated with their BMI. Consistent with these findings, Liang, Kuhle, and Veugelers (2009) have shown that both children's television exposure as well as eating in front of the television are each independently associated with poor dietary choices and BMI.

The media are an integral part of children's home environment. Gilbert-Diamond, Li, Adachi-Mejia, McClure, and Sargent (2014) conducted a nationally representative survey with US children and adolescents, assessing the association between the presence of a TV set in children's and adolescents' bedrooms and their weight gain. Although the clear majority of the respondents (59%) reported having a bedroom TV set, the likelihood of weight gain increased with older male respondents of lower socioeconomic status. Having a TV set in the bedroom at baseline, the authors report, was positively associated with children's and adolescents' television viewing hours per day. When 2-year and 4-year follow-ups were conducted, having a TV set in the bedroom was also associated with children's and adolescents' weight gain. This indicates that the presence of a TV set in the bedroom, the authors suggest, may be an additional risk to children.

Digital Media Strategies: Sowing Bad Seeds?

In summary, one can say that while being a major instrument of health promotion, the media are also a major health risk factor. Against this background, in our own research on the prevention of childhood obesity we concentrated with a "dual focus" on new digital media strategies of public healthcare (for a detailed version of the following condensed report see Meitz, Ort, Kalch, Zipfel, & Zurstiege, 2016). This means that we looked at both the opportunities and the risks of health communication in digital media. From the perspective of public healthcare, digital media offer new, powerful, and affordable means of overcoming the attention threshold in mediatized societies. On the other hand, digital media also present new challenges to public healthcare. While provocative messages "hammering on the point" have always been an essential tool of health communication, campaigns in digital media have to rely increasingly on provocative content, because they aim to motivate unpaid peer-to-peer communication. Therefore, the tone of the initial message has to be even more aggressive (Porter & Golan, 2006).

Social media platforms provide a prominent online environment to spread public healthcare campaigns, raise awareness for health-related messages, and thus energize public health discourse. However, when applying digital media strategies, public healthcare runs the risk of losing control over the effects of its campaigns. Follow-up communication in different media platforms such as blogs, online news sites, or social networking sites picks up the campaign, comments, re-contextualizes, and evaluates. This may positively strengthen the attention focused on the campaign. The question emerges, however, how different contexts affect campaign perceptions negatively. Even if the core message of a given campaign remains the same, the argumentative re-framing of the initial message in different media contexts is likely to alter the recipients' overall perception of the initial message (for an overview of the concept of framing see Entman, 1993 and Iyengar, 1991). Thus, digital campaigns against childhood obesity, using provoking or even shocking motives, run the risk, contrary to their own intention, of offending and stigmatizing the children and families affected by obesity.

Generally speaking, in terms of how campaigns are embedded, two aspects stand out. First, the context of the initial message could either reinforce the argumentation of the campaign or impair it. Thus, the arguments of the original message would either be supported or rejected. This mechanism resembles the so-called "value framing" in political communication (Schemer, Wirth, & Matthes, 2012). Second, follow-up communication regarding the given campaign is likely to happen on different online media platforms, ranging from more traditional journalistic contexts, such as online newspapers, to social-media environments, such as Facebook or blogs. The credibility of each of these platforms is different. Both factors may cause re-contextualization which would affect the credibility and trustworthiness of the message and of the communicator (e.g., Hovland & Weiss, 1951). This is a highly relevant process since, particularly in health communication, any change in recipients' knowledge, preferences, and behavior due to campaign messages strongly depends on the credibility of the source disseminating these health-related messages. Over time, campaigns in digital media may become repurposed, repackaged and remixed in various digital media environments (for the similar problem of "repurposed" news see Vraga, Edgerly, Wang, & Shah, 2011). Thus, the source credibility of the public service initially having disseminated the campaign may fade into the background.

The object we chose for our inquiry was the 2012 Children's Healthcare of Atlanta (CHOA) campaign Strong4Life. This campaign had triggered a highly controversial discussion in the USA. It did so because it used obese children and showed them in online media talking about being stigmatized. According to the premises outlined above, our research is based on the assumption that especially such controversial campaigns disseminated via online media run the risk of being re-contextualized, with the possible negative result of jeopardazing the original communicator's intentions. In our study, we experimentally manipulated the embedment of the Strong4Life social health campaign, varying different online media environments (Facebook, blogs, online newspaper) and context argumentation (reinforcing or impairing). Since the campaign evoked strong online commentaries in blogs, social-networks, and online newspaper websites, we were able to extract realistic arguments on different platforms that reflected either an impairing or reinforcing context for the campaign selected from a previous study (Zurstiege, Meitz, & Ort, 2014).

Our study was designed as an experiment with two periods of measurement. Participating schools were asked to hand out information about the study to seventh and eighth grade students and their parents. For ethical as well as organizational reasons students from lower and higher grades were precluded from the study. After the students' and parents' consent was obtained, questionnaires for the first part of the study (including socio-demographic and health-related measures as well as cognitive and affective self-perceptions of weight) were distributed by the teachers. Two weeks later, the second part of the study was carried out in the computer labs of the schools. Participants were instructed to sit in front of a desktop computer to complete the survey. They were randomly assigned to one of six stimuli for the online experiment (EFS Survey). After viewing the website, participants were asked to complete an online questionnaire on their computer, which included source credibility perceptions, self-relevance measures and again cognitive and affective self-perceptions of body image. 749 children and adolescents aged between 13 and 18 years participated in both parts of the study. How did the type of media and context argumentation affect these children's and adolescents' self-relevance perceptions of the campaign?

In all three media environments reinforcement of the campaign message improved the self-relevance-perceptions (with self-relevance measured according to Burke & Edell, 1989). Also, the type of media predicted the evaluation of self-relevance, with campaign messages reframed on Facebook being perceived as significantly less relevant than the campaign messages reframed in a blog or (marginally significant) in online news. Also, the participants' affective self-perception concerning their own weight (sensu the discrete emotions-based scale by Sassenberg, Fetterman, Krebs, & Neugebauer, 2014) was significantly influenced by the argumentative reframing of the original message on a Facebook site, a Blog, and an online news site. There was overall, however, a reverse effect. The message reinforcing the campaign reduced the children's and adolescents' affective self-perception of their own weight, while in the impairment condition, the children and adolescents had a more confident affective self-perception of their own weight. Concerning the effects of the argumentative reframing, there was, however, no significant difference between the three media context conditions. Obviously, the argumentative contextualization of health campaigns via different digital media platform matters. However, it matters differently for rational, cognitive evaluative processes like the evaluation of relevance and emotionally driven evaluative processes like the affective self-perception of one's own weight. Moreover, the media platform staging the argumentative reframing of a given campaign also matters. But, again, it matters differently for cognitively or emotionally driven evaluative processes.

Obviously, institutions of public healthcare engaging in digital media strategies have to take into account not only that their communication will be commented upon and repackaged in different digital media contexts, they also have to consider that these contexts matter in different ways and alter the overall perception of the initial communication. In a nutshell, one can say, that only a few media platforms communicate information we perceive to be relevant for us. It is crucial for public healthcare institutions, while speculating about which digital media platforms to use as multipliers for their messages, to reach out specifically to these credible platforms and to meet their tastes in order to spread their health-related messages effectively. Many media platforms, however, communicate information that touches us emotionally, even though we do not consider them to be particularly credible.

It is important for public healthcare institutions to take into consideration that, when pursuing a digital media campaign, there is no way to bypass these media platforms. This means that public healthcare campaigns that are in some cases shocking or provoking have to be formulated in a way that stimulates rather than inhibits functional appropriation of the communication by its recipients. This clearly defines the limits of the use of provoking and shocking messages in health communication. Given the fact that one of the most frequently used strategies in health communication is to play on fear, this is particularly important. Therefore, it is worthwhile to take a closer look at how appealing to feelings of fear in health communication affects motivation.

Obesity and Overweight as Health Threats: How to Put Fear into Health Communication?

The theoretical "mainframe" of research on the effects of fear appeals in health communication is provided by the *Extended Parallel Process Model (EPPM)* (Witte, 1992). The *EPPM* (see Fig. 5.1) incorporates previous models that have tried to explain the effects of media on health-related behavior, e.g., the *Health Belief Model (HBM)* (Rosenstock, 1960), the *Parallel Response Model (PRM)* (Leventhal, 1970), and the *Protection Motivation Theory (PMT)* (Rogers, 1975). The *EPPM*, however, goes further in trying to overcome deficiencies and restrictions of its forerunners by combining their strengths in order to form a valid reflection of the processes underlying fear appeals.

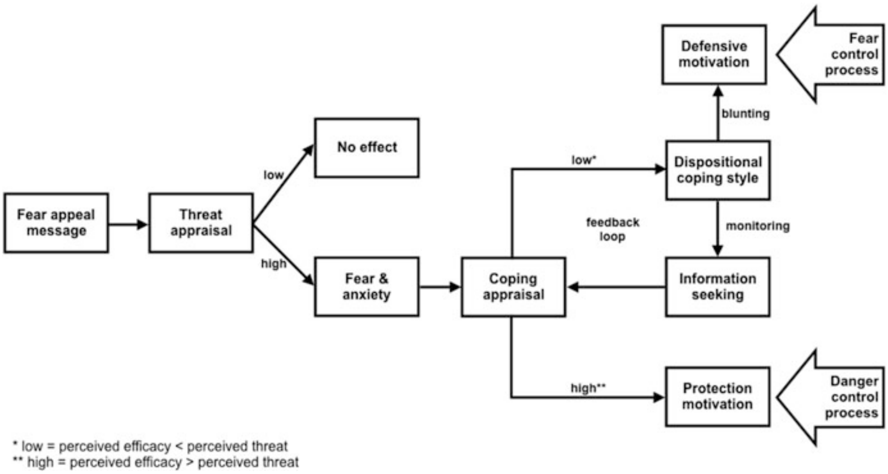

Fig. 5.1 The Extension of the *Extended Parallel Process Model* (EPPM) (Witte, 1992), adapted by So (2013, p. 79)

In Witte's EPPM model, attributes already conceptualized in the *PMT* are also relevant, namely: the recipients' perception of severity, vulnerability, self-efficacy, and response-efficacy in a given health communication message. These perceptions determine whether or not the message will be perceived to be threatening. They also influence the recipients' evaluation as to how effectively they individually will be able to avoid the given threat. The difference in the *EPPM*, however, is that these processes do not run parallel; instead, they emerge one after the other. According to this premise, the perception of a given fear appeal in health communication leads to the perception of threat in a first step. If this threat is perceived to have little or no relevance, no further message processing will take place. In the case of high threat perception, however, the result will be an emotional reaction (most likely fear), which results in further processing of the message and its content. Then, in a second step, individuals are likely to evaluate their resources to master and overcome the threat. This evaluative process determines whether a fear appeal will lead to a defensive motivation or protection motivation, which accordingly results in maladaptive or adaptive behavior, as specified in the *PRM* by Leventhal (1970).

In the case of a given threat that is accompanied by a higher level of perceived efficacy, adaptive danger control behavior is likely to occur. This behavior is motivated by the individuals' need to shield oneself from the threat. In practice, this means that the recipient is more likely to accept and thereupon follow the recommended course of action. If the recommended course of action is, however, perceived to be ineffective, fear-control processes will be stimulated. These processes are likely to spur message resistance and, finally, defiance of the recommended course of action. In extreme cases the result might even be a behavior change diametrically opposed to the originally intended behavior.

In the *EPPM*, fear as an emotional reaction to a fear appeal is directly related only to defensive motivation, thus merely indirectly affecting the individuals' motivation to protect themselves from the perceived threat. Perceived threat therefore determines both the amount of motivation and the intensity of the individuals' reaction, whereas perceived efficacy is crucial for the chosen course of action (danger control vs. fear control). The interaction between perceived threat and perceived efficacy determines the result of message processing. According to Witte (1992), both perceptions contribute to an adaptive behavior change. Making these specifications, the *EPPM* amends the assumptions of the *PRM* and *PMT* and thus allows for a reasonable prognosis of danger control and fear control reactions to a given stimulus.

The *EPPM* has frequently been adopted in order to explore the persuasive effects of health communication with regard to various health-related topics and audiences (for a meta-analytic overview see: Casey, Timmermann, Allen, Krahn, & Turkiewicz, 2009; Ruiter, Kessels, Peters, & Kok, 2014; Ruiter, Verplanken, Kok, & Werrij, 2003; Witte & Allen, 2000). Within EPPM's framework, research has shown that threatening health messages carry the inherent risk of causing contraindicated effects like avoidance, denial, or reactance. According to the *EPPM* as well as to other existing research, these contraindicated effects mainly depend on whether or not the message includes information or advice about effective possibilities to avoid the potential threat (see for a meta-analytic overview Popova, 2012). Thus, scaring people into behavior change by focusing only on the threat component of a message does not seem to be effective. On the contrary, focusing on efficacy appears to be more central and promising in achieving adaptive behavioral changes. Hence the perception of threat and the perception of efficacy are central when it comes to stimulating behavior change via health communication using provocative messages.

In summary, faced with the fact that our society is permeated with media and communication, health communication must be loud and sometimes even provocative to overcome the current threshold of attention. Yet, the messages must also tell the people they are targeting what they can do to avoid health threats. What does this mean in the context of a preventive campaign against obesity and overweight in childhood and adolescence? What can the people affected by such a campaign do to effectively prevent obesity and overweight?

Addressing Children's "Health Managers": What Can They Do to Make a Change?

Health communication concerning obesity, nutritional behavior and weight change touches on deeply rooted habits, on routines, and on lifestyles. Media-based attempts to overcome barriers to prevention and treatment of obesity, therefore, have to account for the length of time it takes to change these habits and routines. For this reason, health campaigns have to become and, what is even more challenging, *remain* socially relevant in the lives of their target groups. Health campaigns are efficient when, and insofar as, recipients start to talk about and to act upon them, and keep doing so in their daily lives. This is especially true in the case of overweight and obesity in childhood and adolescence. The reason is that overweight and obesity in childhood and adolescence are embedded in multiple social and institutional contexts involving interactions with medical staff, friends (peers), and family members. In all these relationships interactions with relevant others occur in various forms—e.g., when getting medical treatment, having meals together, spending time together, doing physical exercise, enjoying leisure activities, or exchanging messages on Facebook and other social media. One can say that social interactions, therefore, are crucial follow-up factors determining the overall effect of digital media interventions and have to be understood in order to evaluate and predict their outcomes.

Concerning childhood overweight and obesity, of course, parents are highly relevant for children and adolescents (Walsh et al., 2014). Though they may not be directly affected by obesity and overweight themselves, as "health managers", so to speak, they affect the health status of their children (e.g., by "maternal provisioning" according to Maher, Fraser, & Lindsay, 2010; Neysmith & Reitsma-Street, 2005). Despite their influence they are, however, often wrong when it comes to health-related topics, e.g., judging the weight of their children (Eckstein et al., 2006).

Obviously, this is a serious problem, as parents have a big impact on the eating habits, leisure behavior, and the physical activities of their children. In the case of obesity and overweight in childhood and adolescence, parents are, therefore, often both part of the problem and at the same time part of the solution as an existential source of appreciation and support.

Thomas et al. (2014) have provided initial groundwork exploring how parents and children mutually respond and interact in relation to a digital anti-obesity campaign. The authors conducted qualitative interviews with 159 parents and 184 children across 150 family groups. When asked, parents and children expressed very different interpretations following the campaign. In some cases, children even articulated their interpretation that the anti-obesity campaign had placed responsibility on *them* to remind their parents not to gain weight in order to set a good example for their children. Most obesity campaigns, however, place the responsibility on parents, presuming that addressing parents' perception of responsibility and efficacy is crucial for the success of the communicative intervention.

Past research on the effects of health-related information has been primarily concerned with the question of how health-related information changes attitudes, intentions, and behavior (Ajzen, 1991; Leventhal, 1970; Rogers, 1975, 1983; Rosenstock, 1960; Witte, 1992). Exposure to a single campaign message, however, is rarely able to change complex attitudes and long-established nutritional habits. Therefore, for campaigns that want to accomplish more than just to attract attention, in many cases it is a fruitful strategy not to try to change the dietary habits, but to engage its target audience in the further exploration of the subject, i.e., to trigger and to channel further information-seeking behavior (for a overview see Johnson & Case, 2012, p. 17; see also Chap. 4 in this book). According to the premises of the *EPPM*, one can assume that the information seeking behavior of parents will be determined both by the way parents perceive their children's susceptibility of becoming obese or overweight *and* by the way parents perceive their own efficacy to make a change.

Fear as a consequence of perceived health threats and shame as a consequence of perceived yet abandoned responsibility are strong feelings that can even cause durable reverse effects unintended by health campaigns addressing children and their parents. As we have seen so far there is a considerable body of research dealing with the question of how to "tune" health communication effectively. The majority of this research aims to determine how to raise awareness for the vulnerability of a target audience to particular forms of health threat, yet, at the same time, to motivate intended behavior such as information seeking. Within this framework, gaining attention and stimulating intended behavior are in latent contradiction. Although campaigns evoking strong feelings are likely to travel extensively through digital mediascapes, their recipients tend to oppose these strong feelings when left alone with the question as to what to do in order to circumvent given health threats. Putting on provocative health campaigns, therefore, is a two-edged strategy bearing chances and risks. What are, then, strategies to circumvent these risks? One of these strategies may be choosing media for health communication messages that, per se, engage its recipients in a playful mood, diminishing the fear factor. Against this background employing serious games as an instrument of health communication is one fruitful strategy to engage the target audience and to motivate behavior change.

Designing a Serious Game for Kids' Obesity Prevention (KOP)

While risk communication in media mainly tries to generate motivation for behavior modification towards a healthier lifestyle, applying a threatening scenario has not been shown to educate or motivate children to learn about this topic. Sustainable intrinsic motivation for learning about useful behavior and lifestyles is associated with a significantly lower risk of becoming obese. This motivation should be encouraged by first making the children curious about the topic and then providing them with possibilities for getting information. The next step would be to introduce the information in an experimental environment and give the children reinforcement for showing learning success and behavioral changes. Gamification of learning content may help, as it enhances the motivational aspect of a learning environment. Gamification means using elements of game designs in nongame contexts in order to increase interest, motivation, and enjoyment, and thus promote engagement (Brown et al., 2016; Cugelman, 2013). Gamification may also support forming a new attitude in the learner in which a virtualized reality encourages trying out different behaviors in a safe context. But gamifying content on its own does not necessarily lead to a better outcome. It may even be harmful (Hors-Fraile et al., 2016) if it leads to frustration or demotivation due to inappropriate use of persuasive strategies or gaming elements.

Screen time is regarded to be a risk factor for childhood obesity, as it is associated with a more sedentary lifestyle. There are several studies that have been able to show this association (Arango et al., 2014; Ballard, Gray, Reilly, & Noggle, 2009; Boone, Gordon-Larsen, Adair, & Popkin, 2007; De Craemer et al., 2012; Dietz & Gortmaker, 1985; van Stralen et al., 2012). Therefore, at first sight it looks like developing a serious game to fight obesity is not a smart idea, as computer games certainly contribute to the problem of screen time.

On the other hand, we know from social work and adult education that approaches with low barriers oriented to the real world of the target group are needed to involve people in learning out of institutional contexts. In the last several decades digital media has revolutionized our lives. We live in an information society in which information is widely available and easy for nearly everybody to access with mobile devices like smartphones at nearly any time and place. This development has also changed the way in which children seek information. A precondition is the availability of digital media in households. In 2011, already 25% of all teenagers from 12 to 19 years old in Germany owned a smartphone. This increased rapidly to 47% in 2012 and reached 92% in 2015 (Medienpädagogischer Forschungsverbund Südwest, 2015). Only a few children possess a smartphone when they start elementary school. In 2014, 29% of the children from 10 to 11 years old owned a smartphone, and 55% of the children from 12 to 13 years old. Children access the internet or play games more often using a computer or tablet. While in 2007, 73% of private German households had a computer and 65% had access to the internet (Czajka & Mohr, 2008), this increased in 2014 to 97% of households with children having a computer and 98% of the children having the possibility to access the internet at home. About a quarter of 10- to 11-year-olds and nearly half of the 12- to 13-year-olds have their own computers. Tablet computers, which are as easy as smartphones to use for gaming or internet access, are available in 19% of all households with children from 6–13 years old (Medienpädagogischer Forschungsverbund Südwest, 2014). Children's usage of new media is quite high. Seventy-two percent of the children between 8 and 9 years old and 91–98% of the children between 10 and 13 years old use a computer (Medienpädagogischer Forschungsverbund Südwest, 2014). In total, more than a quarter of the 11- to 17-year-olds use a computer or the internet for more than 2 h per day. This statistic increases with age and with lower social status (Robert Koch-Institut, 2015).

As children in Germany are used to playing computer and smartphone games and the technology is available for nearly all of them, this medium can be used for gamification, as it is highly attractive, possibly interactive, and it is used daily. It is clear that digital media do have an advantage that could be used in obesity prevention. The disadvantage of digital media promoting a sedentary lifestyle could be possibly reduced by using special interfaces to play a computer game, e.g., motion control.

Does Digital Media Help Children to Learn About Barriers for Obesity Prevention?: Serious Games

Following the national obesity prevention and treatment guidelines (Hauner, Hebebrand, Ried, Müller, & Stumvoll, 2012; Wabitsch & Kunze, 2015), and the consecutive intervention program to fight childhood obesity of BZgA (Bundeszentrale für gesundheitliche Aufklärung (BZgA), 2006, 2012) a learning game should have three pillars as a basis:

– Nutrition: knowledge about healthy nutrition and regular, balanced meals
– Psychosocial factors: generate ideas about how to reduce stress and to practise relaxation
– Physical activity: encourage physical activity.

The main emphasis of many programs lies on the first or last pillar: knowledge about healthy nutrition and regular balanced meals or physical activity. While the aim of serious games (or educational video games) is to improve knowledge and to support behavioral changes, exergames provide a mixture of video gaming and physical activity by using the player's motion to control the game. In a systematic literature review (Mack et al., 2017), we found a group of studies where the games have been developed to teach topics like improving the intake of fruits and vegetables or reducing the intake of sweets. The use of computer games in nutrition education showed a greater increase of knowledge and improvement of eating habits in children compared to a traditional teaching approach or the usage of multimedia presentations. Using an educational multimedia game or internet-based interventions, some studies have been able to increase the learners' intake of fruits and vegetables or decrease the consumption of sweetened beverages and snacks. A computer game to evaluate diets individually did not show any superior knowledge about nutrition in this group, but it did result in better self-care practices and increased activity time. In one study, the use of an online video game in a school setting to increase activity and knowledge about healthy eating and to improve health behaviors showed a high acceptability of the game. It also enhanced diet self-efficacy, positive attitudes towards healthy foods, and nutrition knowledge. Some serious games focus on a healthy amount of exercise and a positive attitude towards it, but the promotion of physical activity using an interactive multimedia approach did not show any significant impact on physical activity in everyday life.

The second pillar, the psychosocial factor, was not addressed directly by the games themselves but sometimes it was part of the whole interventional program if the games were not the only tool used. The involvement of parents is very important in this regard, as parents have a role model function for their children and influence them in their daily behavior and thinking. To date parents have been involved in these types of programs mainly for obtaining information about their children—in only a few studies did they participate in a program or receive information themselves.

Serious games are attractive for children and help them to learn about nutrition topics or even set behavioral changes in motion, but effects on weight loss or the increase of activity time are rare. Whereas the cognitive approach is quite often used to teach children about nutrition, the focus is mainly on teaching the food groups or raising the level of vegetable and fruit consumption. An approach to teaching the concept of energy density is lacking, and subsequently there is little guidance for deciding whether a food item fits into balanced and healthy daily meals, or for enabling children to choose between high and low energy items of similar taste. Psychosocial aspects are so far also lacking, and the involvement of children's parents should be cultivated. They serve as role models and they have the most control over the child's environment, including the food offerings and spare time activities.

Is It Possible to Physically Activate Children While Playing the Game?: Exergames

Most often games are controlled with a mouse and keyboard or specific controllers for gaming consoles that require or at least suggest that a player be sitting. More recently, the control of a game using body movements has become an interesting innovation for a completely new generation of games. As a first step, sensoric mats were used to make the player move and press small switches associated with movements or actions in the game. In a second developmental stage, sensoric handles tracked the movements of the arms, e.g., to play tennis. The third stage used the reflection of an infrared signal to calculate a body model of a person in front of the sensor and track their movements to steer something, for instance, an avatar, in the game. With the availability of these body-motion controlled interfaces, exergames have been invented where the player's motion is tracked and transferred into the game in order to control it. Their developers promise to make the computer games not only less responsible for the growing problem of obesity but also a positive possibility for fighting against it (Biddiss & Irwin, 2010; Staiano & Calvert, 2011).

In our review (Mack et al., 2017) we identified several studies that tried to improve the activity level and showed an increase in physical exercise while playing. Many of them used a sensor to transmit the player's motion into the game. However, these studies have so far not included any control groups. Compared to non-gaming or non-active conditions in a pre-post design, exergames show small interventional effects on weight or activity or other metabolic-related parameters.

This type of game is able to stimulate physical activity, but most often the activity is limited to the gaming time. Without proper instruction, just providing an exergame will not change anything in the physical activity of a child's everyday life. For the player, the possibility to try an exergame seems to be highly inviting, as it is chosen more often compared to an inactive online game. Exergames are widely accepted and impressive because of the use of new technology. But as this novelty effect passes, the interest in the game decreases over time. The combination of exergame and serious game may be more effective in changing behavior. Only a few studies have tried to combine the approaches of exergames with classical serious games about nutrition. In this combined type of game, parents are often involved not only in getting information about their children but also in actively playing with them in a multiplayer-setting, or in supporting their child.

How Can We Translate Obesity Prevention Topics into a "Serious Exergame"?

Besides their cognitive function of providing information about healthy nutrition for the learning individual, serious games and exergames mainly have a high motivational function: They provide a motivating environment for knowledge transfer. To reach the objective of developing a serious game for prevention of obesity, our approach was to set the player in a physically active state with the use of a motion sensor. The priority would be, unlike the entertainment objective of exergames, to provide knowledge in a gamified way. Therefore the task is to avoid making the player sweat through jumping. There are two reasons for this: To motivate overweight or already obese children strenuous physical challenge would be a huge barrier to motivating overweight or obese children to play the game. The other reason is that motivating and sensitizing them for only a bit more physical activity everyday does not have to be synonymous with taking sports to the physical limit.

For the Kids' Obesity Prevention (KOP) game, a framing story of a child living in a medieval fantasy-world is told by a storyteller showing pictures in an ancient book. The story is about the competition of two villages to regain knowledge about nutrition and a healthy lifestyle. The competitive element of the story is meant to motivate particularly boys to engage in the game. The player is supposed to support the child of the smith in mastering the tasks—a supportive element that is meant to motivate particularly girls to engage in the game. After the presentation of the story in a video, the player has to move an avatar in a three-dimensional medieval world to walk from the site of one task to another. The player's physical movement that is necessary to make the avatar move in the game is to walk on the spot, lifting the knees upwards. There is no need to do this very fast—a slow but intense walk (similar to Nordic walking) stimulates the circulation while carefully activating the joints. This element of controlling parts of the game with body movement fits the third pillar (physical activity) of the obesity prevention guidelines described above.

As motion control devices are still very novel and a big version of the 3D-world is projected onto a wall or screen, this fascinating technical aspect is very inviting to the children. But using such a motion control for a serious game has limitations. Not only does the effect of novelty shrink quite fast but current motion devices also present problems in steering a cursor precisely. This may be suitable for menu-like structures in other games, where a cursor has to be moved to select quite big buttons. Using intuitive gestures, like taking a virtual object with the hand, is at the moment too complex a control for use in the typical tasks needed in learning games. For the target group of children aged 9–11 years, the challenge of steering the game using tracked body movements runs the risk of overpowering the main objective, which is to expand the knowledge about relevant topics in obesity prevention. A second limitation is that the motion control sensor is not commonly available in households, which would restrict the possibilities for using the game to one-player settings in research or therapeutic environments. Because of these limitations, the "mini games" need a more common input device. We therefore decided to use touch screens as a very intuitive possibility of controlling a computer game.

Gamifying the Learning Content to Mini Games

As classical computer games are common leisure-time activities for most of today's children and represent an important part of their daily life, a playful approach is expected to make computer learning games accessible to children and to promote learning outside of an institutional context (see also Chap. 7). The cognitive part of the learning game consists of different mini games that make the information accessible to the player by giving feedback about the players' activities.

Different types of gaming and edutainment elements are used to communicate and practice the learning content:

- Animated presentations to introduce new information
- Experimental games to deepen the understanding of the presented information and promote curiosity about the topic
- Exercises to practice or experience introduced topics
- Reflective games to make the player think about their own behavior
- Feedback on reflective games with information about possibilities for moderately changing it
- Competitions with time pressure to show new knowledge and for reinforcement
- Quiz to show new knowledge and for reinforcement.

A common concept in programs about healthy nutrition is to introduce the healthy eating pyramid or circle to explain the food groups and the recommended food intake out of each group (Serrano, 2004). In the nutrition game we therefore start with an animated explanation of these groups. With regard to the target group of children who attend primary school, the mini games should not contain too much written text, so a speaker presents the instructions and information. To encourage better understanding and as a reminder, the groups are represented by animals as symbols for each group, e.g., cereals and potatoes by a hamster. These animals are intended to help the children remember the categories more easily and to sort through new food items. To make the children have a good look at the groups and deepen their knowledge about food groups, they have to pop balloons carrying food items above the correct baskets in the sorting game. If they make a mistake, they get direct feedback, as the wrongly placed item jumps out of the basket (Figs. 5.2 and 5.3).

Most often learning content about healthy nutrition is limited to this level. This leaves the subject with only rudimentary knowledge where there seems to be a strong tendency to divide food into good and bad items. Differences within one group aren't mentioned. There is a big difference in energy intake, for instance, between whether someone eats an apple or eats dried apple rings (Rouhani, Haghighatdoost, Surkan, & Azadbakht, 2016). At another game level, the concept of energy density is introduced in an animated presentation with comparisons of food items, e.g., one 100-g bar of chocolate contains as much energy as 2.8 kg of tomatoes or 1 kg of apples. Equipped with this idea, the learners get access to a food laboratory where the components of different food items can be analysed and their relationship to the energy density is visualized and explained for each food group (Fig. 5.4). The idea of replacing favorite food items containing high energy with similar low-energy ones is introduced (e.g., a chocolate bar is replaced by chocolate pudding). This knowledge is tested afterwards in a quiz to prepare them to think in high- to low-energy alternatives. The ability to transfer the knowledge to new unknown food items is tested in the game "turtle run." The children are encouraged to help the turtle win the race against the kangaroo. They are presented pairs of food items that are similar and could be alternatives. Every time they choose the correct one with less energy the turtle driving a roller skate gets a boost of

Fig. 5.2 Explanation of food groups using animals as a reminder

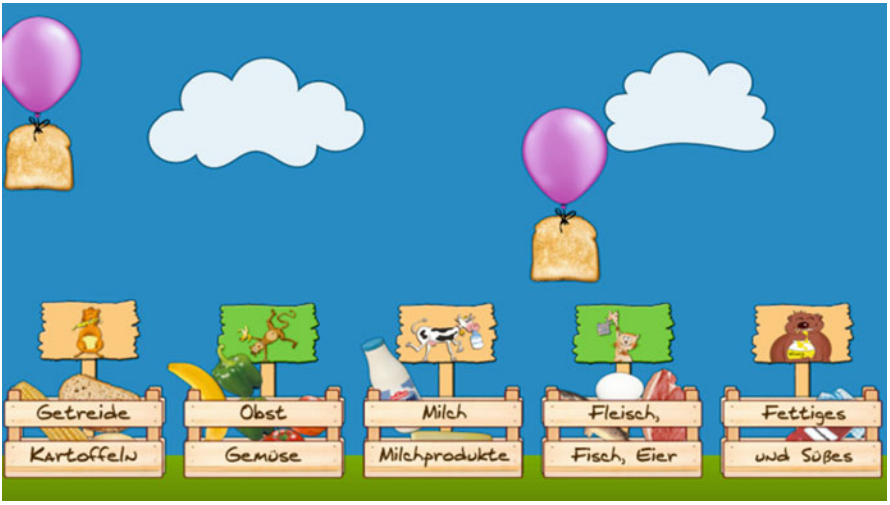

Fig. 5.3 Balloon game to assign food items to groups

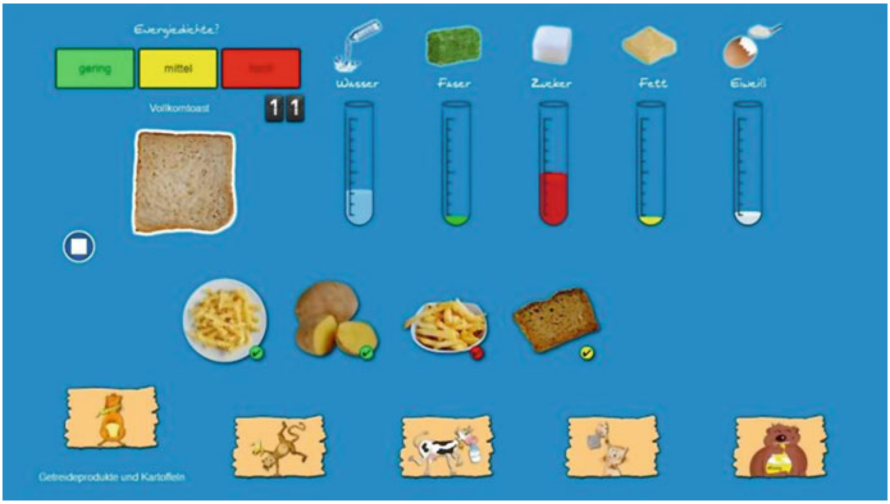

Fig. 5.4 Food laboratory to analyse components and rate the energy density

rocket propulsion—choosing the wrong one with more energy gives the kangaroo a considerable start, then there is some delay until the next pair appears.

The learning content about nutrition is framed by the reflective task of entering food into a day's plan for three main meals and two between-meals. The provisions included in the game have been analysed for gender, age, activity levels, size, and weight in accordance with the recommendations of the German Nutrition Society (OGE). Participants get feedback about whether they have entered too much or too little food in the plan, and whether the food groups are balanced well. If there is too much food out of the meat group, for example, the feedback gives alternative ideas about what to eat instead. Prohibiting food has been shown not to be a good way to promote behavior change (Fig. 5.5).

In a second nutritional track, the energy intake through liquids is introduced and explained critically, as this energy does not add to a feeling of satiation. Besides animated presentations (e.g., depicting how long you have to play soccer or swim to burn the energy of a soft drink), a quiz is provided to guess the sugar content of different beverages. These games give "food for thought": The children learn that pure apple juice contains as much sugar as soft drinks like cola. To burn the energy of 1 L of soft drinks or pure juice one has to swim about 90 min or play soccer for about 2 h. This introduces the idea of balancing energy intake and energy expenditure with being physically active.

The second pillar of obesity prevention regarding psychosocial factors is introduced by games explaining the difference between good stress and distress, with examples out of the daily life of primary school children. They are encouraged to balance their leisure-time activities in choosing activities that involve physical exercise like cycling, creative ones like doing handicrafts, and passive ones like listening to music. Possibilities for including more moderate physical activity in

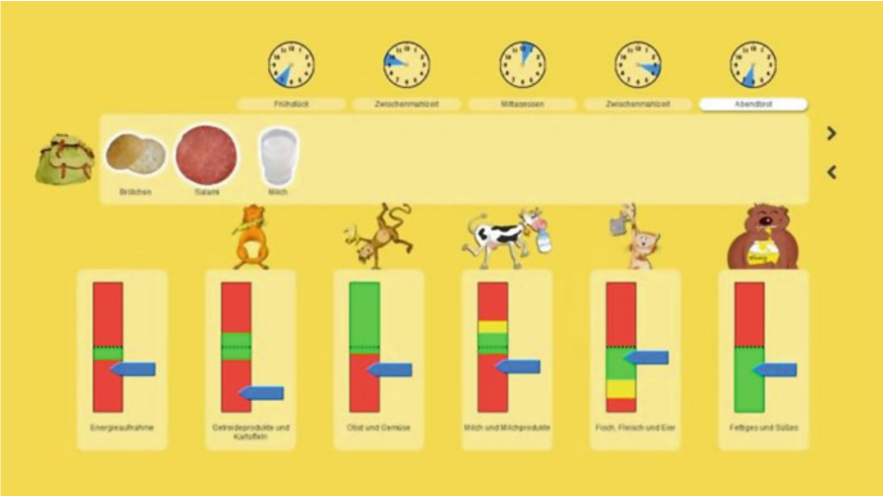

Fig. 5.5 Analysing provisions for 1 day

their daily routine (e.g., taking the stairs instead of using the elevator, cycling or walking instead of taking the bus) are intended to help them change to a more physically active lifestyle. The intention is also to show possibilities that do not necessarily mean doing strenuous sports. These considerations are made deeper in a reflection about activities they are doing and planning a balanced week. The third pillar of obesity prevention regarding physical activity is supported through this lifestyle reflection and intervention as well as through the example of the moderate activity that is necessary to move the avatar through the 3D world of the game. A relaxation exercise completes the mini games in this track.

Does It Work in a School Context?

In a first study the serious game was offered in a fourth class of a primary school to test its feasibility as well as the acceptance by the children. $N = 13$ children were of normal weight, $n = 2$ male and $n = 1$ female were overweight with a weight above the 90th percentile, and $n = 1$ female was obese with a weight above the 97th percentile of the age-related average. The children as a class were shown the trailer story. In one-player sessions they had to walk in front of the motion control to steer the avatar to the site of the first mini game. The mini games were presented on tablet computers. In this study the participants played seven out of 14 mini games in three blocks. After each of these blocks they again walked in front of the screen to steer the avatar with motion control to the next site. In the semi-structured interviews that followed, the KOP-game got a good to very good assessment from the children. The motion control of the avatar was mentioned by half of the children as a very good and funny part, seven highlighted the mini games as the best thing, and two were very impressed with the medieval landscape used in the 3D motion-controlled part. In a paper-pencil test before and after the game they showed a significant increase of knowledge about healthy nutrition.

Due to the external conditions of a school context, it was necessary to separate the elements with motion control from the mini games. For a cluster-randomized controlled interventional study with $N = 83$ pupils ($n = 8$ were overweight, $n = 6$ obese) in six 4^{th}-year classes of a primary school, it was necessary to have them play the mini games as a whole class. The 14 mini games were presented to the children in two sessions of about 45 min each (Zipfel, 2015). Even with this setting, it took 8 weeks to finish the study. In our experience this requires exceptional effort from the involved teachers and principal, which is not feasible in the normal school routine. Even under these conditions, the acceptance of pupils and teachers was again very good. It was possible to show that the interventional group that played the KOP-game showed a better outcome in learning about healthy nutrition than the control group, which simply got a brochure designed for primary school children to learn about healthy nutrition and reflect about their own behavior (Brüggeman, Braukmann, & Tust, 2014).

As discussed in the preceding sections, interaction of children with their parents is very important for an intervention that tries to change the nutrition and lifestyle of primary school children. Therefore the KOP-game should be embedded in a web platform in order to enable interaction with others (social-interactive function). This makes it possible for the children to show the mini games to their parents and learn and discuss the topics within the family. Further research will deal with a clinical context to enhance the clinical treatment of obesity in children as well as in adults who attend programs for weight reduction or for preparation of bariatric surgery.

In Summary: When Do Digital Media Help More Than Harm?

Scrutinizing the social contexts of childhood obesity, it is clear that the media in general, and digital media in particular, are without any doubt crucial factors. We have argued, however, that the media are a double-edged sword. This means that with regard to the "global epidemic" of obesity, the media can be both highly problematic and, at the same time, highly beneficial. There is ample empirical evidence indicating that the excess in media exposure, particularly in screen media exposure, is significantly associated with gains in a child's BMI. As research has repetitively shown, this effect may partly be due to an excess in the exposure to advertising-financed media that markets energy-dense and nutrient-poor food. As some observers have stated, the negative impact of the media *themselves* may be even greater than the effects of problematic media content. "It is the *medium* itself that should concern us, and not merely the content of young children's experiences with screen media" (Sigman, 2012, p. 90).

The consistent evidence of the media's potentially harmful effects notwithstanding, the media are, likewise, frequently conceptualized as an effective instrument for overcoming barriers to prevention and treatment of obesity. Digital media have empowered public healthcare to reclaim a considerable voice in a media-saturated world. Digital media have empowered public healthcare to reach its target groups efficiently despite notoriously low communication budgets. Digital media have thus empowered public healthcare to energize public health discourse. Against the background of our research we are convinced that digital media campaigns as well as digital exergames can be employed to literally mobilize public healthcare's target groups, to inform not to patronize them, to entertain not to frighten them, and eventually to change sedentary lifestyles.

However, we have also clearly recognized that all media-based attempts to overcome barriers to prevention and treatment of obesity face the same challenge. And this challenge comes to a head in the concept of attention. In order to be effective, every health campaign has to overcome an attention threshold. If it does so by becoming too provocative, however, it runs the risk of being misused as semantic raw material for counterproductive follow-up processes in social media, such as the stigmatization of affected individuals. Furthermore, if a health campaign puts too much emphasis on fear appeals in order to overcome the attention threshold of its target groups, it runs the risk of being filtered out by the attempts of the target groups to prevent cognitive dissonance. Thus, the core problem for public healthcare campaigns can be pinpointed with the question of how to *gain* the attention of its target groups properly. The core problem for digital media campaigns employing digital exergames can be pinpointed with the question of how to *keep* the attention of its target groups. The goal is to engage the target groups with effective tools for a longer time, allowing them to consolidate knowledge, to appropriate information, and to change habits. The fascination triggered by new media technologies and platforms such as exergames is ephemeral. Given the rapid changes in the esthetic and narrative appeal of digital games in general, serious games constantly run the risk of "esthetic aging." Providing stimulating incentives, especially those that can be converted into values and actions in real-world contexts, is a promising strategy to make digital experiences lasting and real.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. doi:10.1016/0749-5978(91)90020-T

Altheide, D. L., & Snow, R. P. (1979). *Media logic*. Beverly Hills, CA: Sage Publications.

Arango, C. M., Parra, D. C., Gomez, L. F., Lema, L., Lobelo, F., & Ekelund, U. (2014). Screen time, cardiorespiratory fitness and adiposity among school-age children from Monteria, Colombia. *Journal of Science and Medicine in Sport*, *17*(5), 491–495. doi:10.1016/j.jsams.2013.09.010

Ballard, M., Gray, M., Reilly, J., & Noggle, M. (2009). Correlates of video game screen time among males: Body mass, physical activity, and other media use. *Eating Behaviors*, *10*(3), 161–167. doi:10.1016/j.eatbeh.2009.05.001

Biddiss, E., & Irwin, J. (2010). Active video games to promote physical activity in children and youth: A systematic review. *Archives of Pediatrics & Adolescent Medicine*, *164*(7), 664–672. doi:10.1001/archpediatrics.2010.104

Boone, J. E., Gordon-Larsen, P., Adair, L. S., & Popkin, B. M. (2007). Screen time and physical activity during adolescence: Longitudinal effects on obesity in young adulthood. *International Journal of Behavioral Nutrition and Physical Activity*, *4*(Act. 4), 26. doi:10.1186/1479-5868-4-26

Bragg, M. A., Yanamadala, S., Roberto, C. A., Harris, J. L., & Brownell, K. D. (2013). Athlete endorsements in food marketing. *Pediatrics*, *132*(5), 805–810. doi:10.1542/peds.2013-0093

Brown, M., O'Neill, N., van Woerden, H., Eslambolchilar, P., Jones, M., & John, A. (2016). Gamification and adherence to web-based mental health interventions: A systematic review. *JMIR Mental Health*, *3*(3), e39. doi:10.2196/mental.5710

Brüggeman, I., Braukmann, M., & Tust, D. (2014). *So macht Essen Spaß. Entdeckerheft für Grundschulkinder*. Bonn: Aid Infodienst Ernährung, Landwirtschaft, Verbraucherschutz e.V..

Bundeszentrale für gesundheitliche Aufklärung (BZgA). (2006). *Tut Kindern gut! Ernährung, Bewegung und Entspannung*. Köln: Bundeszentrale für gesundheitliche Aufklärung.

Bundeszentrale für gesundheitliche Aufklärung (BZgA). (2012). *Leitfaden Qualitätskriterien für Planung, Umsetzung und Bewertung von gesundheitsfördernden Maßnahmen mit dem Fokus auf Bewegung, Ernährung und Umgang mit Stress*. Köln: Bundeszentrale für gesundheitliche Aufklärung.

Burke, M. C., & Edell, J. A. (1989). The impact of feelings on ad-based affect and cognition. *Journal of Marketing Research*, *26*(1), 69–83. doi:10.2307/3172670

Cairns, G., Angus, K., Hastings, G., & Caraher, M. (2013). Systematic reviews of the evidence on the nature, extent and effects of food marketing to children. A retrospective summary. *Appetite*, *62*, 209–215. doi:10.1016/j.appet.2012.04.017

Carson, V., & Janssen, I. (2012). The mediating effects of dietary habits on the relationship between television viewing and body mass index among youth. *Pediatric Obesity*, *7*(5), 391–398. https://doi.org/10.1111/j.2047-6310.2012.00049.x

Casey, M. K., Timmermann, L., Allen, M., Krahn, S., & Turkiewicz, K. L. (2009). Response and self-efficacy of condom use: A meta-analysis of this important element of AIDS education and prevention. *Southern Communication Journal*, *74*(1), 57–78. doi:10.1080/10417940802335953

Cugelman, B. (2013). Gamification: What it is and why it matters to digital health behavior change developers. *JMIR Serious Games*, *1*(1), e3. doi:10.2196/games.3139

Czajka, S., & Mohr, S. (2008). Informations-und Kommunikationstechnologien in privaten Haushalten. *Wirtschaft und Statistik*, *9*, 764–771.

De Craemer, M., De Decker, E., De Bourdeaudhuij, I., Vereecken, C., Deforche, B., Manios, Y., … ToyBox-study Group, ToyBox-study Group. (2012). Correlates of energy balance-related behaviours in preschool children: A systematic review. *Obesity Reviews*, *1*(13 Suppl), 13–28. doi:10.1111/j.1467-789X.2011.00941.x

De Niet, J. E., & Naiman, D. I. (2011). Psychosocial aspects of childhood obesity. *Minerva Pediatrica*, *63*(6), 491–505.

Dietz, W. H., & Gortmaker, S. L. (1985). Do we fatten our children at the television set? Obesity and television viewing in children and adolescents. *Pediatrics*, *75*, 807–812.

Eckstein, K. C., Mikhail, L. M., Ariza, A. J., Thomson, J. S., Millard, S. C., & Binns, H. J. (2006). Parents' perceptions of their child's weight and health. *Pediatrics*, *117*(3), 681–690. doi:10.1542/peds.2005-0910

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, *43*(4), 51–58. doi:10.1111/j.1460-2466.1993.tb01304.x

Falbe, J., Rosner, B., Willett, W. C., Sonneville, K. R., Hu, F. B., & Field, A. E. (2013). Adiposity and different types of screen time. *Pediatrics*, *132*(6), e1497–e1505. doi:10.1542/peds.2013-0887

Federal Trade Commission. (2012). Facing facts. Best practices for common uses of facial recognition technologies.

Federal Trade Commission. (2015). *Enforcement policy statement on deceptively formatted advertisements*. Washington, DC: FTC. Retrieved from https://www.ftc.gov/system/files/documents/public_statements/896923/151222deceptiveenforcement.pdf.

Giel, K. E., Zipfel, S., Alizadeh, M., Schaffeler, N., Zahn, C., Wessel, D., … Thiel, A. (2012). Stigmatization of obese individuals by human resource professionals: An experimental study. *BMC Public Health*, *12*, 525. doi:10.1186/1471-2458-12-525

Gilbert-Diamond, D., Li, Z., Adachi-Mejia, A. M., McClure, A. C., & Sargent, J. D. (2014). Association of a television in the bedroom with increased adiposity gain in a nationally representative sample of children and adolescents. *JAMA Pediatrics*, *168*(5), 427–434. doi:10.1001/jamapediatrics.2013.3921

Haslam, D. W., & James, W. P. (2005). Obesity. *Lancet*, *366*(9492), 1197–1209. doi:10.1016/S0140-6736(05)67483-1

Hauner, H., Hebebrand, J., Müller, M. J., Ried, J., Sorensen, T. I. A., & Stumvoll, M. (2012). *Roadmap for tackling obesity in HORIZON 2020. German competence network obesity.* Retrieved from http://www.ngfn.de/upload/mediapool/Strategy_Paper_Obesity_HORIZON_2020_final2.pdf.

HBSC-Studienverbund Deutschland. (2015). *Studie health behaviour in school-aged children—Faktenblatt "Körpergewicht von Kindern und Jugendlichen* Retrieved from http://www.gbe-bund.de/gbe10/owards.prc_show_pdf?p_id=22744&p_sprache=d&p_uid=gast&p_aid=31059359&p_lfd_nr=1.

Hors-Fraile, S., Atique, S., Mayer, M. A., Denecke, K., Merolli, M., & Househ, M. (2016). The unintended consequences of social media in healthcare: New problems and new solutions. *Yearbook of Medical Informatics*, (1), 47–52. doi:10.15265/iy-2016-009

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, *15*, 635–650.

Iyengar, S. (1991). *Is anyone responsible? How television frames political issues*. Chicago, IL: University of Chicago Press.

Johnson, J. D., & Case, D. O. (2012). *Health information seeking*. New York, NY: Peter Lang.

Kahn, S. E., Hull, R. L., & Utzschneider, K. M. (2006). Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature*, *444*(7121), 840–846. doi:10.1038/nature05482

Krotz, F. (2001). *Die Mediatisierung des kommunikativen Handelns. Der Wandel von Alltag und sozialen Beziehungen, Kultur und Gesellschaft durch Medien*. Wiesbaden: Westdeutscher Verlag.

Latner, J. D., Rosewall, J. K., & Simmonds, M. B. (2007). Childhood obesity stigma: Association with television, videogame, and magazine exposure. *Body Image*, *4*(2), 147–155. doi:10.1016/j.bodyim.2007.03.002

Leventhal, H. (1970). Findings and theory in the study of fear communications. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 5, pp. 119–186). New York, NY: Academic Press.

Liang, T., Kuhle, S., & Veugelers, P. J. (2009). Nutrition and body weights of Canadian children watching television and eating while watching television. *Public Health Nutrition*, *12*(12), 2457–2463. doi:10.1017/S1368980009005564

Livingstone, S., & Helsper, E. (2004). *Advertising foods to children: Understanding promotion in the context of children's daily lives. A review of the literature prepared for the Research Department of the Office of Communications (OFCOM)*. London: OFCOM.

Luhmann, N. (2000). *The reality of the mass media*. Cambridge: Polity Press.

Lundby, K. (Ed.). (2009). *Mediatization: Concepts, changes, consequences*. New York, NY: Peter Lang.

Mack, I., Bayer, C., Schäffeler, N., Reiband, N., Brölz, E., Zurstiege, G., … Zipfel, S. (2017). Chances and limitations of video games in the fight against childhood obesity—A systematic review. *European Eating Disorders Review*, *25*(4), 237–267. doi:10.1002/erv.2514

Maher, J., Fraser, S., & Lindsay, J. (2010). Between provisioning and consuming? Children, mothers and 'childhood obesity'. *Health Sociology Review*, *19*(3), 304–316. doi:10.5172/hesr.2010.19.3.304

Manz, K., Schlack, R., Poethko-Muller, C., Mensink, G., Finger, J., Lampert, T., & KiGGS Study Group. (2014). [Physical activity and electronic media use in children and adolescents: Results of the KiGGS study: First follow-up (KiGGS wave 1)]. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz, 57(7), 840–848. doi:10.1007/s00103-014-1986-4

Medienpädagogischer Forschungsverbund Südwest. (2014). *KIM-Studie 2014. Kinder + Medien, Computer + Internet. Basisuntersuchung zum Medienumgang 6-bis 13-Jähriger*. Stuttgart: Medienpädagogischer Forschungsverbund Südwest.

Medienpädagogischer Forschungsverbund Südwest. (2015). *JIM-Studie 2015. Jugend, Information, (Multi-) Media. Basisuntersuchung zum Medienumgang 12-bis 19-Jähriger*. Stuttgart: Medienpädagogischer Forschungsverbund Südwest.

Meitz, T. G. K., Ort, A., Kalch, A., Zipfel, S., & Zurstiege, G. (2016). Source does matter: Contextual effects on online media-embedded health campaigns against childhood obesity. *Computers in Human Behavior*, *60*, 565–574. doi:10.1016/j.chb.2016.02.067

Murer, S. B., Saarsalu, S., Zimmermann, J., & Herter-Aeberli, I. (2016). Risk factors for overweight and obesity in Swiss primary school children: Results from a representative national survey. *European Journal of Nutrition*, *55*(2), 621–629. doi:10.1007/s00394-015-0882-5

Neysmith, S. M., & Reitsma-Street, M. (2005). "Provisioning": Conceptualizing the work of women for 21st century social policy. *Women's Studies International Forum*, *28*(5), 381–391. doi:10.1016/j.wsif.2005.06.001

Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2014). Prevalence of childhood and adult obesity in the United States, 2011–2012. *JAMA*, *311*(8), 806–814. doi:10.1001/jama.2014.732

Popova, L. (2012). The extended parallel process model: Illuminating the gaps in research. *Health Education & Behavior*, *39*(4), 455–473. doi:10.1177/1090198111418108

Porter, L., & Golan, G. J. (2006). From subservient chickens to Brawny Men: A comparison of viral advertising to television advertising. *Journal of Interactive Advertising*, *6*(2), 30–38.

Reilly, J. J., Methven, E., McDowell, Z. C., Hacking, B., Alexander, D., Stewart, L., & Kelnar, C. J. (2003). Health consequences of obesity. *Archives of Disease in Childhood*, *88*(9), 748–752.

Robert Koch-Institut. (2015). *Mediennutzung. Faktenblatt zu KiGGS Welle 1: Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland – Erste Folgebefragung 2009–2012*. Berlin: Robert Koch Institut.

Rogers, R. W. (1975). A protection motivation theory of fear appeals and attitude change. *Journal of Psychology*, *91*(1), 93. doi:10.1080/00223980.1975.9915803

Rogers, R. W. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In J. T. Cacioppo & R. E. Petty (Eds.), *Social psychophysiology* (pp. 153–177). New York, NY: Guilford Press.

Rosenstock, I. M. (1960). What research in motivation suggests for public health. *American Journal of Public Health and the Nations Health*, *50*(3 Pt 1), 295–302. doi:10.2105/AJPH.50.3_Pt_1.295

Rouhani, M. H., Haghighatdoost, F., Surkan, P. J., & Azadbakht, L. (2016). Associations between dietary energy density and obesity: A systematic review and meta-analysis of observational studies. *Nutrition*, *32*(10), 1037–1047. doi:10.1016/j.nut.2016.03.017

Ruiter, R. A. C., Kessels, L. T. E., Peters, G.-J. Y., & Kok, G. (2014). Sixty years of fear appeal research: Current state of the evidence. *International Journal of Psychology*, *49*(2), 63–70. doi:10.1002/ijop.12042

Ruiter, R. A. C., Verplanken, B., Kok, G., & Werrij, M. Q. (2003). The role of coping appraisal in reactions to fear appeals: Do we need threat information? *Journal of Health Psychology*, *8*(4), 465–474. doi:10.1177/13591053030084006

Sassenberg, K., Fetterman, A., Krebs, M., & Neugebauer, J. (2014). *Skala zur Erfassung von Emotionen gegenüber dem eigenen Gewicht bei Kindern [Scale for capturing childrens emotions towards their own weight]*. Tübingen, DE.

Schemer, C., Wirth, W., & Matthes, J. (2012). Value resonance and value framing effects on voting intentions in direct-democratic campaigns. *American Behavioral Scientist*, *56*(3), 334–352. doi:10.1177/0002764211426329

Serdula, M. K. (1993). Do obese children become obese adults? A review of the literature. *Preventive Medicine*, *22*, 167–177.

Serrano, E. L. (2004). The evaluation of food pyramid games, a bilingual computer nutrition education program for Latino youth. *Journal of Family and Consumer Sciences Education*, *22*(1).

Sigman, A. (2012). The impact of screen media on children: A Eurovision for parliament. In C. C. B. Heys, M. Matthes, & P. Sullivan (Eds.), *Improving the quality of childhood in Europe 2012* (Vol. 3, pp. 88–121). Kidbrooke Park: European Parliament Working Group on the Quality of Childhood in the European Union.

Silverstone, R. (1999). *Why study the media?* Thousand Oaks, CA: Sage.

So, J. (2013). A further extension of the extended parallel process model (E-EPPM): Implications of cognitive appraisal theory of emotion and dispositional coping style. *Health Communication*, *28*(1), 72–83. doi:10.1080/10410236.2012.708633

Staiano, A. E., & Calvert, S. L. (2011). Exergames for physical education courses: Physical, social, and cognitive benefits. *Child Development Perspectives*, *5*(2), 93–98. doi:10.1111/j.1750-8606.2011.00162.x

Thiel, A., Alizadeh, M., Giel, K., & Zipfel, S. (2008). Stereotypisierung von adipösen Kindern und Jugendlichen durch ihre Altersgenossen. *Psychotherapie, Psychosomatik. Medizinische Psychologie*, *58*(12), e16–e24. https://doi.org/10.1055/s-2008-1067340

Thiel, A., Giel, K., Thedinga, H., & Zipfel, S. (2016). Körperlichkeit als Devianz. Zur sozialen Konstruktion des übergewichtigen Körpers und ihrer Folgen. *Zeitschrift für Kulturwissenschaften*.

Thomas, S. L., Olds, T., Pettigrew, S., Yeatman, H., Hyde, J., & Dragovic, C. (2014). Parent and child interactions with two contrasting anti-obesity advertising campaigns: A qualitative analysis. *BMC Public Health*, *14*(151). doi:10.1186/1471-2458-14-151

van Stralen, M. M., te Velde, S. J., van Nassau, F., Brug, J., Grammatikaki, E., Maes, L., … ToyBox-study, g. (2012). Weight status of European preschool children and associations with family demographics and energy balance-related behaviours: A pooled analysis of six European studies. Obesity Reviews 1, 13 Suppl, 29–41. doi:10.1111/j.1467-789X.2011.00959.x

Vraga, E. K., Edgerly, S., Wang, B. M., & Shah, D. V. (2011). Who taught me that? Repurposed news, blog structure, and source identification. *Journal of Communication*, *61*(5), 795–U746. doi:10.1111/j.1460-2466.2011.01581.x

Wabitsch, M., & Kunze, D. (2015). *Konsensbasierte (S2) Leitlinie zur Diagnostik, Therapie und Prävention von Übergewicht und Adipositas im Kindes- und Jugendalter. Version 15.10.2015*.

Walsh, A. D., Lioret, S., Cameron, A. J., Hesketh, K. D., McNaughton, S. A., Crawford, D., & Campbell, K. J. (2014). The effect of an early childhood obesity intervention on father's obesity risk behaviors: The Melbourne InFANT Program. *International Journal of Behavioral Nutrition and Physical Activity*, *11*, 18. doi:10.1186/1479-5868-11-18

Witte, K. (1992). Putting the fear back into fear appeals: The extended parallel process model. *Communication Monographs*, *59*(4), 329–349. doi:10.1080/03637759209376276

Witte, K., & Allen, M. (2000). A meta-analysis of fear appeals: Implications for effective public health campaigns. *Health Education & Behavior*, *27*(5), 591–615.

Zipfel, S. (2015). *Kids obesity prevention program—Study (KOP)*. Retrieved from https://clinicaltrials.gov/ct2/show/NCT02551978?term=kids+obesity&rank=2

Zurstiege, G., Meitz, T. G. K., & Ort, A. (2014). «So ashamed»—Die kommunikative Rekontextualisie- rung einer provokanten Kampagne gegen Adipositas bei Kindern und Jugendlichen ["So ashamed." Re-contextualizing a provocative anti-childchood obesity campaign]. In C. Schwender, D. Schlütz, & G. Zurstiege (Eds.), *Werbung im sozialen Wandel [Advertising in the course of social change]* (pp. 206–223). Köln: Halem.

Chapter 6
Using Digital Media to Assess and Promote School and Adult Education Teacher Competence

Thamar Voss, Annika Goeze, Christian Marx, Verena Hoehne, Viola Klotz, and Josef Schrader

Introduction

National and international large-scale studies such as PISA raise fundamental questions about the effectiveness, efficiency, and social equality of national educational systems and lifelong learning (e.g., Organisation for Economic Cooperation and Development, 2016). In this context, a major finding of educational research is that students' performance varies substantially depending upon the different teachers in the classrooms (e.g., Rivkin, Hanushek, & Kain, 2005). As a consequence, teachers have more and more become a focus of empirical educational research (Cochran-Smith & Zeichner, 2005). On the one hand, they are the central players

T. Voss (✉)
Eberhard Karls Universität Tübingen, Hector-Institut für Empirische Bildungsforschung, Tübingen, Germany

Universität Freiburg, Institut für Erziehungswissenschaft, Freiburg, Germany
e-mail: thamar.voss@ezw.uni-freiburg.de

A. Goeze • C. Marx
Deutsches Institut für Erwachsenenbildung-Leibniz-Zentrum für Lebenslanges Lernen e.V, Bonn, Germany

V. Hoehne
Eberhard Karls Universität Tübingen, Hector-Institut für Empirische Bildungsforschung, Tübingen, Germany

V. Klotz
Universität Mannheim, Fakultät für Betriebswirtschaftslehre, Mannheim, Germany

J. Schrader
Deutsches Institut für Erwachsenenbildung-Leibniz-Zentrum für Lebenslanges Lernen e.V, Bonn, Germany

Eberhard Karls Universität Tübingen, Institut für Erziehungswissenschaft, Tübingen, Germany

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_6

in the educational system who create and structure learning environments for students. On the other hand, teachers are learners themselves who must adapt their abilities and develop their competence based on their experiences in the classrooms. Classrooms in all educational sectors are informational environments that are highly complex in nature. Teachers face multidimensional situations in which events happen simultaneously, often taking unpredictable turns (Doyle, 2006). They have to react immediately, and their actions are witnessed by a group of students who will—at least in the school sector—meet continuously for several weeks, months, or years (Doyle, 2006). These intrinsic properties of classroom environments shape the task of teaching, putting constant pressure on the teacher. Teachers in all educational contexts need to consider and make use of many information sources in the classroom in order to adapt to these environments and grow. Hence, effective designs in teacher education research have to take into account this high complexity (Seidel & Thiel, 2017). Video technologies make it possible to approximate the complex and contextualized nature of teaching, not only for assessing but also for promoting teacher competence. Approaches to promoting teacher competence in teacher education and further education have been shown to be especially effective when using authentic and lively materials, which take into account the variety and complexity of teaching (e.g., Janík et al., 2009; Seago, 2004). Written case narratives reduce the complexity of classroom situations and direct the attention of those being tested directly to given events. Video case studies, in contrast, have the potential to enhance the validity of test instruments in assessing teacher competence because of their more inclusively complex presentation. Video technology, thus, has become a powerful tool for creating effective designs to assess teacher competence and promote teachers' professional development (e.g., Brophy, 2004; Goeze, 2016; Janík & Seidel, 2009).

In the chapter presented here, we describe a research program that makes use of the potential of digital media in the assessment and promotion of teacher competence in two different educational sectors, school and adult education. Subsequent to an overview of video-based digital media used for the assessment and promotion of teacher competence, we describe (a) the design of a web-based learning and testing environment with video case studies, (b) empirical research results on the assessment of teacher competence (i.e., one crucial aspect of teacher competence, namely teachers' pedagogical/psychological knowledge, PPK), and (c) empirical research results on the promotion of teacher competence (i.e., one crucial aspect of teacher competence, namely analytical competence) based on the web-based environment.

The Potential of Digital Media to Assess Teacher Competence in Different Educational Contexts

The research interest in digital media tools to assess teacher competence is growing, mainly inspired by new technological developments. Today's continuous technological improvements provide great potential for the construction of innovative measurement instruments for the assessment of teacher competence (e.g., see Janík & Seidel, 2009).

The main advantages of computer-based assessments over conventional paper-pencil testing are (Jurecka, 2008): (a) *Economic benefits* due to time-saving and cost-saving test administration with automatic scoring techniques, plus easy access to large samples made possible through testing that is independent of time and location, (b) measurement of *additional parameters* such as reaction time, test time, or time per item, and (c) higher *psychometric quality* through enhanced objectivity (standardized test administration, instruction, and test scoring), improved reliability (reduced scoring and measurement error), and enhanced validity. The use of digital media, especially digital videos, can strengthen the ecological validity, an advantage particularly relevant for teacher testing. Furthermore, innovative item formats make it possible to assess procedural knowledge that might be directly relevant for teaching or even the actual application of knowledge in authentic teaching situations (Darling-Hammond, 2006; Jahn, Prenzel, Stürmer, & Seidel, 2011). In these respects, the use of classroom videos has become a promising tool to assess teacher competence (Brophy, 2004; Seidel et al., 2009).

In one approach, for instance, researchers developed a classroom video analysis (CVA) assessment of teaching fractions with standardized video clips as item prompts (e.g., Kersting, 2008; Kersting, Givvin, Thompson, Santagata, & Stigler, 2012). Participating teachers watched 13 short (3–5 min) video clips that covered relevant subtopics within the larger domain of fractions. Afterwards, participants had to analyze the teaching episodes on the basis of standardized questions. The following aspects of the teachers' written responses were given a score: mathematical content, student thinking, suggestions for improvement, and depth of interpretation. Based on expert-novice research (Carter, Cushing, Sabers, Stein, & Berliner, 1988; Putnam & Borko, 2000), it was assumed that a teacher's ability to reason about classroom videos indicated differentiated and integrated professional knowledge of the teaching process (Goodwin, 1994). For instance, Kersting and colleagues (Kersting, Givvin, Sotelo, & Stigler, 2010) found that teachers' video assessment was strongly related to their pedagogical content knowledge.

Another prominent example of a classroom video tool is the *Observer*. The Observer is a standardized instrument to measure teachers' *professional vision* (e.g., Seidel, Blomberg, & Stürmer, 2010; Seidel & Stürmer, 2014), defined as the ability to notice and interpret relevant features of classroom events for student learning (van Es & Sherin, 2002). Similar to the approach of Kersting (2008), participants watch short video clips featuring three components of effective teaching and learning (goal clarity, teacher support, and learning climate). The clips are followed by standardized ratings that capture two components of participants' professional vision from their responses to the video clips (see Stürmer, Könings, & Seidel, 2013; van Es & Sherin, 2002): First, the ability to pay attention to events that are of importance for teaching and learning in classrooms (noticing), and second, the ability to take a reasoned approach to the events noticed in the classroom (knowledge-based reasoning). The authors compared participants' responses with expert ratings. The results from this research supported the assumption that knowledge-based reasoning is an aspect of professional vision (Seidel & Stürmer, 2014) and that there would be a gain in professional vision over one semester through exposure to the video clips (Stürmer et al., 2013; see e.g., Hamre et al., 2012 for another video-based test for early childhood teachers).

Taken together, digital media, and videos in particular, are a promising tool to measure aspects of teacher competence taking into account the high complexity of classroom situations and the contextualized nature of teacher competence. However, it remains unclear what kinds of psychological constructs were measured with the video-based instruments. Are the measures indicators of knowledge (and what kinds of knowledge?), skills, and abilities, or the application of knowledge in classroom examples? Hence, further research is needed on the nature of the constructs that were assessed as well as their validity.

Furthermore, the studies described above and approaches to the assessment of specific aspects of teacher competence were conducted in school classes in a given educational context (e.g., early childhood teaching solely, see Hamre et al., 2012, or mathematics teaching solely, see Kersting, 2008). In our research program, we take a comparative approach with teachers from different educational contexts. We focus not only on school teachers, but also on adult-education teachers and trainers, a group that represents a large, often underestimated portion of the teaching staff in educational systems. The few studies on the competence of adult education teachers are mainly qualitative field studies (e.g., Hof, 2001; Pratt & Nesbit, 2000). These studies focus on teachers' subjective theories as well as their pedagogical beliefs in relation to their biographical backgrounds (e.g., Bastian, 1997; Hartig, 2008; Hof, 2001; Pratt & Associates, 1998; Richardson, 1994). Consequently, one starting point of our research program was to develop a standardized digital measurement to assess teacher competence across educational sectors, to break new ground in research.

The Potential of Digital Media to Promote Teacher Competence in Different Educational Contexts

A second research field of growing interest is the use of digital media to foster the competence of teachers. Teachers' abilities to "think like a teacher" (Shulman, 1992, p. 1) is considered crucial in many empirical studies (e.g., Sherin, Jacobs, & Philipp, 2011). "Thinking like a teacher" relates to teachers' cognitive skills in grasping the complexity of pedagogical situations, especially from the students' point of view (Shulman, 1992). In particular, the ability to observe, analyze, and "diagnose" classroom situations by becoming immersed in student and teacher perspectives as well as by applying conceptual knowledge to better understand classroom situations ("analytical competence") can be seen as an essential aspect of teacher expertise and professionalism (Boshuizen, Bromme, & Gruber, 2004; Goeze, Zottmann, Vogel, Fischer, & Schrader, 2014; Hogan, Rabinowitz, & Craven, 2003; Sherin et al., 2011). Presumably, this analytical competence is a necessary (albeit not in itself sufficient) precondition for the promotion of teacher competence resulting in professional classroom practice.

In the context of initial teacher education and teachers' professional development as in-service teachers, video technologies have been identified as having great potential since the early 1960s. Today, digital video technology allows for the visualization of dynamic processes, approximating a fuller representation of complex reality to learners. Video usage (e.g., for the design of digital video case studies) is regarded as especially beneficial when entering ill-structured domains such as teacher education (Goldman, Pea, Barron, & Derry, 2007). In recent years, a growing body of empirical studies on teacher training shows that in particular pre-service teachers' cognitive skills can be enhanced using video case studies (see Gaudin & Chaliés, 2015, and Goeze, 2016, for overviews). Video examples can support meaningful learning from practice through illustrating classroom practices and providing opportunities to observe and analyze examples of authentic classroom situations (Brophy, 2004; Kramer & Reusser, 2005). Furthermore, videos can be used to guide or coach teachers' teaching (Janík et al., 2009). The use of digital media allows not only for a more authentic presentation of cases, it also enables learners to work with them in diverse ways, e.g., by annotating (Fu, Schaefer, Marchionini, & Mu, 2006; Mu, 2010) or editing videos (Pea, Lindgren, & Rosen, 2008; Zahn, Krauskopf, Hesse, & Pea, 2010).

Despite these advantages, however, several authors have emphasized the importance of providing adequate instructional support for learners (Brophy, 2004; Kirschner, Sweller, & Clark, 2006), since the usage of video case studies alone does not ensure learning (Moreno & Valdez, 2007). Video case-based learning is considered (and partially empirically proven) to have great potential for the promotion of teachers' analytical competence in initial teacher education as well as for their ongoing professional development. However, investigations of the effects of this kind of instructional support have been rare. In addition, "the empirical literature supporting the effectiveness of using classroom video in pre-service teacher education (especially in terms of controlled experimental design studies) is relatively thin" (Blomberg, Renkl, Sherin, Borko, & Seidel, 2013, p. 94).

Development of a Web-Based Learning and Test Environment

In policy documents as well as in research, the quality of teachers' and trainers' performance is regarded as a key factor for the quality of education in different educational contexts. Therefore, the EU, for example, has made efforts to create standards for the qualifications and competence not only of school teachers but also of the adult education personnel (Commission of the European Communities, 2007).

In German-speaking countries, there are also attempts to establish frameworks for adult educators' qualification requirements (Lattke & Jütte, 2014). These initiatives have been especially useful in the field of university didactics and further education, since—in contrast to the school sector—becoming a teacher in these sectors does not depend on any specific standard, qualification, or degree in many European countries. In fact, especially the initial education of people deciding to become adult educators is extremely heterogeneous (e.g., Deutsches Institut für Erwachsenenbildung, 2016), and their previous training does not necessarily include any pedagogical aspects. Moreover, for many freelance adult educators, investing time and money to participate in train-the-trainer-seminars at certain (distant) locations increasingly comes into conflict with the constraints to make ends meet or to make a living with a teaching job. Under these circumstances, it is understandably difficult to educate and professionalize teachers, especially in the adult education sector.

Considering these circumstances, new answers for the assessment of teacher competence and for teachers' professional development have to be found that are flexible in time and location, particularly for the adult education sector. In order to meet these challenges, a web-based environment was designed that can be integrated into formal and non-formal lifelong learning contexts (Digel, Goeze, & Schrader, 2012; Schrader, Hohmann, & Hartz, 2010) to create effective designs to assess and promote teacher competence. The web-based environment makes it possible to integrate split screen videos as well as hyperlinks and different response formats. In the environment closed item formats (e.g., multiple-choice-items where participants have to choose the correct answer) as well as open formats (e.g., short-answer items where participants have to write down short answers to open questions) can be implemented for the assessment of teacher competence. Furthermore, it allows for the integration of more innovative item formats, such as drag-and-drop items (e.g., to sort statements) or video-based items to tap procedural aspects of teacher competence (see Figs. 6.1 and 6.2 for a sample screenshot of a design to *assess* teacher competence).

In order to create a design that is effective in promoting teacher competence, the web-based environment was designed with the possibility of self-directedly exploring complex fields of knowledge. Hypertexts can be integrated to provide participants with additional information (see Fig. 6.3 for a sample screenshot), and flexible work—alone and in (online-) groups—with the videos is possible. For example, not only can participants pause or fast-forward, etc., but they can also directly annotate the case videos and share these annotations. In these ways the tool greatly expands the potential for assessing and promoting teacher competence with video case studies.

By designing this web-based environment we have tried to realize the idea behind *use-inspired basic research* (Fischer, Waibel, & Wecker, 2005; Stokes, 1997). Doing use-inspired basic research means on the one hand focusing on the practical relevance by looking at the field of practice and its professionals and determining

The following video clip shows the beginning of the first session of a seminar.

The situation can be regarded as problematic in terms of communication and interaction.

What could be the two main reasons for this?

*Please mark the two answers. You are free to watch the video **multiple** times.*

(A) The teacher's instruction at the very beginning.	☐
(B) The teacher's behavior at the end of the round of introductions.	☐
(C) The teacher's static position at the front of the room.	☐
(D) The sequence of the course: overview first, round of introductions second.	☐

Fig. 6.1 Video-based item for the dimension *knowledge about communication and interaction*

a starting point for research from their practical needs and unsolved questions. On the other hand, it means simultaneously reframing these questions in state-of-the-art research designs—thus contributing to scientific progress beyond direct practical usage.

As a consequence, we designed the learning and testing environment to also be in line with the *participatory design* approach (Könings, Brand-Gruwel, & van Merriënboer, 2005, 2007). Participatory design “stresses collaboration between stakeholders, in which stakeholders work on better understanding of each other’s perspectives and improving instructional design by close collaboration” (Könings, Seidel, & van Merriënboer, 2014, p. 3). This collaboration was designed into almost every step of the development of the learning environment resulting in high acceptance rates among users (Digel, Herbrechter, & Schmitt, 2013).

In the next paragraph, we describe how this web-based environment with video case studies has been used to assess a crucial aspect of teacher competence (namely teachers’ pedagogical/psychological knowledge) as well as to promote teacher’s analytical competence.

A positive self-perception and high learning motivation have proven to be beneficial for the learning processes.

Which of the following instruments and methods would you implement to measure the students' self-perception and motivation?

Please check one box in each line.	Rather incorrect	Rather correct
(A) Sociograms (mapping social networks, e.g. students' seating arrangement or observation of conversation contacts)	☐	☐
(B) Written examinations	☐	☐
(C) Self-reports	☐	☐
(D) Group work	☐	☐

Imagine covering a topic which the students are diversely familiar with. You want to adjust the learning objectives to the students' prior knowledge.

Please answer with short key words.

(1) Which learning objective would you phrase for students unfamiliar with the topic?

----

(2) Which learning objective would you phrase for students highly familiar with the topic?

Fig. 6.2 Text-based items of the dimensions *knowledge about assessment methods and individual diagnostics* and *knowledge about learning goals and their transfer/implementation*

Development of an Instrument to Measure Teachers' Pedagogical/Psychological Knowledge with the Web-Based Learning and Testing Environment

In the following sections, we describe the development of a measurement instrument embedded within the web-based learning and testing environment to tap the PPK of teachers in different educational contexts. First, we introduce a detailed conceptualization of PPK based on a literature research and its validation through a written expert investigation. We then describe the development of the digital items designed to tap teachers' PPK and the steps in the piloting process. Lastly, we present results of a study with 212 teachers in two educational sectors (namely school and adult education).

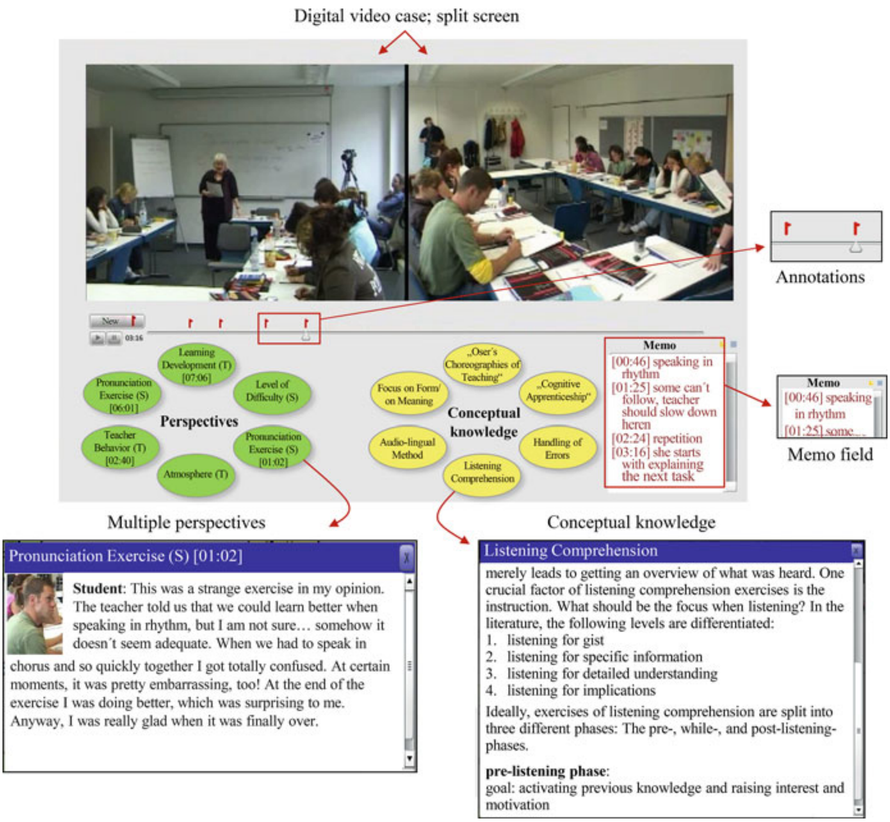

Fig. 6.3 Screenshot of the training program to promote teacher's analytical competence

Conceptualization of Pedagogical/Psychological Knowledge (PPK)

The development of the instrument was based on the COACTIV-model on teachers' professional competence ("Professional Competence of Teachers, **Co**gnitively **Activ**ating Instruction, and the Development of Students' Mathematical Literacy"; see Kunter, Baumert, et al., Kunter et al., 2013), which has become very influential in research on teacher competence and in teacher education (especially in Germany) over the last few years. The model combines concepts of competence (Klieme, Hartig, & Rauch, 2008; Weinert, 2001) with taxonomies of Shulman (1986, 1987) on teacher knowledge, with approaches from motivational psychology. At the most general conceptual level, the model distinguishes among four aspects of teachers' professional competence: professional knowledge, beliefs, motivational characteristics, and self-regulation (Baumert & Kunter, 2013). Teachers' professional knowledge is further differentiated into three domains (Bromme, 1992; Shulman, 1986, 1987): *content knowledge*, *pedagogical content knowledge* (e.g., Baumert et al., 2010; Krauss et al., 2008), and *pedagogical/psychological knowledge* (PPK).

PPK is defined as generic cross-curricular knowledge needed to create and optimize teaching and learning situations across domains and educational sectors (Voss, Kunina-Habenicht, Hoehne, & Kunter, 2015; Voss, Kunter, & Baumert, 2011). We made use of digital media based on an existing instrument that measures secondary school teachers' PPK from the COACTIV project (see Voss et al., 2011) to measure teachers' procedural knowledge with digital technologies, such as videos integrated in the web-based learning and testing environment described above. After a review of the literature (more than 9000 titles), we conceptualized teachers' PPK and its constituting dimensions. To validate and find blind spots in the conceptualization, a paper-pencil expert survey was conducted with practitioners, experts, and members of the scientific community from the school sector and from the field of adult education ($N = 44$). The experts were asked to rate the identified dimensions of PPK regarding generality (i.e., generality across subjects and educational sectors) and relevance and were also asked to comment freely on the conceptualization. The data were analyzed qualitatively as well as quantitatively. Results showed that the experts perceived the proposed dimensions as highly generic ($M = 3.36$, scale: 1–4) and relevant for teaching ($M = 3.47$). The results of the expert survey gave us for adaptation into the design the following eight dimensions: PPK is defined as the knowledge needed to create and optimize teaching-learning situations, including declarative and procedural generic knowledge of effective teaching that is potentially applicable in different subjects and educational sectors (Voss et al., 2015). It encompasses knowledge about (1) learning processes, (2) methods and concepts of learning and instruction, (3) learning goals and their transfer/implementation, (4) assessment methods and individual diagnostics, (5) classroom/courseroom management, (6) heterogeneity of learners and how to deal with it, (7) how to interact and communicate with learners, and (8) how to design learning environments (Marx, Goeze, & Schrader, 2014a, 2014b; Marx et al., 2017; Voss et al., 2015).

Item Construction

For all of these eight dimensions, we constructed items and embedded them in the web-based learning and testing environment. The developed items have different formats: Some were dichotomous (multiple-choice-items with one correct answer and multiple distractors), some polytomous (true-false-items with multiple statements), and short-answer items (participants had to write down short answers to open questions). For some of the polytomous and short-answer items participants watched short videos and rated statements related to the videos (video-based items). Some of the videos were compiled by case studies in the program to promote teacher competence described below. Furthermore, we reenacted classroom situations with real secondary school classes, which had also been successfully done for other studies (Voss et al., 2011). Overall, eight videos were reenacted covering critical classroom situations related to efficient classroom management. The situations were based on classroom management literature. Classroom management (in adult education: courseroom management) is defined as the efficient use of instructional time to enable meaningful learning and promote students' growth (Emmer & Sabornie, 2015). Core strategies that have been empirically shown to be efficient are related to the prevention, correction, and redirection of inappropriate behavior with the goal of promoting students' cognitive, social, and emotional development (e.g. Emmer & Evertson, 2013; Marzano, 2005). Knowledge of classroom management is thus an important aspect of the knowledge base of teachers in different educational sectors. Research on classroom management often refers to the seminal work of Kounin (1970), who identified a set of preventive and proactive teacher behaviors associated with the absence of misbehavior. For instance, effective classroom managers are always aware of what is going on in all parts of the classroom (with-it-ness, monitoring), they are able to effectively handle two or more classroom events at the same time (overlapping), and are able to activate all students (group alerting and accountability). The construction of our items to assess knowledge of classroom and courseroom management is closely guided by Kounin's work. The videotaped classroom situations present critical classroom situations in terms of preventive/proactive classroom management. In the PPK-test, we presented the short videos to the participating teachers and administered questions after each sequence. In most cases, the first question related to with-it-ness by addressing monitoring competence. In a second step, we assessed strategies for preventing or dealing with misbehavior (e.g., overlapping strategies). We additionally constructed video items measuring aspects of teachers' knowledge on students' heterogeneity and teaching methods.

The items were pre-tested in three pilot studies with school teachers and teachers from adult education. In the first pilot study, we used think-aloud-protocols ($N = 20$ teachers) in order to examine whether the instructions, items, and technical usability of the current test version were clear. Based on the second pilot study with $N = 332$ student teachers, we conducted classical item analysis to get preliminary information about the psychometric quality of the items. The third pilot study was conducted to test the newly developed video-based items which measured teachers' classroom management knowledge of $N = 32$ school teachers (Luzius, 2016).[1]

After adapting step-by-step the items based on the results of the pilot studies, the resulting 67 items were implemented in a teacher study which will be presented in the following section. Example items are depicted in Figs. 6.1 and 6.2.

[1]The items on classroom management were constructed later than most of the other items. Therefore, an additional pilot study was carried out.

Main Study with Teachers from Adult Education and School

Data Base and Analytic Approach

The sample consisted of 212 teachers, 66 school teachers and 146 teachers from adult education. The teachers from adult education were on average 48.70 years old (SD = 12.39, range = 23–83 years), with on average of 14.36 years teaching experience (SD = 11.17, range = 0–45), and 68 % female. Teachers from the school context were 25–63 years old (*M* = 45.41, SD = 10.60), with 1.5–37 years of teaching experience (*M* = 16.42, SD = 10.23). The majority taught at an academic track school (Gymnasium) and 59 % of them were female.

We analyzed the data based on a partial credit model using the Conquest software (Wu, Adams, Wilson, & Haldane, 2007).

Results of the Measurement of Overall PPK

Analyses showed that PPK was satisfactorily described by a unidimensional partial credit model: $0.87 \leq$ wMNSQ ≤ 1.17 (*weighted mean square*, values close to 1 indicate a good fit to the data, values >1.25 indicated a bad fit and should be excluded); $-3.557 \leq$ item difficulty $\leq +2.145$ and an EAP/PV reliability of .78 (*Expected A Posteriori Estimation* reliability, see Wu et al., 2007). A short version of the test instrument to measure overall PPK more economically resulted in a 35-item instrument with a homogeneity of .84 (Mayer, 2015). Hence, we were able to measure overall PPK of teachers with satisfactory reliability.

However, based on our conceptualization, we had assumed a multidimensional structure of PPK with eight dimensions. The rather low (in relation to the number of items) reliability resulting in the model with all 67 items can be interpreted as a first indication that we were measuring different aspects of PPK with our items. Because of the limited sample size of this study, we were not able to systematically test the multi-dimensional structure based on the comprehensive conceptualization with all eight dimensions. However, to get a first insight in PPK's dimensionality, we conducted multi-dimensional analyses for two of the eight dimensions as examples, which will be described in the following section.

Results of Multidimensional Analyses with Two Dimensions as Examples

Based on the two dimensions *learning goals and their transfer/implementation* and *assessment methods and individual diagnostics* (14 items), we compared two models: a one-dimensional model and a two-dimensional model (Marx et al., 2017). The comparison of both models based on an index of likelihood and number of parameters (Wu et al., 2007) revealed that the one-dimensional model had a significantly poorer fit than the two-dimensional model. The descriptive comparison

Table 6.1 Fit statistics of the models

	Deviance	Parameter	AIC	BIC	EAP/PV reliability
Model 1: 1-dim.	5456.0	39	5534.0	5562.0	.70
Model 2: 2-dim.	5440.7	41	5522.7	5552.1	.70 and .60
Difference	*15.3**	*2*	*11.3*	*9.9*	

$^{*}p < .05$

of the Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) supported this assumption, with small differences in favor of the two-dimensional model (see Table 6.1).

The weighted fit indices of the items (WMNSQ) in the two-dimensional model ranged from .89 to 1.07. Hence, all items reached satisfactory WMNSQ-values. The difficulties of the items ranged from −1.372 to 2.000, indicating good values in relation to the person's abilities. Both dimensions were highly correlated with $r = .89$.

Running both models separately in the two subsamples (teachers from school and adult education) supported the assumption that the two-dimensional model displayed the data more accurately than the one-dimensional model: In both subsamples, the fit of the two-dimensional model exceeded the fit of the one-dimensional model. Interestingly, the reliabilities of the two dimensions were markedly higher in the subsample of the school teachers (.81 and .70) than in the subsample of the adult educators (.58 and .44). The intercorrelations of both dimensions did not differ substantially between the subsamples ($r = .88$ school teachers, .82 adult education teachers).

A crucial indicator in comparing items across subgroups is differential item functioning (DIF). DIF effects of items indicate to what extent items might be measuring different abilities for members of different groups. Hence, items display DIF if participants from different groups with the same underlying ability have a different probability of giving a correct answer. We implemented the common routine for assessing DIF with the Conquest 2.0 software (Wu et al., 2007). We found medium-sized DIF effects for four items and a small effect (Santelices & Wilson, 2012) for another item. Hence, the majority of the items show no DIF effect. The ConQuest implementation routine to check for DIF also includes the comparison of the mean levels (Wu et al., 2007). For both dimensions, teachers from the school context statistically significantly outperform teachers from adult education (*assessment methods and individual diagnostics*: .124 logits (SE = 0.023), $p \leq .001$; *learning goals and their transfer/implementation*: .464 logits (SE = 0.059), $p \leq .001$).

To sum up, we were able to measure these two chosen aspects of PPK in both educational sectors reliably. In line with our expectations, on average school teachers outperformed adult education teachers, who have often not been prepared in any professional teacher education program.

Development of a Training Program to Promote Teacher Competence Based on the Web-Based Learning and Testing Environment

Theoretical Background and Implementation of the Training Program

Cognitive Flexibility Theory (CFT) can be drawn upon as a basis for instructional support that seeks to promote flexible knowledge application in different real situations, improve awareness of one's own perspective, and allow for the construction of connections to alternative perspectives (Spiro, Collins, Thota, & Feltovich, 2003; Spiro & Jehng, 1990). In general, CFT seeks to (1) explain instructional support that helps people to learn important yet difficult subject matter, (2) inform the design of environments that foster adaptively flexible use of knowledge in real-world settings, (3) provide guidelines that support modification of underlying ways of thinking, and (4) guide the development of technology-enhanced non-linear learning environments for the promotion of complex learning and flexible knowledge application. This approach is based on the idea that "knowledge that has to be used in many ways has to be represented in any ways" (Spiro et al., 2003, p. 6). Having had experience with many situations and cases is assumed to be advantageous in order to prepare people to apply knowledge across a variety of real-world situations. Providing instructional support that comprises multiple perspectives may enhance self-directed flexible changes of perspective. The ability to shift from one's own perspective to that of the different actors in a teaching case (i.e., teachers and learners) is crucial for the analytical competency of teachers (van Es & Sherin, 2002). CFT further recommends the use of computer-supported learning environments to allow for a non-linear, multi-dimensional presentation of contents, which should be organized in the form of hypermedia to support the principle of the so-called "criss-crossed landscape." This landscape refers to learners' self-directed exploration of a complex field of knowledge by exploring varying views and contexts by means of the hypermedia and finding connections to other concepts, in other words, by "criss-crossing" the knowledge field. Further benefits associated with the use of hypermedia include the possibility of quickly revisiting scenes or perspectives (for this summary on CFT see also Zottmann et al., 2012, p. 515).

The training program that was developed allowed for this "criss-crossing," taking on new perspectives and applying knowledge in a self-directed way (Olleck, 2010). A "criss-crossed landscape" was realized in the user interface of the web-based environment through its 10- to 15-min split screen videos and its integrated hypertexts. By clicking on "multiple perspectives" buttons, the learner (here: a pre-service or in-service-teacher) can "read the mind" of a student or teacher who participated in the authentic video episodes and were later asked in an interview to comment on them. When clicking on “conceptual knowledge” buttons the learner could briefly inform oneself about relevant models or theories chosen by experts as a background for better understanding the situation depicted in the video. Thus, the web-based training program enables the learners to self-directedly access a complex field of knowledge not by linear, step-by-step instructions, but by criss-crossing the knowledge landscape (see Fig. 6.3, in which the circular layout of the hyperlink-buttons avoids any priority or order).

The training program allowed the participants to work with the video case studies flexibly via pause, fast-forward, and rewind functions, as well as giving them the option to annotate the case videos with markers (flags placed above the time bar; see Fig. 6.3). With each annotation, a time stamp appeared automatically in the memo field, where related remarks could be filled in directly. If a participant clicked on a hyperlink button, a pop-up window appeared that displayed the selected item of conceptual knowledge or one of the multiple perspectives; simultaneously, the video froze and shaded to avoid split attention effects (Chandler & Sweller, 1992).

Research Results on the Promotion of Teacher Competence with the Developed Training Program

In this section, we present studies on the effectiveness of the 4-day training program. Firstly, we present the results of a study using experimentally varied instructional support to promote pre-service teachers’ competence in analyzing and diagnosing pedagogical situations during this 4-day training. Secondly, we describe the results of studies conducted using the same training concept with adult educators, and thirdly, we present results on the use of digital media to promote student teachers’ professional development *in real practice*.

Results of Using Digital Media to Promote Pre-service Teachers’ Analytical Competence

A pre-post-follow-up-intervention study was conducted using seven video case studies from adult education courses (three for pre, post, and follow-up tests, four as treatment). The experiment was done to investigate the effects of two types of instructional support used during the 4-day teacher training program: Hyperlinks to conceptual knowledge and hyperlinks to multiple perspectives (see Fig. 6.3 above) were varied in a 2×2 factorial design by not providing versus providing them. We examined whether these types of instructional support could (together) help to counteract some of the known deficits of video case-based learning and teacher thinking, such as limited change of perspective and knowledge usage. The study was carried out as a 4-day university training course for pre-service teachers

($N = 100$). As outcomes, we analyzed collaborative learning processes (small-group discussions) and as data, we analyzed written case narratives. The results indicated that both types of instructional support were beneficial. In particular, hyperlinks to multiple perspectives affected small-group case discussions and written post-tests, as they led to increased changes in perspective. Hyperlinks to conceptual knowledge furthered the application of this knowledge, especially in the written post-tests (Goeze et al., 2014). Whether the effects on written case analyses would be sustained over time was also examined, in our study after 3 months, without further training in the learning environment. This was the case, suggesting a kind of "heureka-effect" as the basis for the development of competence; once something new was discovered, no significant fading was discernible (Goeze, 2016).

Moreover, in linear regression analyses we investigated whether and how strongly personal characteristics (individual learning conditions, e.g., prior knowledge, tolerance of ambiguity) and process characteristics (cognitive load; cognitive, emotional, and motivational activation) had an impact on the development of teachers' analytic competence. The results showed that this type of video case-based learning placed little demand on personal factors and that the quality of the learning process played a more decisive role for the promotion of teachers' analytical competence (Goeze, Hetfleisch, & Schrader, 2013).

Results of Using Digital Media to Promote Adult Educators' Analytic Competence

In the described studies, the training yielded empirical evidence in the field of (pre-service) teacher education in the school sector. In further studies, we implemented the training program in different educational contexts, such as teacher education in university didactics and in adult education. By replicating similar research designs in other educational contexts, we tested the reproducibility of the empirical findings regarding teacher competence development from the school sector. Subsequently, a field study examining (a) the effects and (b) the acceptance of the concept of video-case-based learning in a broad variety of everyday further education courses was conducted. The field study further served to test the usability of the newly developed learning environment among organizational (i.e., adult education providers) as well as individual users (i.e., teachers or trainers and their course participants). In addition, what kinds of pedagogical, professional, and technical support multiplicators required were investigated (Digel & Goeze, 2010). Study results confirmed (a) the positive effects on the development of trainers' analytic competence as well as (b) high acceptance rates of the concept of video-case-based learning and its correspondent learning environment among adult education providers as well as diverse users in different educational contexts (Digel & Schrader, 2013; Digel et al., 2013).

Results of Using Digital Media to Promote Teacher Candidates' Analytic Competence in Real Practice

Since evidence-based training interventions (such as the above-mentioned video case- based training program) that can improve student teachers' professional development *in real practice* are rare, it is especially interesting to find out how these interventions can be implemented into everyday teacher education practice *by teacher educators themselves*. It is a challenge to be able to use these interventions without losing their initial, lab-tested effectiveness as a result of teacher educators' exercising their own practical educational autonomy in their own courses. On the basis of the results mentioned above, a study was conducted in the second phase of the German teacher training system, i.e., the teachers' training period after their university master's degree. In this quasi-experimental field study, we investigated what effects different degrees of educational autonomy (low/middle/high) had on the development of competence to diagnose classroom situations. Nineteen teacher educators who were granted different degrees of autonomy implemented the experimentally proven training into their own teacher training courses with 261 teacher candidates. ANCOVAs using planned contrasts indicated that the effectiveness of the training could be sustained in "the real world" of teacher education practice, even when practitioners—no longer researchers—were responsible for the realization of the training program. The result was that the highest increase in teacher candidates' analytic competence occurred when low educational autonomy was granted, i.e., when the training was implemented most closely to the original (Hetfleisch, Goeze, & Schrader, 2014).

Establishing Infrastructure for the Promotion of Teachers' and Trainers' Professional Development

The empirical findings outlined above resulted not only in the further explication of the video case-based learning environment itself, but also inspired the composition of a more comprehensive and wide-spreading learning platform that includes video upload and didactical materials, as well as chats and forums, among other functions (Digel et al., 2012; www.videofallarbeit.de). This learning platform, in turn, is being integrated within an even larger portal "wb-web", which has been established by the German Institute for Adult Education–Leibniz-Center for Lifelong Learning and the Bertelsmann Foundation (www.wb-web.de). It serves as an exchange forum and learning platform for the more than 700,000 teachers and trainers working in the adult education sector in Germany.

A survey of demands with more than 1000 trainers revealed great interest in (a) freely accessible didactical materials, in (b) networking activities across individuals, education providers, professional communities and associations, as well as in (c) the comprehensive certification of trainers' individual competences (Digel, 2014). Against this backdrop, www.wb-web.de uses social media and provides a broad and growing variety of informational and educational materials as open educational resources (OER) for teachers and trainers (especially in the adult education sector).

Conclusion and Outlook

In this chapter, we presented results of a research program that makes use of the potential of digital media to assess and promote teacher competence in different educational sectors. A web-based learning and testing environment was designed. In several studies, we investigated whether (and how) this web-based environment can be implemented to assess and to promote aspects of teacher competence. The web-based learning and testing environment has two main advantages: (a) It allows developing and implementing approaches that are flexible in time and location to assess and promote teacher competence in different educational sectors. (b) It allows for the integration of video-based and interactive elements in the approaches to assess and promote teacher competence. In the introduction, we elaborated on the need for approaches that take into account the complexity of teaching-learning situations and the contextualized nature of teacher competence.

The main results of the research program were:

(a) There was wide acceptance of the web-based learning environment with video case studies among pre-service school teachers and teachers from adult education.
(b) We made use of the advantages of the web-based learning and testing environment with integrated classroom videos in the assessment of teacher competence. We successfully developed a test instrument to measure the pedagogical/psychological knowledge (PPK) of teachers in two educational sectors (school and adult education). With this instrument, we were able to measure aspects of teachers' PPK reliably on a general level in both educational sectors. Furthermore, research results indicated the multidimensional structure of the PPK measured with our instrument. In an additional study, we implemented the instrument in a sample with a total of almost 300 school and adult education teachers. With these data, we aim to replicate our finding that our instrument is suited to measure teachers' PPK in an economical and reliable manner in different educational sectors. Furthermore, we aim to test further the multidimensional structure of PPK and its relations to other aspects of teacher competence. In a supplemental study, we investigated instructional quality from the student perspective to get first indications of the criterial validity of the instrument.
(c) Results of several intervention studies indicated that the web-based learning and testing environment that integrated video case studies is a powerful tool to promote taking on new perspectives and applying knowledge among school

> and adult education teachers. Two types of instructional support turned out to be effective: hyperlinks to multiple perspectives and hyperlinks to conceptual knowledge. The effects were sustained over time as indicated in a follow-up study 3 months after the intervention.

Several implications for practice and research resulted from the research project. The web-based learning and testing environment is helpful and beneficial for research on teacher competence: Economic benefits resulted from cost-saving administration and automatic scoring of participants' responses. Furthermore, and probably more importantly, with this tool researchers have easier access to large samples because of administration that is independent of time and location. Teachers often work to capacity and are highly stressed (Johnson et al., 2005; OECD, 2005), so that, in general, teachers' willingness to additionally support empirical research is rather low (especially among experienced in-service teachers). With a time and location independent of administration the expenses are lower and the access to teachers is more likely.

Furthermore, the instrument to measure teachers' PPK is (to the best of our knowledge) the first instrument that—from the outset—takes the assumed generality of PPK into account. We first conceptualized PPK as generic, i.e., as important for teaching in different subjects and different educational sectors. Secondly, we constructed items for the measurement of the PPK of teachers in different subjects and educational sectors. Lastly, we empirically tested for DIF-effects. The results supported our claim and indicated that the items were appropriate to measure PPK in the two educational sectors that we chose to compare. Hence, with our instrument, comparative research in the many different educational sectors and subjects is possible. The development of the instrument prepares the ground for research in educational sectors—such as adult or vocational education—where little empirical evidence on teacher competence is available so far. Such instruments are crucial to investigate systematically the impact of teacher competence on instructional quality and, finally, on students' learning success. Therefore, the instrument will contribute to results and ideas for the enhancement of teaching quality in different educational sectors.

With regard to practical implications, teachers can use the instrument to regulate their own diagnosis of their competence. They can identify personal strengths and weaknesses, and based on these self-diagnoses, they can, for instance, choose courses for professional development.

In the long run, based on the measurement instrument and the web-based learning environment we constructed, it will be possible to develop training programs not only for teacher candidates, but also for in-service teachers and trainers, to enhance their PPK levels.

Furthermore, based on the web-based learning and testing environment that was developed, an even larger web-portal has been generated ("wb-web"). This exchange forum and learning environment will be expanded into a platform that will not only serve as a community building platform but will also allow for assessment and certification processes of teachers' and trainers' knowledge and competence.

To conclude, the web-based learning and testing environment is a powerful tool for implementing effective designs in order to assess and promote teacher competence. It can effectively be integrated into formal and informal learning contexts in different educational settings.

References

Bastian, H. (1997). *Kursleiterprofile und Angebotsqualität*. Bad Heilbrunn: Klinkhardt.

Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. R. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers. Results from the COACTIV project* (pp. 49–62). New York, NY: Springer.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., … Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133–180. doi:10.3102/0002831209345157

Blomberg, G., Renkl, A., Sherin, M. G., Borko, H., & Seidel, T. (2013). Five research-based heuristics for using video in pre-service teacher education. *Journal for Educational Research Online*, *5*(1), 90–114.

Boshuizen, H. P. A., Bromme, R., & Gruber, H. (2004). On the long way from novice to expert and how travelling changes the traveller. In H. P. A. Boshuizen, R. Bromme, & H. Gruber (Eds.), *Professional learning: Gaps and transitions on the way from novice to expert* (pp. 3–8). Dordrecht: Kluwer Academic Publishers.

Bromme, R. (1992). *Der Lehrer als Experte: Zur Psychologie des professionellen Wissens*. Bern: Huber.

Brophy, J. (2004). *Using video in teacher education* (Vol. 10). West Yorkshire: Emerald Group Publishing Limited.

Carter, K., Cushing, K., Sabers, D., Stein, P., & Berliner, D. (1988). Expert-Novice differences in perceiving and processing visual classroom information. *Journal of Teacher Education*, *39*(3), 25–31. doi:10.1177/002248718803900306

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*(2), 233–246.

Cochran-Smith, M., & Zeichner, K. M. (2005). *Studying teacher education: The report of the AERA panel on research and teacher education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Commission of the European Communities. (2007). *Action plan on adult learning*. Brussels. Retrieved from http://ec.europa.eu/education/policies/adult/com558_en.pdf.

Darling-Hammond, L. (2006). Assessing teacher education: The usefulness of multiple measures for assessing program outcomes. *Journal of Teacher Education*, *57*(2), 120–138.

Deutsches Institut für Erwachsenenbildung, Bundesinstitut für Berufsbildung, & Universität Duisburg-Essen. (2016). *Personal in der Weiterbildung: Beschäftigungsverhältnisse und Tätigkeiten—wb-personalmonitor*. Bielefeld: W. Bertelsmann.

Digel, S. (2014). Bilden und Forschen mit Videofällen—Stand und Erfahrungen der Projektgruppe Videofallarbeit der Universität Tübingen. In E. Feyerer, K. Hirschenhauser & K. Soukup-Altrichter (Hrsg.), *Last oder Lust? Forschung und Lehrer_innenbildung* (pp. 159–166). Münster: Waxmann.

Digel, S. & Goeze, A. (2010). Aufgaben und Qualifizierung von ModeratorInnen fallbasierter Fortbildungen. In J. Schrader, R. Hohmann, & S. Hartz (Hrsg.), *Mediengestützte Fallarbeit—Konzepte, Erfahrungen und Befunde zur Kompetenzentwicklung von Erwachsenenbildnern* (pp. S.147–S.166). Bielefeld: W. Bertelsmann.

Digel, S., Herbrechter, D., & Schmitt, T. (2013). Förderung professioneller Kompetenz Lehrender. Studierende und erfahrene Lehrkräfte im Vergleich. In H. V. Felden, C. Hof, & S. Schmidt-Lauff (Hrsg.), *Erwachsenenbildung im Spannungsfeld von Wissenschaft, Politik und Praxis* (pp. S.62–S.76). Baltmannsweiler: Schneider Hohengehren.

Digel, S., Goeze, A., & Schrader, J. (2012). *Aus Videofällen lernen: Einführung in die Praxis für Lehrkräfte, Trainer und Berater*. Bielefeld: Bertelsmann.

Digel, S., & Schrader, J. (Eds.). (2013). *Diagnostizieren und Handeln von Lehrkräften: Lernen aus Videofällen in Hochschule und Schule*. Bielefeld: W. Bertelsmann.

Doyle, W. (2006). Ecological approaches to classroom management. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 97–125). Mahwah, NJ: Lawrence Erlbaum Associates.

Emmer, E. T., & Evertson, C. M. (2013). *Classroom management for middle and high school teachers*. Boston, MA: Pearson.

Emmer, E. T., & Sabornie, E. J. (2015). *Handbook of classroom management* (2nd ed.). New York, NY: Routledge.

Fischer, F., Waibel, M., & Wecker, C. (2005). Nutzenorientierte Grundlagenforschung im Bildungsbereich: Argumente einer internationalen Diskussion. *Zeitschrift für Erziehungswissenschaft*, *8*(3), 427–442.

Fu, X., Schaefer, J. C., Marchionini, G., & Mu, X. (2006). Video annotation in a learning environment. In *Proceedings of the annual meeting of the American Society for Information Science & Technology*, *Vol. 1* (pp. 1–22).

Gaudin, C., & Chaliés, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review*, *16*, 41–67.

Goeze, A. (2016). *Professionalitätsentwicklung von Lehrkräften durch videofallbasiertes Lernen—Voraussetzungen, Prozesse, Wirkungen*. Bielefeld: W. Bertelsmann.

Goeze, A., Hetfleisch, P., & Schrader, J. (2013). Wirkungen des Lernens mit Videofällen bei Lehrkräften: Welche Rolle spielen instruktionale Unterstützung, Personen- und Prozessmerkmale? *Zeitschrift für Erziehungswissenschaft*, *16*(1), 79–113.

Goeze, A., Zottmann, J., Vogel, F., Fischer, F., & Schrader, J. (2014). Getting immersed in teacher and student perspectives? Facilitating analytical competence using video cases in teacher education. *Instructional Science*, *42*(1), 91–114. doi:10.1007/s11251-013-9304-3

Goldman, R., Pea, R. D., Barron, B., & Derry, S. (Eds.). (2007). *Video research in the learning sciences*. Mahwah, NJ: Erlbaum.

Goodwin, C. (1994). Professional vision. *American Anthropologist*, *96*(3), 606–633.

Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T., … Scott-Little, C. (2012). A course on effective teacher-child interactions: Effects on teacher beliefs, knowledge, and observed practice. *American Educational Research Journal*, *49*(1), 88–123. doi:10.3102/0002831211434596

Hartig, C. (2008). *Berufskulturelle Selbstreflexion. Selbstbeschreibungslogiken von ErwachsenenbildnerInnen*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Hetfleisch, P., Goeze, A., & Schrader, J. (2014). Implementation eines wissenschaftlich erprobten, didaktischen Konzepts: Der Einfluss pädagogischer Autonomie auf die Wirksamkeit in der Praxis. *Zeitschrift für Erziehungswissenschaft*, *17*(2), 297–322.

Hof, C. (2001). *Konzepte des Wissens. Eine empirische Studie zu den wissenstheoretischen Grundlagen des Unterrichtens*. Bielefeld: W. Bertelsmann Verlag.

Hogan, T. M., Rabinowitz, M., & Craven, J. A. (2003). Representation in teaching: Inferences from research of expert and novice teachers. *Educational Psychologist*, *38*, 235–247.

Jahn, G., Prenzel, M., Stürmer, K., & Seidel, T. (2011). Varianten einer computergestützten Erhebung von Lehrerkompetenzen: Untersuchungen zu Anwendungen des Tools Observer. *Unterrichtswissenschaft*, *39*(2), 136–153.

Janík, T., & Seidel, T. (Eds.). (2009). *The power of video studies in investigating teaching and learning in the classroom*. Münster: Waxmann.

Janík, T., Janíková, M., Knecht, P., Kubiatko, M., Najvar, P., Najvarová, V., & Šebestová, S. (2009). Exploring different ways in using video in teacher education: Examples from CPV Video Web. In T. Janík & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 207–224). Münster: Waxmann.

Johnson, S., Cooper, C., Cartwright, S., Donald, I., Taylor, P., & Millet, C. (2005). The experience of work-related stress across occupations. *Journal of Managerial Psychology*, *20*, 178–187. doi:10.1108/ 02683940510579803

Jurecka, A. (2008). Introduction to the computer-based assessment of competencies. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 193–213). Toronto: Hogrefe.

Kersting, N. B. (2008). Using video clips of mathematics classroom instruction as item prompts to measure teachers' knowledge of teaching mathematics. *Educational and Psychological Measurement*, *68*(5), 845–861. doi:10.1177/0013164407313369

Kersting, N. B., Givvin, K., Sotelo, F., & Stigler, J. W. (2010). Teacher's analysis of classroom video predicts student learning of mathematics: Further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education*, *61*(1), 172–181.

Kersting, N. B., Givvin, K. B., Thompson, B. J., Santagata, R., & Stigler, J. W. (2012). Measuring usable knowledge: Teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, *49*(3), 568–589. doi:10.3102/0002831212437853

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86.

Klieme, E., Hartig, J., & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of comptencies in educational contexts* (pp. 3–22). Cambridge: Hogrefe.

Könings, K. D., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2005). Towards more powerful learning environments through combining the perspectives of designers, teachers and students. *British Journal of Educational Psychology*, *75*(4), 645–660.

Könings, K. D., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2007). Teachers' perspective on innovations: Implications for educational design. *Teaching and Teacher Education*, *23*(6), 985–997.

Könings, K., Seidel, T., & van Merriënboer, J. (2014). Participatory design of learning environments: Integrating perspectives of students, teachers, and designers. *Instructional Science*, *42*(1), 1–9.

Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York: Holt, Rinehart & Winston.

Kramer, K., & Reusser, K. (2005). Unterrichtsvideos als Medium der Aus- und Weiterbildung von Lehrpersonen. *Beiträge zur Lehrerbildung*, *23*(1), 35–50.

Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M., & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology*, *100*(3), 716–725. doi:10.1037/0022-0663.100.3.716

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2013). *Cognitive activation in the mathematics classroom and professional competence of teachers. Results from the COACTIV project*. New York: Springer.

Lattke, S. & Jütte, W. (Hrsg.) (2014). *Professionalisation of adult educators: International and comparative perspectives*. Frankfurt a. M.: Peter Lang.

Luzius, K. (2016). Klassenführungskompetenz bei Lehrkräften der Sekundarstufe: Erfassung und Zusammenhang zu Beanspruchungserleben. Unveröffentlichte Zulassungsarbeit, Universität Tübingen.

Marx, C., Goeze, A., & Schrader, J. (2014a). Adult education teachers' pedagogical- psychological knowledge. Potential elements and test development. In S. Lattke & W. Jütte (Eds.), *Professionalisation of adult educators. International and comparative perspectives* (pp. 165–182). Frankfurt a.M.: Peter Lang.

Marx, C., Goeze, A., & Schrader, J. (2014b). Pädagogisch-psychologisches Wissen zur Gestaltung von Lehr-Lernsituationen: (Wie) Unterscheidet es sich in Erwachsenenbildung/ Weiterbildung und Schule? *Hessische Blätter für Volksbildung Erwachsenenbildung*, *3*, 238–251.

Marx, C., Goeze, A., Voss, T., Hoehne, V., Klotz, V. K., & Schrader, J. (2017). Pädagogisch-psychologisches Wissen von Lehrkräften aus Schule und Erwachsenenbildung: Entwicklung und Erprobung eines Testinstruments. *Zeitschrift für Erziehungswissenschaft*, *20*(Suppl 1), 165–200. doi:10.1007/s11618-017-0733-7

Marzano, R. J. (2005). *A handbook for classroom management that works*. Alexandria, Va.: Association for Supervision and Curriculum Development.

Mayer, J. (2015). Entwicklung eines Testinstruments zur Erfassung pädagogisch-psychologischen Wissens bei Lehrkräften unterschiedlicher Bildungsbereiche—Ergebnisse einer Pilotstudie, Unveröffentlichte Masterarbeit. Universität Tübingen.

Moreno, R., & Valdez, A. (2007). Immediate and delayed effects of using a classroom case exemplar in teacher education: The role of presentation format. *Journal of Educational Psychology*, *99*(1), 194–206.

Mu, X. (2010). Towards effective video annotation: An approach to automatically link notes with video content. *Computers & Education*, *55*, 1752–1763.

Olleck, R. (2010). Mediengestützte Fallarbeit in computerunterstützten Lernumgebungen: Technische Anforderungen und Funktionalitäten für Einzelarbeit, Gruppenarbeit und Blended-Learning-Szenarien. In J. Schrader, R. Hohmann, & S. Hartz (Eds.), *Mediengestützte Fallarbeit: Konzepte, Erfahrungen und Befunde zur Kompetenzentwicklung von Erwachsenenbildnern* (pp. 191–207). Bielefeld: W. Bertelsmann.

Organisation for Economic Cooperation and Development. (2005). *Attracting, developing and retaining effective teachers*. Paris: OECD.

Organisation for Economic Cooperation and Development. (2016). *PISA 2015 Ergebnisse (Band I): Exzellenz und Chancengerechtigkeit in der Bildung*. Bielefeld: W. Bertelsmann Verlag. doi:10.1787/9789264267879-de

Pea, R., Lindgren, R., & Rosen, J. (2008). Cognitive technologies for establishing, sharing and comparing perspectives on video over computer networks. *Social Science Information*, *47*(3), 353–370.

Pratt, D. D., & Associates. (1998). *Five perspectives on teaching in adult and higher education*. Malabar: Krieger Publishing.

Pratt, D. D., & Nesbit, T. (2000). Discourses and cultures of teaching. In E. R. Hayes & A. L. Wilson (Eds.), *Handbook of adult and continuing education* (pp. 117–132). San Francisco, CA: Jossey-Bass.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning. *Educational Researcher*, *29*(1), 4–15. doi:10.3102/0013189X029001004

Richardson, V. (1994). The consideration of teachers' beliefs. In V. Richardson (Ed.), *Teacher change and the staff development process: A case in reading instruction* (pp. 90–108). New York, NY: Teachers College Press.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458.

Santelices, M. V., & Wilson, M. (2012). On the relationship between differential item functioning and item difficulty. An issue of methods? Item response theory approach to differential item functioning. *Educational and Psychological Measurement*, *72*(1), 5–36.

Schrader, J., Hohmann, R., & Hartz, S. (Hrsg.). (2010). *Mediengestützte Fallarbeit: Konzepte, Erfahrungen und Befunde zur Kompetenzentwicklung von Erwachsenenbildnern*. Bielefeld: W. Bertelsmann.

Seago, N. (2004). Using video as an object of inquiry mathematics teaching and learning. In J. Brophy (Ed.), *Using video in teacher education* (pp. 259–285). Oxford: Elsevier.

Seidel, T., & Stürmer, K. (2014). Modeling the structure of professional vision in pre-service teachers. *American Educational Research Journal*, *51*(4), 739–771. doi:10.3102/0002831214531321

Seidel, T., Blomberg, G., & Stürmer, K. (2010). “Observer”—Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. *Zeitschrift für Pädagogik*, *56. Beiheft*, 296–306.

Seidel, T., Prenzel, M., Schwindt, K., Stürmer, K., Blomberg, G., & Kobarg, M. (2009). LUV and observe: Two projects using video to diagnose teachers’ competence. In T. Janík & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 243–258). Münster: Waxmann.

Seidel, T., & Thiel, F. (2017). Standards und Trends der videobasierten Lehr-Lernforschung. *Zeitschrift für Erziehungswissenschaft*, *20*(1). doi:10.1007/s11618-017-0726-6

Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (Eds.). (2011). *Mathematics teacher noticing: Seeing through teachers’ eyes*. New York, NY: Routledge.

Shulman, J. H. (Ed.). (1992). *Case methods in teacher education*. New York, NY: Teachers College Press.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14. doi:10.3102/0013189X015002004

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–22.

Spiro, R. J., & Jehng, J. (1990). Cognitive flexibility and hypertext: Theory and technology for the non-linear and multidimensional traversal of complex subject matter. In D. Nix & R. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology* (pp. 163–204). Hillsdale, NJ: Erlbaum.

Spiro, R. J., Collins, B. P., Thota, J. J., & Feltovich, P. J. (2003). Cognitive flexibility theory: Hypermedia for complex learning, adaptive knowledge application, and experience acceleration. *Educational Technology*, *43*(5), 5–10.

Stokes, D. (1997). *Pasteur’s quadrant—Basic science and technological innovation*. Washington, DC: Brookings Institution Press.

Stürmer, K., Könings, K. D., & Seidel, T. (2013). Declarative knowledge and professional vision in teacher education: Effect of courses in teaching and learning. *British Journal of Educational Psychology*, *83*(3), 467–483. doi:10.1111/j.2044-8279.2012.02075.x

van Es, E., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers’ interpretations of classroom interactions. *Journal of Technology and Teacher*, *10*(4), 571–596.

Voss, T., Kunina-Habenicht, O., Hoehne, V., & Kunter, M. (2015). Stichwort Pädagogisches Wissen von Lehrkräften: Empirische Zugänge und Befunde. *Zeitschrift für Erziehungswissenschaft*, *18*(2), 187–223.

Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates’ general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, *103*(4), 952–969. doi:10.1037/a0025125

Weinert, F. E. (2001). A concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45–65). Seattle, WA: Hogrefe & Huber.

Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. A. (2007). *ACER ConQuest version 2.0. Generalised item response modelling software*. Camberwell: ACER Press.

Zahn, C., Krauskopf, K., Hesse, F. W., & Pea, R. (2010). Digital video tools in the classroom: How to support meaningful collaboration and critical advanced thinking of students? In M. S. Khine & I. M. Saleh (Eds.), *New science of learning: Cognition, computers and collaboration in education* (pp. 503–523). New York, NY: Springer.

Zottmann, J. M., Goeze, A., Frank, C., Zentner, U., Fischer, F., & Schrader, J. (2012). Fostering the analytical competency of pre-service teachers in a computer-supported case-based learning environment: A matter of perspective? *Interactive Learning Environments*, *20*(6), 513–532.

Chapter 7
Behavioral and Neurocognitive Evaluation of a Web-Platform for Game-Based Learning of Orthography and Numeracy

Mojtaba Soltanlou, Stefanie Jung, Stephanie Roesch, Manuel Ninaus, Katharina Brandelik, Jürgen Heller, Torsten Grust, Hans-Christoph Nuerk, and Korbinian Moeller

Introduction

Over the last decade, it has become increasingly evident that learning is not limited to traditional, formal, and institutional contexts such as schools, but is diffusing into informal contexts. As was seen in Chap. 5 of this book, informal game-based interventions might be helpful not only for academic achievement but also for public health issue (see Zurstiege et al., 2017). This is closely associated with the rise of modern digital media, above all the World Wide Web as an informational environment, allowing for accessing information anywhere and at any time. These

M. Soltanlou (✉)
Department of Psychology, University of Tübingen, Tübingen, Germany

Leibniz-Institut für Wissensmedien, Tübingen, Germany

Graduate Training Centre of Neuroscience/IMPRS for Cognitive and Systems Neuroscience, Tübingen, Germany

S. Jung
Department of Psychology, University of Tübingen, Tübingen, Germany

Leibniz-Institut für Wissensmedien, Tübingen, Germany

S. Roesch
Leibniz-Institut für Wissensmedien, Tübingen, Germany

K. Brandelik
Department of Psychology, University of Tübingen, Tübingen, Germany

M. Ninaus
Leibniz-Institut für Wissensmedien, Tübingen, Germany

LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany

J. Heller
Department of Psychology, University of Tübingen, Tübingen, Germany

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_7

developments seem to have the power and relevance to change education because learners can now specifically search for and find information according to their interests, needs, and abilities allowing for a personalization of their informational environment.

In the present chapter, we describe the developmental and behavioral as well as neurocognitive evaluation of a web-based platform for game-based learning of orthography and numeracy—reflecting the idea of providing informal, low-threshold, and adaptive learning opportunities for what have been called key competencies for our knowledge and information societies. To promote such key competencies, diagnostic assessment of individual abilities and needs is essential. Automatized adaptive methods as provided by some web-based learning platforms can provide such an assessment.

First we give a brief introduction into the relevance of orthography and numeracy as well as the idea of using digital games for learning. Thereafter, we describe the web-based learning platform and its evaluation with a focus on neurocognitive effects of training. We conclude with a discussion of the results and their implications for considering neurocognitive correlates of learning.

The Relevance of Literacy and Numeracy

Literacy and numeracy (i.e., basic reading and writing as well as numerical and arithmetical abilities) are key competencies for navigating life in our post-industrial, knowledge-based societies. Consequently, it seems obvious that insufficient literacy and numeracy carry severe disadvantages for the affected individuals (Hanushek & Woessmann, 2010; Parsons & Bynner, 2005).

Empirical data clearly indicate that poor literacy increases an individual's risk for school dropout, low educational achievement, and unemployment (Esser, Wyschkon, & Schmidt, 2002). Moreover, children with poor reading or writing abilities may be aware of their disadvantages and the social relevance of sufficient literacy. They are more inclined to suffer from behavioral and emotional problems (e.g., Daniel et al., 2006). Similarly, insufficient numerical and arithmetical abilities are known to be detrimental to an individual's educational achievement and career, as well as their health prospects (Parsons & Bynner, 2005).

Despite these individual disadvantages associated with poor literacy and numeracy, it also needs to be noted that poor literacy and numeracy lead to immense

T. Grust
Department of Computer Science, University of Tübingen, Tübingen, Germany

H.-C. Nuerk • K. Moeller
Department of Psychology, University of Tübingen, Tübingen, Germany

Leibniz-Institut für Wissensmedien, Tübingen, Germany

LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany

socio-economic costs (e.g., Butterworth, Varma, & Laurillard, 2011). In the case of insufficient literacy, the estimated annual costs range between £45,000 and £53,000 per affected individual, accumulating to a total of £1.73 to £2.05 billion every year for the UK (Gross, 2006). A similar study addressing the costs of poor numeracy estimated those costs to sum up to £2.4 billion per annum, again for the UK (Gross, Hudson, & Price, 2009).

These results clearly indicate that consequences of insufficient literacy and numeracy are not only relevant for the affected individuals but also for society on a larger scale—particularly Western knowledge-based societies (Beddington et al., 2008). Against this background, it is obvious that ensuring successful numeracy and literacy education is of major societal importance. Additionally, it is necessary to develop effective training programs to benefit students with learning difficulties (i.e., dyslexia and/or dyscalculia). The potential returns from investing in such efforts are considerable. For instance, the OECD (Hanushek & Woessmann, 2010) estimated that "an improvement of one-half standard deviation in mathematics and science performance at the individual level implies, by historical experience, an increase in annual growth rates of GDP [gross domestic product] per capita of 0.87 %" (p. 17). Applied to Germany, this reflects a sum of about 25 billion EUR.

However, as learning disorders were found to be largely resistant to conventional teaching methods, there is the need to look for new and innovative intervention approaches. Therefore, investigating new possibilities to foster literacy and numeracy, such as the application of information technology and new digital media, is a worthwhile endeavor for research and practice.

Over the last decades, digital media have become part of children's everyday life, with considerable influence on free-time activities, interpersonal interactions, and peer-group communication. Nevertheless, formal education and teaching of numeracy and literacy is still largely based on traditional approaches, such as textbooks and worksheets. As a medium for learning, digital games or game-based environments offer interesting possibilities to motivate and engage students in learning (Chen, Liao, Cheng, Yeh, & Chan, 2012).

Digital Games for Learning

Playing digital games can be highly engaging and rewarding, and thus games have become a ubiquitous part of our children's daily lives. According to the Interactive Software Federation of Europe (ISFE), 25 % of Europeans, including both adults and children, play digital games at least once a week (ISFE, 2012). Although most people think of digital games as entertainment, there is increasing interest in using digital games to enhance education (for a review see Boyle et al., 2016), combining the compelling aspects of games with instruction.

Recent studies indicate that the use of game-based tasks or the implementation of game elements in conventional cognitive tasks can not only increase users' motivation and engagement, but also improve their performance (e.g., Mekler, Brühlmann, Tuch, & Opwis, 2017; Ninaus et al., 2015; Prins, Dovis, Ponsioen, Ten

Brink, & Van der Oord, 2011; for a review see Lumsden, Edwards, Lawrence, Coyle, & Munafò, 2016). Thus, digital games provide increasingly important strategies for learning, educational interventions, and cognitive training. The compelling nature of games keeps users motivated to play or interact with the learning application (Erhel & Jamet, 2013; Ninaus et al., 2013), and this can be attributed to certain mechanics within the game itself. For instance, Garris, Ahlers, and Driskell (2002) emphasized the importance of immediate feedback, reflection, and active involvement in games. Interestingly, according to the Self-Determination Theory (Deci & Ryan, 2000), the intrinsic appeal of games can be explained by their ability to satisfy basic psychological needs for competency, autonomy, and relatedness which—when experienced—increase users' motivation and engagement (Przybylski, Rigby, & Ryan, 2010).

Unfortunately, though, only a minority of games for learning were developed based on recent research findings, and even fewer are evaluated empirically. In the domains of literacy (e.g., Tintenklex: http://www.legasthenie-software.de/) and numeracy (e.g., Semideus, http://www.flowfactory.fi/semideus/; The Number Race: http://www.lacourseauxnombres.com/; for an overview see also Moeller, Fischer, Nuerk, & Cress, 2015) game-based solutions seem to be particularly promising. Since research on game-based learning in the domains of literacy and numeracy is still in its infancy, only a few studies have examined positive effects on learning and their comparability to conventional learning methods (e.g., Kast, Baschera, Gross, Jäncke, & Meyer, 2011; Ninaus, Kiili, McMullen, & Moeller, 2016; Wilson, Dehaene et al., 2006; Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006). Thus, the current project is aimed at designing, implementing, and evaluating a web-based learning platform hosting various games for learning orthography and numeracy. In the following section, key features of the web-based platform and the embedded learning games are described.

A Web-Platform for Game-Based Orthography and Numeracy Learning

The web-based learning platform for orthography and numeracy (http://lernplattform.iwm-kmrc.de) was developed in an interdisciplinary collaboration between psychologists, linguists, and computer scientists of the University of Tübingen and of the Leibniz-Institut für Wissensmedien in Tübingen. The learning platform hosts several learning games that aim at fostering orthography and numeracy skills in secondary school students, which starts from the fifth grade in Germany. The learning games are embedded in a web browser to facilitate cross-platform playability, and, thus, to enable individual learning independent from formal learning environments—anytime and anywhere.

Features of the Learning Platform

The learning platform contains various features to personalize its use and encourages social interaction (a detailed description of the learning platform can be found in Jung et al., 2015).

On a front page, students are invited to create a profile and to register (Fig. 7.1, "User Login"). Setting up profiles on the learning platform offers the opportunity to monitor an individual's learning progress. Once the profile has been set up, an individual profile page is generated (Fig. 7.1, "Mein Konto"). This profile page allows students to add and to share personal information (e.g., age, gender, preference for school subjects, etc.) without being obliged to disclose personal data. The front page enables access to all hosted learning games (Fig. 7.1, "Spielen") and informs about currently logged in users and new users who recently joined the community (Fig. 7.1, "Logged in users" and "New users").

Moreover, a chat forum is provided that invites chat with other users or contact with psychologists and programmers when necessary (Fig. 7.1, "Forum").

These implemented features are provided by a client-server architecture based on the Google Web Toolkit (GWT). GWT is a development tool that enables implementation of complex browser-based applications. The included Java-to-Javascript compiler allows developers to design rich internet applications solely written in Java. Additionally, GWT provides a large library of widgets and panels and comprises several built-in methods to communicate with a server (e.g., remote procedure calls). A database server, which runs the object-relational database management system PostgreSQL, logs all user-generated data (e.g., textual input, movements, and current game states). These logs are archived and provide crucial information about users' performance and development (Giorgidze, Grust, Schreiber, & Weijers, 2010).

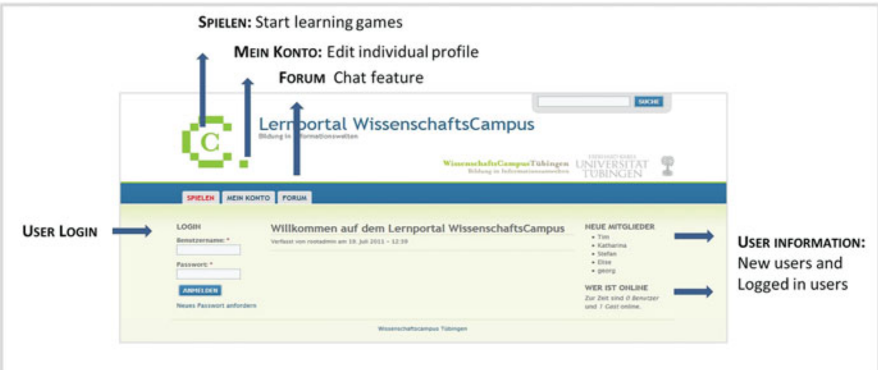

Fig. 7.1 Front page of the learning platform

Learning Games

The learning games rest on the identification of individual difficulties and developmental trajectories for orthography and arithmetic competencies assessed in more than 400 fifth and sixth graders (Huber, Fischer, Moeller, & Nuerk, 2013; Huber, Moeller, & Nuerk, 2012). Moreover, theoretical considerations including recent findings in the relevant domains should ensure optimally effective and individually tailored learning. Altogether, the learning platform hosts four numerical and four spelling games. A special feature of these games is their social interactivity: most of the games are developed as multiplayer games for up to four players and at least one computer-controlled opponent (e.g., number line game and multiplication game, see below for a description). To win a game, both fast and accurate responses are required.

Orthography games are based on in-depth linguistic analyses covering relevant characteristics of the German language, such as phonological (e.g., the perceptual sensitivity of vowel duration) and morphological (e.g., word-formation) aspects. These games are designed to foster an understanding of various German spelling rules.

In German, double consonants (e.g., ff, mm, etc.) usually follow short vowels. As awareness of vowel duration (e.g., offen vs. Ofen) is vital for correct gemination (Ise & Schulte-Körne, 2010; Landerl, 2006), the "Gemination Game" aims at enhancing students' awareness of short and long vowels. The game procedure is threefold: First, vowel duration in an audibly presented word (e.g., Ball) needs to be identified by clicking on a respective symbol (i.e., an exploding ball for "short vowel" or a little man pulling a rope for "long vowel"). Second, in a mini-game, coins depicting double consonants fall from the top of the screen. Students should collect as many coins as possible showing the same double consonant as identified in the target word (e.g., ll). They gain a score point for each collected coin. Third, students are requested to spell the target word correctly by typing it in.

The game "LeTris" also addresses correct spelling. It was designed in the style of the well-known game "Tetris," but instead of geometric figures, falling letters need to be rotated and arranged according to their order in a word, which is presented audibly. When the word is spelled correctly, the line of letters clears away and score points are gained. When the word is spelled incorrectly, the letters remain on the screen. Consequently, each spelling error yields one more line of letters piling up, finally resulting in "game "

The notion of word families (i.e., words that share the same root or morpheme such as read, readable, etc.) is important in language education and associated with literacy performance (Carlisle & Katz, 2006). Word families might be used to derive the spelling of related words. The game "Word Families" fosters this strategy as follows: First, a target word is presented (e.g., child). Second, this target is followed by a range of other words of which some are part of the same word family. Students are requested to remember only those words that belong to the same word family as the target (e.g., children, childish, childhood, etc.). Finally, students are asked to type in and spell correctly all the words they can remember.

The game "Word Building Blocks" also focuses on the approach of word families and was designed as a choice-reaction task. At first, students are requested to select one out of four morphemes (e.g., "old"). Subsequently, they are asked to find and highlight as many words as possible belonging to the same word family, out of various other words (e.g., old timer, oldie, car, spoon, cold, etc.).

Numerical games were designed in line with recent results regarding number processing (e.g., Huber et al., 2013; Link, Huber, Nuerk, & Moeller, 2014; Link, Nuerk, & Moeller, 2014) and cover different numerical principles. The games are predominantly developed as choice-reaction tasks presenting two or more possible solutions to a given problem.

The game "Multiplication" aims at enhancing students' ability to solve multiplication problems. To this end, students are requested to select a multiplication problem (i.e., 3×9) that corresponds to a previously given multiplication result (e.g., 27) out of several response options. The theoretical basis of this game is research indicating that multiplication facts are retrieved from long-term memory once they are learnt by heart (e.g., Delazer et al., 2003; Domahs, Delazer, & Nuerk, 2006).

Addition problems are trained in the game "Partner Number." First, students are asked to pick one out of four given numbers (e.g., 6). Subsequently, they need to select the corresponding number that adds up to 10 (i.e., 4) out of several response options. Only the player who is the first to choose the correct solution gains a score point. The theoretical background of this game is research showing that mastery of the base-10-place-value structure of the Arabic number system is essential for multi-digit number processing (Nuerk, Moeller, Klein, Willmes, & Fischer, 2011; Nuerk, Moeller, & Willmes, 2015 for reviews).

The "Carry Game" extends the training of addition problems. This game specifically trains carry-over operations that can be solved by the following strategy: adding up to the next decade before adding the rest of the addend. The procedure in the "Carry Game" is threefold: First, students choose an addition problem (e.g., $4 + 9$). Second, they are requested to select the number that needs to be added to the first summand (i.e., 4) to add up 10 (solution: 6). Third, they are required to indicate which number remains to be added for the correct solution (i.e., 13; solution: 3). This game is based on research indicating that the carry operation poses a particular difficulty in mental arithmetic in both children and adults (Moeller, Klein, & Nuerk, 2011a, 2011b).

In the "Number Line Game" students are trained in mapping numbers to space (Link, Huber et al., 2014). To pursue this aim, students are required to mark the correct position of a given number (e.g., 43) on a plain number line by means of the mouse cursor. Once a marker has been set, slower opponents cannot place their marker at or around the same location (i.e., 5 % deviation). Thus, players need to answer both quickly and accurately. The theoretical basis of the number line game are findings showing that (a) performance in the number line estimation task is associated reliably with arithmetic performance (e.g., Booth & Siegler, 2006) and (b) that training this task also improves arithmetic performance (Link, Moeller, Huber, Fischer, & Nuerk, 2013; Whyte & Bull, 2008).

Behavioral Evaluation

In a first evaluation of the learning platform, intervention effects of three of the arithmetic ("Multiplication," "Partner Number," and "Carry Game") and three of the spelling games ("Gemination Game," "Word Families," and "Word Building Blocks") were appraised. Two fifth- and sixth-grade classes ($n = 47$, 19 female) of two public secondary schools participated in a study following a crossover design. One class (Group 1, see Fig. 7.2) received three sessions of the spelling training first, followed by three sessions of the arithmetic training (both about 45 min each). This procedure was reversed for the other class (Group 2).

Both arithmetic and spelling performance were assessed at three time points to evaluate learning progress: at the beginning, and then after three and six training sessions (see Fig. 7.2). At these three time points, arithmetic competencies were assessed by a speeded paper-pencil test covering the basic arithmetic operations (addition, subtraction, multiplication, division) with 36 problems each. In the spelling assessment, children had to perform a writing-to-dictation task and completed 28 fill-in-the-blank sentences by inserting a respective target word. Target words covered all central orthographic aspects of the German written language such as capitalization of initial letters, germination, and lengthening signs.

Results indicated that playing the learning games had a beneficial effect on children's arithmetic performance. However, this effect was not specific to training with the arithmetic games. This indicated that some aspects of the orthography training may have improved arithmetic performance as well (i.e., sequential, symbolic, and spatial processing of items). Finally, training effects were only observed for comparably easy problems across all basic arithmetic operations (see Roesch et al., 2016 for a more detailed description of the results). Thus, although arithmetic games successfully trained specific abilities, other training may be beneficial by training more domain-general aspects such as sequential problem processing, sustained attention, or working memory; this needs to be further studied in more detail.

Concerning orthography, training effects could be separated from other temporal effects by an analysis based on a Linear Logistic Test Model for measuring change (Fischer, 1995). The general temporal trend indicated an initial improvement, and a subsequent decrease back to the starting level (see Fig. 7.3a, no intervention). In addition to this trend, the analysis revealed a significant effect of the spelling training, which specifically improved children's writing performance (solid and dashed line in Fig. 7.3a). This model closely predicts the data at the group level. As shown in Fig. 7.3b, however, the training effect was not strong enough to counteract

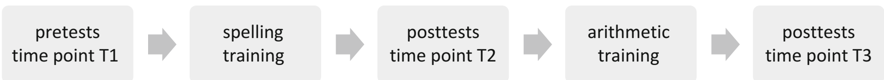

Fig. 7.2 Pretests, post-tests and training schedule, Group 1

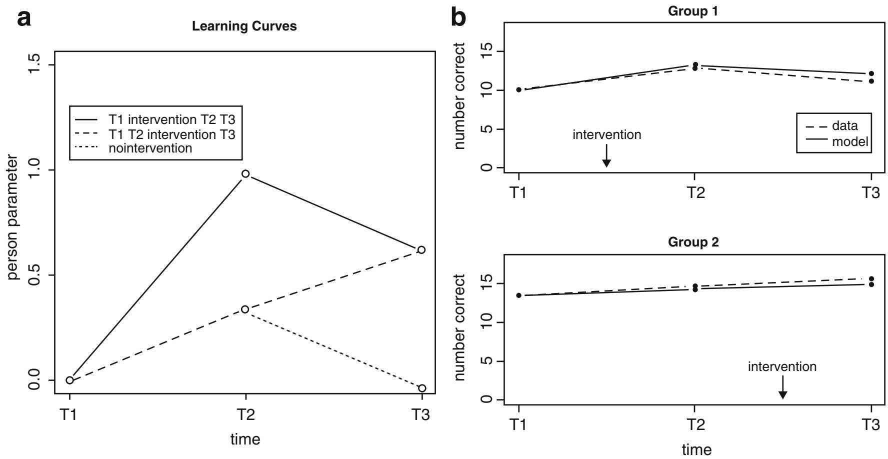

Fig. 7.3 Panel (**a**) provides an illustration of the general temporal trend (*dotted line*) and the intervention effects. Panel (**b**) shows that the observed number of correctly written items is closely predicted by the LLTM results at the group level

the general temporal trend (Group 1). This may be due to a lack of motivation in the post-test at T3, or the limited training time, making long-term effects unlikely.

In sum, the observed learning effects were promising, especially for orthography. The behavioral results demonstrated the general applicability of the developed learning platform and its games, as well as their effectiveness in conveying orthography and arithmetic competencies. One major advantage of digital training on our web-based platform is the possibility to meet individual requirements by introducing adaptivity: implementations similar to those proposed, i.e., implementing adaptivity based on a Brain-Computer Interface (BCI) in an educational context, as it is shown in Chap. 8 of this book (Spüler et al. 2017), may further enhance the already obtained training effects in the future. In the following sections, we discuss neurocognitive data reflecting learning effects, including our initial results based on the online learning platform.

Neurocognitive Evaluation of Arithmetic Learning

Investigating the neurocognitive foundations of arithmetic development and learning is a complementary approach to uncover how children improve their behavioral math skills. For instance, Supekar et al. (2013) observed that while neural correlates predicted math learning in children, behavioral measures failed to do so. Therefore, it is essential to go beyond behavioral investigation of arithmetic development and learning in children in order to come up with appropriate educational and therapeutic interventions for each individual. As is seen in Chap. 8 of this book, Spüler et al. (2017) elaborate on an important application of neurocognitive data over and beyond behavioral assessment for adapting learning environments to individual needs. Neurocognitive studies have already shown that individuals may rely on different brain networks to solve arithmetic problems (Grabner et al., 2007; Grabner, Ansari et al., 2009). Grabner, Ansari et al. (2009) observed higher activation of language-related parietal areas in the left hemisphere in individuals with higher math competency as compared to individuals with lower math competency. Activation of this area is usually interpreted to reflect retrieval strategies (e.g., knowing that 2×3 gives 6 without effortful calculation) in arithmetic problem solving (Zamarian, Ischebeck, & Delazer, 2009), which is one of the dominant strategies used after sufficient training. Delazer et al. (2003) found that different training methods all led to a successful use of retrieval strategies in adults. However, these different training methods led to differing brain activation patterns, whereas behavioral performance improvement was similar.

Neurocognitive Foundation of Arithmetic Learning in Adults

According to Poldrack (2000), learning is a shift from general purpose processes to more task-specific processes. In line with this definition, arithmetic learning is characterized by a strategy shift from more effortful and algorithm-based calculations to more economic retrieval processes, which results in characteristic changes in brain activation patterns (Zamarian et al., 2009). The fronto-parietal network usually found to be active during number processing involves both areas associated with domain-general and domain-specific processing engaged in arithmetic problem solving. In this network, frontal areas are associated with supplementary domain-general cognitive processes such as working memory (WM) and planning in mental calculation, while parietal areas are associated with magnitude processing of numerals and domain-specific processes (for a review see Arsalidou & Taylor, 2011). According to the triple-code model of number processing (Dehaene & Cohen, 1997; Dehaene, Piazza, Pinel, & Cohen, 2003), domain-specific processing of number magnitude information is subserved by neurons in and around the bilateral intra-parietal sulci, while visuo-spatial demands of number processing are associated with the superior parietal lobule, and language-related demands of number processing (e.g., retrieving arithmetic facts from long-term memory) are dedicated to the left angular gyrus.

In adults, Delazer et al. (2003) found that arithmetic learning goes along with a decrease of cognitive load in verbal and visuo-spatial WM, less engagement of attentional control, strategy planning and self-monitoring, and also less application of mathematical rules and algorithms in calculation associated with frontal activation (for a review see Zamarian et al., 2009). Furthermore, it is assumed to induce more specific processing of number magnitude information in parietal and language-related processes in left temporal areas of the brain. Such a shift in solution strategy was repeatedly observed to be accompanied by reduced activation in the fronto-parietal network of number processing and increased activation in language-related parieto-temporal areas in the left hemisphere in adults (Delazer et al., 2003, 2005; Grabner, Ischebeck et al., 2009; Ischebeck et al., 2006; Ischebeck, Zamarian, Egger, Schocke, & Delazer, 2007; Ischebeck, Zamarian, Schocke, & Delazer, 2009; Pauli et al., 1994; for a review see Zamarian et al., 2009). Generally, there seems to be a shift from frontal to parietal regions, and then within parietal regions with increasing proficiency reflecting reduced demands on domain-general cognitive processes and increased domain-specific numerical processes (Zamarian et al., 2009). Interestingly, it was found that this shift in brain activation can already happen after eight repetitions of arithmetic problems in adults (Ischebeck et al., 2007).

However, there is no agreement on the specificity of this shift in brain activation with training. Importantly, the shift seems to depend on the learning method (Delazer et al., 2005), on the arithmetic operation (Ischebeck et al., 2006), and on the experimental design (Bloechle et al., 2016), and may not even be specific to arithmetic learning (Grabner & De Smedt, 2012; Grabner, Ischebeck et al., 2009). Moreover, recent studies suggested a pivotal role of hippocampal systems associated with long-term memory functioning in arithmetic learning in adults (for a review see Klein et al., 2016) as well as in children (e.g., Qin et al., 2014).

Neurocognitive Foundation of Arithmetic Development and Learning in Children

Neuroimaging studies of arithmetic learning in children are still scarce and most of our knowledge is drawn from studies comparing children with adults using cross-sectional designs or involving math tutoring and assessing longitudinal age- and/or training-related effects. There is agreement on an increase of automated processes in arithmetic problem solving with age, which is reflected by enhanced activation of domain-specific parietal areas and reduced activation of complementary domain-general frontal areas (e.g., Kaufmann, Wood, Rubinsten, & Henik, 2011; Kucian, von Aster, Loenneker, Dietrich, & Martin, 2008). This developmental increase of automated processing is reflected by a frontal-to-parietal shift of brain activation (i.e., reduced activation of frontal areas and increased activation of parietal areas (e.g., Ansari, 2008; Rivera, Reiss, Eckert, & Menon, 2005)—very similar to that observed in adults after training.

For instance, Kawashima et al. (2004) found greater activation of right parietal areas during subtraction and multiplication problem solving in adults compared to children, but no difference in frontal areas (see also Kucian et al., 2008). Moreover, they found reduced activation of areas associated with attentional processing in adults as compared to children (Kucian et al., 2008; see also Cantlon et al., 2009). Moreover, Rivera et al. (2005) suggested that adolescents recruit left parietal areas, and the left occipito-temporal area (see also Emerson & Cantlon, 2015), while children still recruit bilateral frontal areas and attention-related areas (see also

Prado, Mutreja, & Booth, 2014). This more pronounced frontal activation indicates that greater demands on WM and executive function in children are required to achieve similar performance to adolescents. Furthermore, stronger activation of the left hippocampus was observed in children, which was interpreted as revealing higher demands on declarative and procedural memory systems (Rivera et al., 2005).

The transitional role of the hippocampal system in arithmetic development and learning was investigated by Menon and colleagues. Supekar et al. (2013) found that pre-training hippocampal volume, and the functional association of the hippocampus with frontal areas subserving domain-general processes, predicted training-related arithmetic improvement in children. Surprisingly, no behavioral measures, including IQ, WM, and general math abilities, did so (see also Evans et al., 2015). Moreover, in a longitudinal study on children, Qin et al. (2014) found a critical transient role of the medial temporal lobe, including the hippocampus, in arithmetic learning. They suggested that the hippocampal system is pivotal in the shift from procedural to retrieval strategies, as shown by the increasing involvement of the hippocampus and the decreased involvement of fronto-parietal networks in arithmetic problem solving with age (Qin et al., 2014).

Furthermore, it was also suggested that the neurocognitive correlates of different arithmetic operations are not necessarily identical during development. In a cross-sectional study in children, Prado et al. (2014) found a grade-related activation increase in left language-related temporal areas for multiplication, but a grade-related increase of right quantity- and magnitude-related parietal activation for subtraction. Prado et al. (2014) concluded that fluency in arithmetic problem solving is achieved through different strategies depending on the arithmetic operation: by increasing retrieval, in the case of multiplication, but by increasing efficient procedural strategies in the case of subtraction. In sum, a developmental frontal-to-parietal shift of activation along with increased engagement of hippocampal areas seems to accompany arithmetic development and learning in children. These results can be compared with the results of our own studies investigating children's numerical learning using the learning platform, which are described below.

Neurocognitive Foundation of Arithmetic Learning on the Learning Platform

Three studies were conducted to investigate changes in brain activation due to arithmetic learning by means of the learning platform: short-term effects of a 2-week training, immediate effects of one session training, and continuous changes during training. In order to evaluate changes in brain activation, functional near-infrared spectroscopy (fNIRS) and electroencephalography (EEG) were utilized.

fNIRS comes with some advantages over fMRI involving the possibility of measurement in ecologically valid situations, like school settings (Dresler et al., 2009; Obersteiner et al., 2010), or with whole body movement (Bahnmueller, Dresler, Ehlis, Cress, & Nuerk, 2014). It is suitable for measuring brain activation in children in upright body postures, like sitting behind a desk and in front of a computer. fNIRS is a comparably cheap method in contrast to other brain-imaging methods such as fMRI; it is also easily applicable and highly versatile, which altogether allows frequent measurement repetitions (for more see Ehlis, Schneider, Dresler, & Fallgatter, 2014). With regard to its functioning, it can be said that activation of certain brain areas leads to increased cerebral blood flow in these areas, reflecting increased oxygen consumption (Scholkmann et al., 2014). This results in changes in oxyhemoglobins (O_2Hb) and deoxyhemoglobins (HHb). Non-invasively, fNIRS records these changes as an indirect measure of brain activation.

However, it needs to be mentioned that fNIRS has some limitations, including restricted penetration depth and spatial resolution (Wabnitz et al., 2010), confounding influences of extracranial signals (Haeussinger et al., 2014), and peripheral hemodynamic parameters such as skin perfusion (for a review see Scholkmann et al., 2014). Balancing advantages and limitations, fNIRS seems to be a promising tool to investigate cognitive development in children and adults (e.g., Dresler et al., 2009; Verner, Herrmann, Troche, Roebers, & Rammsayer, 2013).

On the other hand, EEG offers very high temporal, but low spatial, resolution, and is relatively sensitive to motion. It is much cheaper than many brain imaging tools and because it is portable, it is easily applicable in very different situations such as in schools. An advantage of EEG is that brain oscillations (i.e., neural electric activation) can be recorded and analyzed in different ways. For instance, cognitive and motor processes lead to so-called event-related potentials (ERPs) and also to changes in continuous EEG in the form of event-related synchronization and desynchronization (ERS/ERD) (Pfurtscheller, 2001). ERS/ERD for specific frequency bands has been associated with particular cognitive functions (Pfurtscheller, 2001; Pfurtscheller & Da Silva, 1999) and thus reflect quantifiable measures of brain dynamics (Pfurtscheller & Aranibar, 1977). Previous studies indicated that theta and alpha frequency bands are sensitive to cognitive tasks such as arithmetic processing (e.g., Dolce & Waldeier, 1974). For instance, task complexity, attentional and domain-general cognitive demands, as well as memory load, lead to theta ERS (i.e., an increase in theta power) but alpha ERD (i.e., a decrease in alpha power) (Antonenko, Paas, Grabner, & van Gog, 2010; Gevins, Smith, McEvoy, & Yu, 1997; Klimesch, 1999; Moeller, Wood, Doppelmayr, & Nuerk, 2010; Pfurtscheller & Da Silva, 1999).

Short-Term Neurocognitive Changes

A group of typically developing children received seven sessions of training on simple (one-digit × one-digit, e.g., 3 × 7) and complex (one-digit × two-digit, e.g., 3 × 17) multiplication using the multiplication game implemented on the learning platform over 2 weeks. The game entailed a multiple-choice paradigm in which children had to click on the correct one out of 12 presented solution choices. The task was speeded and had to be performed in competition against a virtual computer player. The effect of a 2-week training period was evaluated using

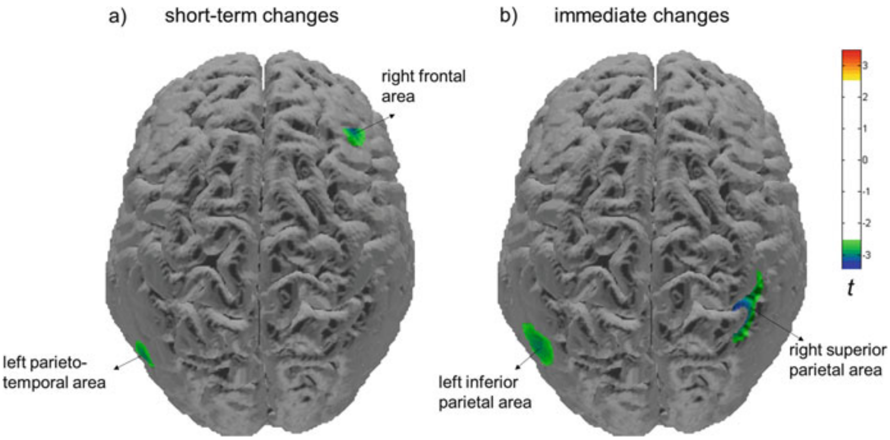

Fig. 7.4 Panel (**a**) illustrates changes in brain activation due to short-term arithmetic training in children. Panel (**b**) illustrates immediate brain activation changes due to arithmetic training in children. Color coding: *blue/green* indicates a reduction of activation

simultaneous fNIRS and EEG before and after the training. The 2-week training data indicated that children became more efficient in the trained compared to untrained multiplication problems, which was reflected in faster responses and fewer errors.

With respect to the trained simple condition, no significant change was observed in the fNIRS data. However, decreased alpha ERD (i.e., increased alpha power) was found at the central frontal electrode. This decrease suggests reduced cognitive effort in general (Pfurtscheller, 2001), probably reflecting more retrieval-based solution strategies in trained simple multiplications after the training (see also Gevins et al., 1997).

For the trained complex condition, fNIRS findings revealed reduced activation in left parieto-temporal as well as right frontal areas after training (cf. Fig. 7.4a). It has been shown that learning changes the relation of domain-general to more domain-specific processing demands, which is indicated by reduced activation in several brain regions (Poldrack, 2000). In agreement with this and also studies in adults (for a review see Zamarian et al., 2009), we observed reduced activation in right frontal areas, associated with executive control and WM, after the training (see also Soltanlou et al., 2017). This indicated an increase in retrieval-based solution strategies, which depend less on domain-general cognitive processes, following complex multiplication training (see also Prado et al., 2014).

Furthermore, the observation of decreased activation in left parieto-temporal areas is in line with longitudinal and training studies in children (Qin et al., 2014; Supekar et al., 2013), but is contradictory to multiplication training studies in adults, which have reported increased activation of left language-related parietal areas after training (for a review Zamarian et al., 2009; but see Bloechle et al., 2016). It seems that although a shift from effortful procedural to retrieval-based strategies is represented by a frontal-to-parietal shift of brain activation and increased activation of language-related parieto-temporal areas in adults, the same might not necessarily be true for children (see also Supekar et al., 2013). This difference might be due to more stable neural substrates of arithmetic processes in adults as opposed to children (Qin et al., 2014). Furthermore, the reduced activation of language-related left parietal areas is in line with an fMRI study by Menon, Rivera, White, Glover, and Reiss (2000), which also reported decreased activation of language-related parietal areas with increasing arithmetic expertise (see also Amalric & Dehaene, 2016). It should be noted, however, that even for adults, several brain areas, and not only left-hemispheric language-related parieto-temporal areas, have been associated with retrieval processes after multiplication training (Bloechle et al., 2016; Delazer et al., 2005).

Immediate Neurocognitive Changes

Immediate training effects were evaluated using simultaneous fNIRS and EEG before and after a single session of the same training by means of a multiplication game implemented on the learning platform. Interestingly, in the absence of any significant behavioral improvement, a pre- and post-test comparison of fNIRS data showed reduced activation at left inferior parietal and right superior parietal areas after training, during the calculation of trained complex multiplication problems (cf. Fig. 7.4b). This finding is in line with the longitudinal fMRI data by Qin et al. (2014), indicating reduced activation of bilateral parietal regions in children after a year of schooling. A decrease in parietal activation, which is usually associated with processing numerical magnitude, may indicate that after the training, children seemed to rely less on manipulations of numerical magnitudes to solve the task. This finding is in line with an fMRI study by Ischebeck et al. (2007) that reported similar brain activation changes in adults after a training on complex multiplication problems. In sum, this decrease seems to indicate reduced demands on exact calculation (e.g., Dehaene, Molko, Cohen, & Wilson, 2004). Moreover, we also observed increased alpha ERD at the central parietal electrode for complex multiplication, which might indicate increased visual attentional processes (Klimesch, Sauseng, & Hanslmayr, 2007). In simple multiplication, no significant difference was observed, probably because children of this age (i.e., fifth grade) were quite advanced in solving simple multiplication problems, and more training than just one session would have been needed to improve their performance.

Ongoing Neurocognitive Changes

In the next step, neurophysiological changes during arithmetic learning using the same multiplication game in the learning platform were investigated in children. Findings showed gradually increasing power in the theta (4–7 Hz) and lower alpha (8–10 Hz) bands, but not in the upper alpha band (10–13 Hz) over six repetitions of multiplication problems (cf. Fig. 7.5).

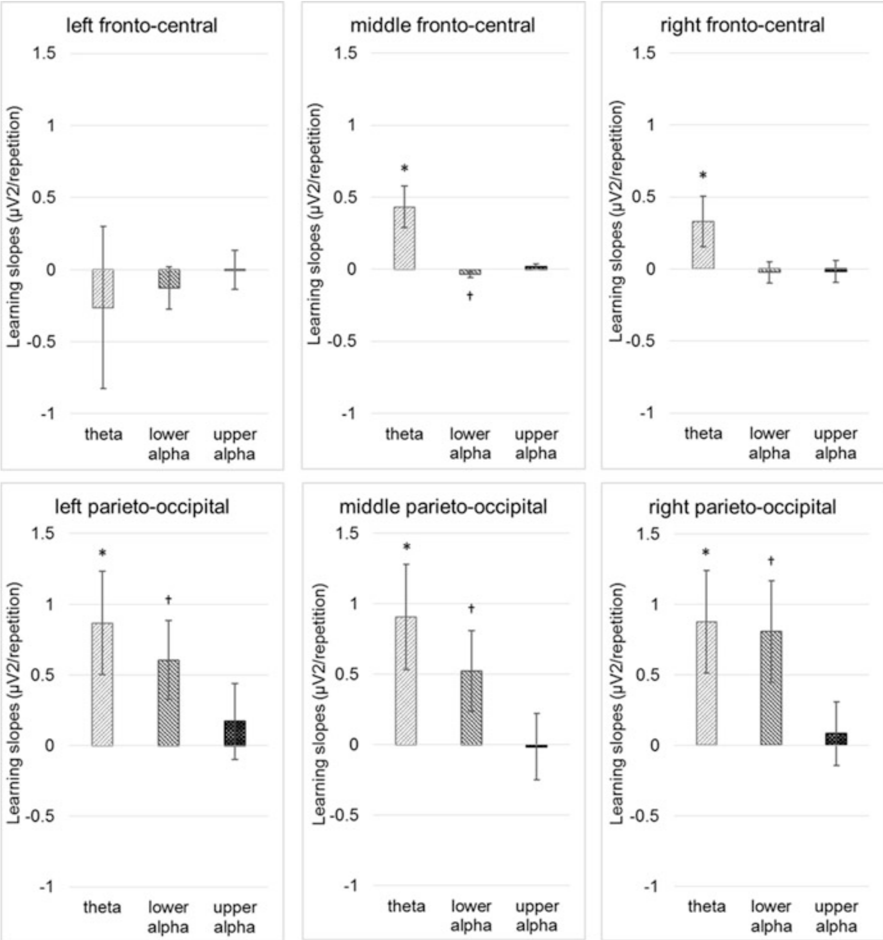

Fig. 7.5 Unstandardized coefficients (learning slopes) of EEG power density in theta, lower alpha, and upper alpha bands for different brain regions. Positive values display increased power density during six repetitions. Negative values display decreased power density during six repetitions. *Error bars* depict 1 SE. *FDR corrected $p < .05$; †FDR corrected $p <. 1$

Importantly, this fits nicely with the results of Klimesch, Vogt, and Doppelmayr (1999), who reported higher theta power in individuals with high as compared to individuals with low calculation skills—corresponding to the training effect observed in the present study. Increased theta power during multiplication training in children is probably associated with acquiring new information (Klimesch, 1999), which usually needs additional attention and mental effort, rather than retrieving existing knowledge (e.g., Gevins et al., 1997; Mizuhara & Yamaguchi, 2007; Sammer et al., 2007). In line with this argument, Gevins et al. (1997) found an increase in theta power resulting from a brief WM training in adults and attributed it to the extra effort required to focus attention over a long time.

According to a model by Baroody (1983), mathematical training entails a shift from slow procedural processes towards compact procedural strategies, which can lead to a temporarily increased involvement of domain-specific and domain-general cognitive processes (see also Núñez-Peña & Suárez-Pellicioni, 2012; Prado et al., 2014). In the present study, it seems that children primarily improved their procedural and algorithm-based strategies over six repetitions of the multiplication problems, which led to fewer errors but also increased theta power (see also Barrouillet & Thevenot, 2013). These findings point to the idea that procedural strategies improve as a function of learning in children in an early phase of building up new arithmetic fact knowledge (see also Lemaire, 2016; Zhou et al., 2011).

Although it seems that these few repetitions led to a shift from more effortful to more efficient procedural processes, there might be a shift to retrieval strategies as well, or at least an increase in retrieval processes as a part of efficient procedural strategies. It has been shown that executive functioning is involved even in arithmetic fact retrieval (Hinault & Lemaire, 2016). Therefore, an increased use of retrieval processes during arithmetic learning might also lead to increased involvement of domain-general cognitive processes, resulting in theta increase. In sum, increased theta power seems to indicate more efficient performance, resulting from not only more efficient procedural but also more retrieval-based strategies.

Furthermore, an increase in lower alpha power at bilateral occipito-parietal electrodes was observed in the present study. This is in line with previous arithmetic and cognitive training studies in adults (Gevins et al., 1997; Grabner & De Smedt, 2012). For instance, Gevins et al. (1997) suggested that increased automaticity through training is associated with increased power in the lower alpha band, reflecting a reduction in domain-general cognitive demands (Pfurtscheller, 2001), task difficulty (e.g., Gevins et al., 1997), and attentional demands (Ray & Cole, 1985). It has been found that procedural strategies demand more cognitive resources than retrieval strategies in multiplication problem solving in children (Lemaire, Barrett, Fayol, & Abdi, 1994). Therefore, in agreement with Grabner and De Smedt (2012), we conclude that decreased alpha power is associated with more procedural strategies. In contrast, increased power in lower alpha as observed in the current study may represent more automatic, presumably retrieval-based, strategies in arithmetic problem solving (e.g., Moeller et al., 2010). It seems that children first shift to more efficient procedural strategies before they shift to retrieval strategies.

Conclusion of Neurocognitive Evaluation

The findings of the above-described studies nicely demonstrated that arithmetic learning using the multiplication game of our learning platform in typically developing children led to measurable neurocognitive changes (even in the absence of reliable improvements in behavioral performance). In particular, the results of our studies indicated that the development of arithmetic competencies occurs in two steps: the first one from slow effortful procedural processes to faster more efficient procedural processes, and the second step to retrieval-based solution strategies. In the following section, the theoretical and practical implications of these results are discussed.

General Discussion, Limitations, and Perspectives of the Neurocognitive Approach

A neurocognitive approach to learning and education, which we have embraced in the above studies and which is specifically pursued in the emerging field of Educational Neuroscience, has been both hyped and debunked. For instance, Bowers (2016a) puts forth his skepticism about such an approach:

“Neuroscience cannot determine whether instruction should target impaired or nonimpaired skills. More importantly, regarding the assessment of instruction the only relevant issue is whether the child learns, as reflected in behavior. Evidence that the brain changed in response to instruction is irrelevant”.

In the on-going debate, there have been fierce responses arguing that *“Bowers' assertions misrepresent the nature and aims of the work in this new field [of Educational Neurosceince]. We suggest that, by contrast, psychological and neural levels of explanation complement rather than compete with each other“(see also Bowers,* 2016b*; Gabrieli,* 2016*; Howard-Jones et al.,* 2016*)*. Without reiterating the back and forth of this controversial debate in more detail, we wish to make two statements about what our approach as described above can and cannot contribute to (educational) science: one about the nature of neurocognitive data, and one on the localization of cognitive functions within the brain.

The Construct and Its Operationalization: Neurocognitive Data are Just Another Dependent Variable

At times researchers or the public may have gained the impression that neurocognitive data reflect the “real” underlying variables of human cognition and behavior, whereas behavioral variables may only be second-hand, indirect indices. It is important to mention that we do not want to promote such an impression nor do we share the underlying notion. Instead, our position on this is that neurocognitive data are at the beginning just another sort of dependent variables such as error rates, reaction times, eye movements, motion parameters, and other so-called *behavioral* measures. Most of these behavioral measures in learning and instruction research are indirect in the sense that they do not directly reflect the ultimate goal of instruction and education. For instance, in math education, our goal is that children can solve math problems correctly.

However, since the beginning of the so-called cognitive era in psychological and educational research (starting in the late 1950s), it is almost undisputed that we do not only need to do behaviorist experiments, in which children simply need to commit fewer errors in math at the end, regardless of why and how. Instead, we are specifically interested in understanding the underlying cognitive processes and representations, the strategies and procedures employed, the contributions of domain (i.e., math)-specific and domain-general (e.g., WM) processes to human development; and, in turn, as we learn to understand these aspects, the best ways to teach such processes, along with developmental issues inherent in learning and instruction.

In fact, however, some of these processes, strategies, and so forth are quite abstract concepts. We cannot study them directly, but only indirectly through operationalization. A particular concept is broken down into graspable questions, and we derive particular hypotheses about specific operationalizations. We usually do not measure WM or mathematics competencies per se, but substitute performance in a task or test that we are confident provides an appropriate proxy measure for the respective construct. The dependent variables in such tasks or tests are not just accuracy (the primary concern in mathematics education), but much more often behavioral measurements like reaction times or eye movement or motion data. In this context, it may not be of interest for education per se whether a particular arithmetic problem is solved 30 ms faster or slower under a given condition or after a particular training—in the end it may have little practical relevance whether one solves a simple arithmetic problem in 800 or 830 ms. However, it may have important theoretical relevance, because it informs us about the way children and adults represent numerical knowledge. For instance, the observation that children represent multi-digit numbers in a decomposed fashion (i.e., separated into units, tens, etc., cf. Nuerk, Kaufmann, Zoppoth, & Willmes, 2004) from early on in their numerical development was derived from reaction time data. Furthermore, eye-tracking data indicated that even a simple addition task involving two-digit numbers seems to comprise at least three different underlying processes (Moeller et al., 2011b). Taken together, this illustrates that most behavioral data are not used to directly measure whether a child has learned something, but to understand what, how, and why a child has learned.

Importantly, the very same goal is pursued by neurocognitive data. We do not want to claim that neurocognitive data are superior to behavioral data. Like behavioral data, neurocognitive data only measure particular indices and not the underlying cognitive processes and representations in the brain—and most certainly not localizations of cognitive functions/representations in the brain (see below for a more detailed discussion of this). However, neurocognitive data are powerful dependent variables, and considering them will help us grasp the underlying processes in children's and adults' learning, by their temporal and spatial characteristics in different conditions and in contrast to other studies.

The above-described studies provide a good example for this argument. We know that adults and children can learn multiplication facts by drill when the association between operands and result is repeated sufficiently often. In fact, they usually get faster and more accurate with training, at least in the trained problems. In case we are not interested in the processes and representations underlying this learning effect, we could simply leave it at that. However, when we are interested in the what, how, and why of the learning process, neurocognitive data may be informative. In previous adult studies, the increase in performance was associated with a shift in brain activation argued to reflect a change in strategy from effortful calculation to overlearned arithmetic fact retrieval (Delazer et al., 2003). No such shift was observed in our data. Rather the data suggested that children automatized and facilitated their calculation procedures first. When we only considered the behavioral data, we would only observe that children and adults improved very similarly after training; and thus, we would most likely assume that the same processes underlie their multiplication learning. However, complementary neurocognitive data suggested that brain activation patterns associated with the learning process may differ between children and adults.

That is why we think that it is essential to understand the what, how, and why of the learning process to enhance and promote optimal learning procedures, instructions, and environments. Neurocognitive data can contribute to this understanding in addition to other (behavioral) data like eye-movement or motion data.

Function is Not Region: The Need to Go Beyond Localistic Modular Ideas of Neurocognitive Functioning

Early approaches in cognitive neuroscience have often been localistic and modular. This means that for a particular function like vision, hearing, motion, or somatosensory perception, the aim has been to identify brain regions that are selectively activated, which are then assumed to subserve the function.

While this approach has worked to a certain degree for basic sensory and motor functions, as reflected by evidence for a visual cortex, an auditory cortex, a motor cortex, and a somato-sensory cortex, the idea of localized, the modular mental function is less realistic for higher cognitive functions, such as numerical cognition. The original triple code model of numerical cognition by Dehaene (Dehaene & Cohen, 1997) has focused—although white matter connections were implicitly suggested—heavily on such anatomo-functional associations. There was the proposition of a representation of number magnitude in the bilateral intra-parietal sulci, a verbal representation of number in the left-hemispheric superior temporal area and the angular gyrus, and a visual number form area in bilateral inferior occipito-temporal areas. In recent years, however, this model has been extended and networks with their white matter connections were included in network models of numerical cognition and arithmetic (Klein et al., 2016). Moreover, brain regions that were supposed to subserve a particular numerical representation are found to actually serve multiple functions when literature from different fields is considered. For instance, the angular gyrus thought to be involved in arithmetic fact retrieval may serve a more general role in a network with hippocampal structures jointly underlying fact retrieval (e.g., Bloechle et al., 2016; see also Harvey, Klein, Petridou, & Dumoulin, 2013 for the involvement of the intra-parietal sulcus).

The more a cognitive function is evolutionarily advanced, the less it is likely to be subserved by only a single brain region. Instead, whole networks seem to underlie complex cognitive functions like arithmetic, which makes neuroscientific research on these functions at the same time challenging but also more interesting and promising. After this discussion of the general benefits of neurocognitive data for educational science, the final section of the chapter focuses on challenges to and the implications of our results for the idea of numerical learning promoted by a web-based learning platform.

Challenges and Implications

A central goal of the development of the learning platform was to provide a personalized learning experience tailored to the specific needs of each individual learner. This is a highly desirable feature given the fact that there is considerable variance in orthography and arithmetic competencies. Moreover, motivation to use the platform needs to be maintained at a high level, in particular for children with specific deficits in orthography or numeracy. Presenting material that is too difficult may cause frustration and should thus be avoided. This may be achieved by adjusting the level of difficulty of the learning games to more directly target existing deficits. Such adaptivity in the platform can be realized at both an intra-individual and an inter-individual level.

For intra-individual adaptivity, the system needs to have a valid model of strengths and weaknesses of the individual user. Therefore, these need to be assessed at a fine-grained level—at least for novices to the system—but also in between playing the learning games, to monitor learning progress. To keep these assessments short, an adaptive procedure seems most suitable.

A prototype of such an adaptive assessment approach was developed based on modeling relevant competencies by Item-Response Theory (IRT), a state-of-the-art methodology (Fleischer, Leutner, & Klieme, 2012; Hartig & Frey, 2012). Rich sets of items were analyzed and spelling/numerical errors were evaluated with respect to particular orthographic/arithmetic markers. Item difficulty parameters were determined separately according to the Rasch model, for each arithmetic operation (addition, subtraction, multiplication, etc.) and orthographic rule (capitalization, gemination, lengthening, etc.), and were used to assess proficiency with respect to all of these dimensions through the respective person parameters. The assessment proceeds in an adaptive manner by presenting the most informative item next. Notice that testing a German word like "Rennbahn" provides information on various competencies simultaneously: capitalization (initial capital letter), gemination (double consonant "nn"), and lengthening (long vowel "ah"). Thus, the corresponding information values were integrated to select the next word. Also, note that the spelling of a typed-in word needs to be checked automatically with respect to each of the orthographic markers. Efficient pattern matching algorithms were applied to provide such an online evaluation of the spelling. The adaptive assessment is stopped as soon as a certain level of confidence on the person parameters is achieved, at which point the learning games can be adapted based on these parameters. The difficulty of items in the learning games can be matched to the person parameters, so that the task is neither too demanding nor too easy. In sum, a proof of concept for computer-based adaptive assessment of spelling could be successfully developed and awaits implementation.

Inter-individual adaptivity refers to balancing the odds of two players in competitive learning games and is a functionality that is currently under development. The estimated person parameters of above-described assessments may be used to match players according to their competencies, or to simulate a computer-controlled opponent of the same ability profile. Moreover, game play behavior may be modified with respect to the student's performance relative to (a) her/his current opponent and (b) the set of all registered players. Individual log data may be exploited to provide an index of competency (like an ELO score in chess; Elo, 1978), which allows for specifying handicaps for the better performing student in any pair of players. Such handicaps can be realized via time-delays in item presentation or answer registration. This makes it possible to balance success rates and keep the motivation of the individual players high, quite independent of their ability level.

Such inter-individual and intra-individual adaptivity remains a central goal of any learning game: Web-based and computer-based games provide—grounded on IRT-based methods—a promising tool for incorporating such adaptivity. In our view, this is essential to prevent poor performers (most in need of learning) from quitting learning games due to lack of success and reward.

Finally, a recapitulating view on our project allows for some important statements about learning and educational practice:

- Informal learning settings (such as our learning platform) can successfully complement formal learning settings in schools as shown by the observed training effects. New media like the internet can be particularly powerful, because the games we designed can be played anywhere where internet connections are available, with any device, inside or outside of school. The application of new digital media and game-based learning for interventions are not limited to academia and education, but also may also be used in a broader field of mental and physical healthcare, as it was shown in Chap. 5 of this book (Zurstiege et al. 2017).
- Game-based learning of cultural competencies like orthography or numeracy works. Math learning (and to a lesser degree, orthography) is usually considered highly aversive by many children, and, unfortunately, also by many teachers. Game-based approaches can provide a new approach to these important learning topics and nevertheless increase learning progress.
- Individualization is not just a matter of human judgment and expertise. While we strongly appreciate individualization and differentiation of the learning progress by teachers, we are also aware that it requires enormous effort. The time needed may not always be available, and in large classes it may sometimes not be easy or even possible to evaluate the particular strengths and weaknesses of each

individual child on each new topic properly and correctly. Data-driven adaptivity as we have developed it for diagnostics of the spelling process may help to identify individual problems and help to tailor individual learning and instruction.
- Neurocognitive data can help to enhance our understanding of the learning and instruction process. In this project, we gained first insights that multiplication learning in children might be different from that in adults, although both improve in accuracy and reaction time during the learning process. Neurocognitive data are certainly not the only data allowing us to understand the what, how, and why of learning, but they are nevertheless a powerful complementary dependent variable to enhance this process.

Taken together, mastering literacy and numeracy remains one of the great challenges in education and instruction in our modern knowledge society. At the same time, respective deficits are a severe obstacle in professional and private life, impacting not only individuals but also societies at large. New, innovative methods for administrating and conducting orthography and numeracy training as well as evaluating and understanding the underlying learning processes revealed promising results in our web-based learning project. Because it is essential to achieve mastery of literacy and numeracy for as many as possible, we should use such computer-supported, adaptive, and game-based methods to foster learning opportunities and thus enrich life and career prospects for as many children as possible.

References

Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *Proceedings of the National Academy of Sciences*, *113*(18), 4909–4917.

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience*, *9*(4), 278–291.

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, *22*(4), 425–438.

Arsalidou, M., & Taylor, M. J. (2011). Is 2+ 2= 4? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage*, *54*(3), 2382–2393.

Bahnmueller, J., Dresler, T., Ehlis, A.-C., Cress, U., & Nuerk, H.-C. (2014). NIRS in motion—Unraveling the neurocognitive underpinnings of embodied numerical cognition. *Frontiers in Psychology*, *5*, 743.

Baroody, A. J. (1983). The development of procedural knowledge: An alternative explanation for chronometric trends of mental arithmetic. *Developmental Review*, *3*(2), 225–230.

Barrouillet, P., & Thevenot, C. (2013). On the problem-size effect in small additions: Can we really discard any counting-based account? *Cognition*, *128*(1), 35–44.

Beddington, J., Cooper, C. L., Field, J., Goswami, U., Huppert, F. A., Jenkins, R., ... Thomas, S. M. (2008). The mental wealth of nations. *Nature*, *455*(7216), 1057–1060.

Bloechle, J., Huber, S., Bahnmueller, J., Rennig, J., Willmes, K., Cavdaroglu, S., ... Klein, E. (2016). Fact learning in complex arithmetic—The role of the angular gyrus revisited. *Human Brain Mapping*, *37*(9), 3061–3079.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *42*(1), 189.

Bowers, J. S. (2016a). The practical and principled problems with educational neuroscience. *Psychological Review*, *123*(5), 600–612.
Bowers, J. S. (2016b). Psychology, not educational neuroscience, is the way forward for improving educational outcomes for all children: Reply to Gabrieli (2016) and Howard-Jones et al. (2016). *Psychological Review*, *123*(5), 628–635.
Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., … Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education*, *94*, 178–192.
Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science*, *332*(6033), 1049–1053.
Cantlon, J. F., Libertus, M. E., Pinel, P., Dehaene, S., Brannon, E. M., & Pelphrey, K. A. (2009). The neural development of an abstract concept of number. *Journal of Cognitive Neuroscience*, *21*(11), 2217–2229.
Carlisle, J. F., & Katz, L. A. (2006). Effects of word and morpheme familiarity on reading of derived words. *Reading and Writing*, *19*(7), 669–693.
Chen, Z.-H., Liao, C. C., Cheng, H. N., Yeh, C. Y., & Chan, T.-W. (2012). Influence of game quests on pupils' enjoyment and goal-pursuing in math learning. *Educational Technology & Society*, *15*(2), 317–327.
Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A., & Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities*, *39*(6), 507–514.
Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, *11*(4), 227–268.
Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex*, *33*(2), 219–250.
Dehaene, S., Molko, N., Cohen, L., & Wilson, A. J. (2004). Arithmetic and the brain. *Current Opinion in Neurobiology*, *14*(2), 218–224.
Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*(3–6), 487–506.
Delazer, M., Domahs, F., Bartha, L., Brenneis, C., Lochy, A., Trieb, T., & Benke, T. (2003). Learning complex arithmetic—An fMRI study. *Cognitive Brain Research*, *18*(1), 76–88.
Delazer, M., Ischebeck, A., Domahs, F., Zamarian, L., Koppelstaetter, F., Siedentopf, C., … Felber, S. (2005). Learning by strategies and learning by drill—Evidence from an fMRI study. *NeuroImage*, *25*(3), 838–849.
Dolce, G., & Waldeier, H. (1974). Spectral and multivariate analysis of EEG changes during mental activity in man. *Electroencephalography and Clinical Neurophysiology*, *36*, 577–584.
Domahs, F., Delazer, M., & Nuerk, H. C. (2006). What makes multiplication facts difficult: Problem size or neighborhood consistency? *Experimental Psychology*, *53*(4), 275–282.
Dresler, T., Obersteiner, A., Schecklmann, M., Vogel, A. C. M., Ehlis, A.-C., Richter, M. M., … Fallgatter, A. J. (2009). Arithmetic tasks in different formats and their influence on behavior and brain oxygenation as assessed with near-infrared spectroscopy (NIRS): A study involving primary and secondary school children. *Journal of Neural Transmission*, *116*(12), 1689–1700.
Ehlis, A.-C., Schneider, S., Dresler, T., & Fallgatter, A. J. (2014). Application of functional near-infrared spectroscopy in psychiatry. *NeuroImage*, *85*, 478–488.
Elo, A. E. (1978). *The rating of chessplayers, past and present*. New York, NY: Arco Pub.
Emerson, R. W., & Cantlon, J. F. (2015). Continuity and change in children's longitudinal neural responses to numbers. *Developmental Science*, *18*(2), 314–326.
Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education*, *67*, 156–167.
Esser, G., Wyschkon, A., & Schmidt, M. H. (2002). Was wird aus Achtjährigen mit einer Lese-und Rechtschreibstörung. *Zeitschrift für Klinische Psychologie und Psychotherapie; Forschung und Praxis*, *31*(4), 235–242.
Evans, T. M., Kochalka, J., Ngoon, T. J., Wu, S. S., Qin, S., Battista, C., & Menon, V. (2015). Brain structural integrity and intrinsic functional connectivity forecast 6 year longitudinal growth in children's numerical abilities. *The Journal of Neuroscience*, *35*(33), 11743–11750.

Fischer, G. H. (1995). *The linear logistic test model Rasch models* (pp. 131–155). New York, NY: Springer.
Fleischer, J., Leutner, D., & Klieme, E. (2012). Modellierung von Kompetenzen im Bereich der Bildung: eine psychologische Perspektive. *Psychologische Rundschau*, *63*, 1–2.
Gabrieli, J. D. (2016). The promise of educational neuroscience: Comment on Bowers (2016). *Psychological Review*, *123*(5), 613–619.
Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming*, *33*(4), 441–467.
Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385.
Giorgidze, G., Grust, T., Schreiber, T., & Weijers, J. (2010). Haskell boards the Ferry.. Paper presented at the Symposium on Implementation and Application of Functional Languages.
Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F., & Neuper, C. (2009). To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia*, *47*(2), 604–608.
Grabner, R. H., Ansari, D., Reishofer, G., Stern, E., Ebner, F., & Neuper, C. (2007). Individual differences in mathematical competence predict parietal brain activation during mental calculation. *NeuroImage*, *38*(2), 346–356.
Grabner, R. H., & De Smedt, B. (2012). Oscillatory EEG correlates of arithmetic strategies: A training study. *Frontiers in Psychology*, *3*.
Grabner, R. H., Ischebeck, A., Reishofer, G., Koschutnig, K., Delazer, M., Ebner, F., & Neuper, C. (2009). Fact learning in complex arithmetic and figural-spatial tasks: The role of the angular gyrus and its relation to mathematical competence. *Human Brain Mapping*, *30*(9), 2936–2952.
Gross, J. (2006). *The long term costs of literacy difficulties*. Montvale, NJ: KPMG Foundation.
Gross, J., Hudson, C., & Price, D. (2009). *The long term costs of numeracy difficulties*. London: Every Child a Chance Trust and KPMG.
Haeussinger, F. B., Dresler, T., Heinzel, S., Schecklmann, M., Fallgatter, A. J., & Ehlis, A.-C. (2014). Reconstructing functional near-infrared spectroscopy (fNIRS) signals impaired by extra-cranial confounds: An easy-to-use filter method. *NeuroImage*, *95*, 69–79.
Hanushek, E. A., & Woessmann, L. (2010). *The high cost of low educational performance: The long-run economic impact of improving PISA outcomes*. Paris: OECD.
Hartig, J., & Frey, A. (2012). Konstruktvalidierung und Skalenbeschreibung in der Kompetenzdiagnostik durch die Vorhersage von Aufgabenschwierigkeiten. *Psychologische Rundschau*, *63*, 43–49.
Harvey, B., Klein, B., Petridou, N., & Dumoulin, S. (2013). Topographic representation of numerosity in the human parietal cortex. *Science*, *341*(6150), 1123–1126.
Hinault, T., & Lemaire, P. (2016). What does EEG tell us about arithmetic strategies? A review. *International Journal of Psychophysiology*, *106*, 115–126.
Howard-Jones, P., Varma, S., Ansari, D., Butterworth, B., De Smedt, B., Goswami, U., ... Thomas, M. (2016). The principles and practices of educational neuroscience: Commentary on Bowers (2016). *Psychological Review*, *123*(5), 620–627.
Huber, S., Fischer, U., Moeller, K., & Nuerk, H.-C. (2013). On the interrelation of multiplication and division in secondary school children. *Frontiers in Psychology*, (4), 740.
Huber, S., Moeller, K., & Nuerk, H.-C. (2012). Differentielle Entwicklung arithmetischer Fähigkeiten nach der Grundschule: Manche Schere öffnet und schließt sich wieder. *Lernen und Lernstörungen*, *1*, 119–134.
Ischebeck, A., Zamarian, L., Egger, K., Schocke, M., & Delazer, M. (2007). Imaging early practice effects in arithmetic. *NeuroImage*, *36*(3), 993–1003.
Ischebeck, A., Zamarian, L., Schocke, M., & Delazer, M. (2009). Flexible transfer of knowledge in mental arithmetic—An fMRI study. *NeuroImage*, *44*(3), 1103–1112.
Ischebeck, A., Zamarian, L., Siedentopf, C., Koppelstätter, F., Benke, T., Felber, S., & Delazer, M. (2006). How specifically do we learn? Imaging the learning of multiplication and subtraction. *NeuroImage*, *30*(4), 1365–1375.

Ise, E., & Schulte-Körne, G. (2010). Spelling deficits in dyslexia: Evaluation of an orthographic spelling training. *Annals of Dyslexia*, *60*(1), 18–39.

ISFE. (2012). *Videogames in Europe: Consumer study. European summary report*. Brussels: ISFE.

Jung, S., Roesch, S., Huber, S., Heller, J., Grust, T., Nuerk, H., & Moeller, K. (2015). *An interactive web-based learning platform for arithmetic and orthography.* Paper presented at the Advances in Computers and Technology for Education–Proceedings of the 11th International Conference on Educational Technologies.

Kast, M., Baschera, G.-M., Gross, M., Jäncke, L., & Meyer, M. (2011). Computer-based learning of spelling skills in children with and without dyslexia. *Annals of Dyslexia*, *61*(2), 177–200.

Kaufmann, L., Wood, G., Rubinsten, O., & Henik, A. (2011). Meta-analyses of developmental fMRI studies investigating typical and atypical trajectories of number processing and calculation. *Developmental Neuropsychology*, *36*(6), 763–787.

Kawashima, R., Taira, M., Okita, K., Inoue, K., Tajima, N., Yoshida, H., … Fukuda, H. (2004). A functional MRI study of simple arithmetic—A comparison between children and adults. *Cognitive Brain Research*, *18*(3), 227–233.

Klein, E., Suchan, J., Moeller, K., Karnath, H.-O., Knops, A., Wood, G., … Willmes, K. (2016). Considering structural connectivity in the triple code model of numerical cognition: Differential connectivity for magnitude processing and arithmetic facts. *Brain Structure and Function*, *221*, 979–995.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2), 169–195.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88.

Klimesch, W., Vogt, F., & Doppelmayr, M. (1999). Interindividual differences in alpha and theta power reflect memory performance. *Intelligence*, *27*(4), 347–362.

Kucian, K., von Aster, M., Loenneker, T., Dietrich, T., & Martin, E. (2008). Development of neural networks for exact and approximate calculation: A FMRI study. *Developmental Neuropsychology*, *33*(4), 447–473.

Landerl, K. (2006). Reading acquisition in different orthographies: Evidence from direct comparisons. In R. M. Joshi & P. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 513–530). Mahwah, NJ: L. Erlbaum.

Lemaire, P. (2016). *Cognitive aging: The role of strategies*. New York, NY: Psychology Press.

Lemaire, P., Barrett, S. E., Fayol, M., & Abdi, H. (1994). Automatic activation of addition and multiplication facts in elementary school children. *Journal of Experimental Child Psychology*, *57*(2), 224–258.

Link, T., Huber, S., Nuerk, H.-C., & Moeller, K. (2014). Unbounding the mental number line—New evidence on children's spatial representation of numbers. 4: 1021

Link, T., Moeller, K., Huber, S., Fischer, U., & Nuerk, H.-C. (2013). Walk the number line—An embodied training of numerical concepts. *Trends in Neuroscience and Education*, *2*(2), 74–84.

Link, T., Nuerk, H.-C., & Moeller, K. (2014). On the relation between the mental number line and arithmetic competencies. *The quarterly journal of experimental psychology*, *67*(8), 1597–1613.

Lumsden, J., Edwards, E. A., Lawrence, N. S., Coyle, D., & Munafò, M. R. (2016). Gamification of cognitive assessment and cognitive training: A systematic review of applications and efficacy. *JMIR Serious Games*, *4*(2), e11.

Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2017). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*, *71*, 525–534.

Menon, V., Rivera, S. M., White, C. D., Glover, G. H., & Reiss, A. L. (2000). Dissociating prefrontal and parietal cortex activation during arithmetic processing. *NeuroImage*, *12*(4), 357–365. doi:10.1006/nimg.2000.0613

Mizuhara, H., & Yamaguchi, Y. (2007). Human cortical circuits for central executive function emerge by theta phase synchronization. *NeuroImage*, *36*(1), 232–244.

Moeller, K., Fischer, U., Nuerk, H.-C., & Cress, U. (2015). Computers in mathematics education–Training the mental number line. *Computers in Human Behavior*, *48*, 597–607.

Moeller, K., Klein, E., & Nuerk, H. C. (2011a). (No) Small adults: Children's processing of carry addition problems. *Developmental Neuropsychology*, *36*(6), 702–720.

Moeller, K., Klein, E., & Nuerk, H. C. (2011b). Three processes underlying the carry effect in addition—Evidence from eye tracking. *British Journal of Psychology*, *102*(3), 623–645.

Moeller, K., Wood, G., Doppelmayr, M., & Nuerk, H.-C. (2010). Oscillatory EEG correlates of an implicit activation of multiplication facts in the number bisection task. *Brain Research*, *1320*, 85–94.

Ninaus, M., Kiili, K., McMullen, J., & Moeller, K. (2016). A game-based approach to examining students' conceptual knowledge of fractions. In R. Bottino et al. (Eds.), *Games and learning alliance* (pp. 37–49). New York, NY: Springer.

Ninaus, M., Pereira, G., Stefitz, R., Prada, R., Paiva, A., Neuper, C., & Wood, G. (2015). Game elements improve performance in a working memory training task. *International Journal of Serious Games*, *2*, 3–16.

Ninaus, M., Witte, M., Kober, S. E., Friedrich, E. V., Kurzmann, J., Hartsuiker, E., ... Wood, G. (2013). Neurofeedback and serious games. *Psychology, Pedagogy, and Assessment in Serious Games*, *82*.

Nuerk, H.-C., Kaufmann, L., Zoppoth, S., & Willmes, K. (2004). On the development of the mental number line: More, less, or never holistic with increasing age? *Developmental Psychology*, *40*(6), 1199.

Nuerk, H.-C., Moeller, K., Klein, E., Willmes, K., & Fischer, M. H. (2011). Extending the mental number line: A review of multi-digit number processing. *Zeitschrift Fur Psychologie-Journal of Psychology*, *219*(1), 3–22. doi:10.1027/2151-2604/a000041

Nuerk, H.-C., Moeller, K., & Willmes, K. (2015). Multi-digits numerical understanding. In R. CohenKadosh & A. Dowker (Eds.), *Oxford handbook of mathematical cognition* (pp. 106–139). Oxford: Oxford University Press.

Núñez-Peña, M. I., & Suárez-Pellicioni, M. (2012). Processing false solutions in additions: Differences between high-and lower-skilled arithmetic problem-solvers. *Experimental Brain Research*, *218*(4), 655–663.

Obersteiner, A., Dresler, T., Reiss, K., Vogel, A. C. M., Pekrun, R., & Fallgatter, A. J. (2010). Bringing brain imaging to the school to assess arithmetic problem solving: Chances and limitations in combining educational and neuroscientific research. *ZDM*, *42*(6), 541–554.

Parsons, S., & Bynner, J. (2005). *Does numeracy matter more?* London: National Research and Development Centre for Adult Literacy and Numeracy.

Pauli, P., Lutzenberger, W., Rau, H., Birbaumer, N., Rickard, T. C., Yaroush, R. A., & Bourne, L. E. (1994). Brain potentials during mental arithmetic: Effects of extensive practice and problem difficulty. *Cognitive Brain Research*, *2*(1), 21–29.

Pfurtscheller, G. (2001). Functional brain imaging based on ERD/ERS. *Vision Research*, *41*(10), 1257–1260.

Pfurtscheller, G., & Aranibar, A. (1977). Event-related cortical desynchronization detected by power measurements of scalp EEG. *Electroencephalography and Clinical Neurophysiology*, *42*(6), 817–826.

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857.

Poldrack, R. A. (2000). Imaging brain plasticity: Conceptual and methodological issues—A theoretical review. *NeuroImage*, *12*(1), 1–13.

Prado, J., Mutreja, R., & Booth, J. R. (2014). Developmental dissociation in the neural responses to simple multiplication and subtraction problems. *Developmental Science*, *17*(4), 537–552.

Prins, P. J., Dovis, S., Ponsioen, A., Ten Brink, E., & Van der Oord, S. (2011). Does computerized working memory training with game elements enhance motivation and training efficacy in children with ADHD? *Cyberpsychology, Behavior, and Social Networking*, *14*(3), 115–122.

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology*, *14*(2), 154.

Qin, S., Cho, S., Chen, T., Rosenberg-Lee, M., Geary, D. C., & Menon, V. (2014). Hippocampal-neocortical functional reorganization underlies children's cognitive development. *Nature Neuroscience*, *17*(9), 1263–1269.

Ray, W. J., & Cole, H. W. (1985). EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science*, *228*(4700), 750–752.

Rivera, S. M., Reiss, A., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex*, *15*(11), 1779–1790.

Roesch, S., Jung, S., Huber, S., Artemenko, C., Bahnmueller, J., Heller, J., … Moeller, K. (2016). Training arithmetic and orthography on a web-based and socially-interactive learning platform. *International Journal of Education and Information Technologies*, *10*, 204–217.

Sammer, G., Blecker, C., Gebhardt, H., Bischoff, M., Stark, R., Morgen, K., & Vaitl, D. (2007). Relationship between regional hemodynamic activity and simultaneously recorded EEG-theta associated with mental arithmetic-induced workload. *Human Brain Mapping*, *28*(8), 793–803.

Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Pavia, J. M., Wolf, U., & Wolf, M. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *NeuroImage*, *85*, 6–27. doi:10.1016/j.neuroimage.2013.05.004

Soltanlou, M., Artemenko, C., Dresler, T., Haeussinger, F. B., Fallgatter, A. J., Ehlis, A.-C., & Nuerk, H.-C. (2017). Increased arithmetic complexity is associated with domain-general but not domain-specific magnitude processing in children: A simultaneous fNIRS-EEG study. *Cognitive, Affective, & Behavioral Neuroscience*, 1–13.

Spüler, M., Krumpe, T., Walter, C., Scharinger, C., Rosenstiel, W., & Gerjets, P. (2017). Brain-computer interfaces for educational applications. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 177–201). New York, NY: Springer.

Supekar, K., Swigart, A. G., Tenison, C., Jolles, D. D., Rosenberg-Lee, M., Fuchs, L., & Menon, V. (2013). Neural predictors of individual differences in response to math tutoring in primary-grade school children. *Proceedings of the National Academy of Sciences*, *110*(20), 8230–8235.

Verner, M., Herrmann, M. J., Troche, S. J., Roebers, C. M., & Rammsayer, T. H. (2013). Cortical oxygen consumption in mental arithmetic as a function of task difficulty: A near-infrared spectroscopy approach. *Frontiers in Human Neuroscience*, *7*, 217.

Wabnitz, H., Moeller, M., Liebert, A., Obrig, H., Steinbrink, J., & Macdonald, R. (2010). Time-resolved near-infrared spectroscopy and imaging of the adult human brain. In E. Takahashi & D. Bruley (Eds.), *Oxygen transport to tissue XXXI* (pp. 143–148). Boston, MA: Springer.

Whyte, J. C., & Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. *Developmental Psychology*, *44*(2), 588.

Wilson, A., Dehaene, S., Pinel, P., Revkin, S., Cohen, L., & Cohen, D. (2006). Principles underlying the design of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*, *2*, 19.

Wilson, A., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*, *2*(1), 1.

Zamarian, L., Ischebeck, A., & Delazer, M. (2009). Neuroscience of learning arithmetic—Evidence from brain imaging studies. *Neuroscience & Biobehavioral Reviews*, *33*(6), 909–925.

Zhou, X., Booth, J. R., Lu, J., Zhao, H., Butterworth, B., Chen, C., & Dong, Q. (2011). Age-independent and age-dependent neural substrate for single-digit multiplication and addition arithmetic problems. *Developmental Neuropsychology*, *36*(3), 338–352.

Zurstiege, G., Zipfel, S., Ort, A., Mack, I., Meitz, T. G. K., & Schäffeler, N. (2017). Managing obesity prevention using digital media—A double-sided approach. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 97–123). New York, NY: Springer.

Chapter 8
Brain-Computer Interfaces for Educational Applications

Martin Spüler, Tanja Krumpe, Carina Walter, Christian Scharinger, Wolfgang Rosenstiel, and Peter Gerjets

Introduction

Currently, there is an ongoing debate in education research on how to optimally support learners during learning (Calder 2015; Askew 2015). Obviously, learning outcome is most promising if the training program and learning content is tailored to the learner's specific needs (e.g., Gerjets & Hesse 2004; Richards et al. (2007); for numerical interventions see Dowker 2004; Karagiannakis & Cooreman 2014). To optimally support the learner's efforts, it is important that the learning content is neither too easy nor too difficult. Therefore, it is crucial for successful learning to keep the cognitive workload in the individual optimal range for each learner (Gerjets, Scheiter, & Cierniak 2009; Sweller, Van Merrinboer, & Paas 1998). This can be achieved by adapting the difficulty of the training content to the individual competencies of the learner.

Computer-supported learning (Kirschner & Gerjets 2006) seems specifically suited for implementing adaptivity, because these informational environments can be easily extended with algorithms that change the difficulty of the presented material based on the learner's behavioral response. This allows for an easy personalization of the learning environment to the user's individual needs, which is assumed to be key for efficient learning. So far, adaptive computer-supported learning environments rely on the user's interaction behavior for adaptation, e.g.,

M. Spüler (✉) • T. Krumpe • C. Walter • W. Rosenstiel
Department of Computer Engineering, Eberhard-Karls University Tübingen, Tübingen, Germany
e-mail: spueler@informatik.uni-tuebingen.de

C. Scharinger • P. Gerjets
Knowledge Media Research Center, Tübingen, Germany

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_8

the difficulty of the training content is adapted based on the number of correct responses (Corbett 2001; Graesser & McNamara 2010; Käser et al. 2013). These behavioral measures are rather indirect and distal measures that are not very specific with respect to inferring the cognitive processes required for performing the task at hand. For instance, more errors in a row may not only be caused by the difficulty of the task itself but also by task-unspecific processes (e.g., lapses of attention, fatigue, or disengagement).

As already outlined in Chap. 7 of this book (Soltanlou et al. 2017), there are several neurocognitive processes involved in learning, which can be found in recorded neurophysiological data. With the advent of Brain-Computer Interfaces (BCIs), which translate brain activity into control signals (Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan 2002), there seems to be a new technology that could use these neurocognitive processes in the context of educational learning environments. For measuring brain activity, electroencephalography (EEG) is a widely used method for BCIs. The measured activity is then processed, filtered, and certain features are extracted out of the brain signals. For translation into control signals, different machine learning methods are used. Those methods need a certain amount of training data (e.g., preprocessed EEG data) and class labels (e.g., user thinks left or right). Each data point has to be assigned one of the two class labels. Based on the data and the corresponding labels, the algorithms find patterns in the data that belong to either of the classes. These patterns can be used to create a prediction model which allows to predict to which class a new data point belongs.

While traditional BCIs allow a user to communicate or control a computer by brain activity only (Spüler 2015), they are more recently also used to extract information about the user and infer his or her mental states (e.g., workload, vigilance).

As BCIs can be used to measure cognitive processes, this allows for a more direct and implicit monitoring of the learner's state and should thereby allow for a better adaptation of the training content to improve learning success of the user. That the amount of cognitive workload can be measured by EEG has been shown by multiple studies (Gevins, Smith, McEvoy, & Yu 1997; Murata 2005; Berka et al. 2007; Wang, Hope, Wang, Ji, & Gray 2012). While workload is reflected in the amplitude of event-related potentials in the EEG (Brouwer et al. 2012; Causse, Fabre, Giraudet, Gonzalez, & Peysakhovich 2015), it also affects the oscillatory activity observed in an EEG (Kohlmorgen et al. 2007; Brouwer et al. 2012). Specific to arithmetic tasks, it was shown that the cognitive demand results in a power increase in the theta band and a decrease in the alpha band (Harmony et al. 1999). As there are known effects of workload in the EEG, these can potentially be detected by a BCI and used for adaptation of a learning environment.

In the course of this chapter, we will give an overview of the progress made in the last 7 years, while trying to develop an EEG-based learning environment. Based on the lessons learned from the first studies for workload assessment, we present results that show that EEG can be used to assess workload during arithmetic learning. By developing a BCI that detects workload in real-time, we could show that this approach can also be used for content adaptation in a learning environment. While this system only detected workload as one component, we could also show that different sub-components of workload can be differentiated and thereby be used to gain more detailed information about the learner's mental state. Finally, we discuss these findings, show open problems, and present ideas how this line of research can be extended in the following years.

The First Steps Are the Hardest: What We Have Learned in the Early Days

When we started the project to develop an EEG-based learning environment, we performed two studies in which we addressed the workload assessment during learning of realistic instructional material and the use of cross-task workload assessment to assess workload in a variety of different tasks. While the results of these studies were mixed, they uncovered several issues in designing an EEG-based learning environment. Here we briefly describe these studies and summarize the lessons we have learned to avoid those issues in the future.

Workload Detection During Studying Realistic Material

In our first study (Walter, Cierniak, Gerjets, Rosenstiel, & Bogdan 2011), we analyzed workload in ten subjects (age 12–14 years) who studied instructional material while EEG was recorded. There were two types of instructional material, which involved processes of learning and comprehension with different levels of workload. The material inducing a high workload consisted of graphical representations and explanations of mathematical angle theorems, while the low workload material consisted of comic strips in which the subjects had to understand the story. While both materials contained a complex graphical display, the comic strips were easy to understand, while the angle theorems were harder and required more effort.

An analysis of the EEG data showed that spectral power in the alpha band (8–13 Hz) over the occipital area showed a strong desynchronisation (decrease of power) for high-workload material. Using a support vector machine (SVM), we tried to classify the data into categories of high and low workload on a single-trial level and obtained an average accuracy of 76%, which shows that EEG can be used to detect workload.

Although the results are promising for developing an EEG-based learning environment, they have to be interpreted with caution as some methodological critique can be applied to that study regarding perceptual-motor confounds. As comic strips and angle theorems differed largely in their appearance, this might lead to differences in perceptual processing beyond imposing different levels of workload. The different perception might also lead to different motor behavior (e.g., different eye movement patterns), which could lead to different motor-related brain activity. Although the results are in line with the literature, some doubts remain if the effects are solely caused by workload modulation.

Cross-Task Workload Prediction

While the previous results have shown that machine learning methods can be used to detect workload from EEG data, another problem becomes apparent when trying to use this approach in a learning environment. For the application of machine learning, training data are needed, which also includes labels that assign each datapoint to one class (such as easy or hard). In the previous study, data from the same task was used. In a learning environment this approach can not be used, due to learning effects of the subject. These learning effects (task is hard at the beginning, easy after learning) change the labels of the data, which makes it difficult to apply machine learning approaches.

As an alternative approach to calibrate an EEG-based learning environment, we wanted to use basic psychological tasks (n-back, go-nogo, reading span) that induce workload and use data from these tasks to train a classifier, which can afterwards be used in a realistic learning scenario. Therefore, we recorded EEG data from 21 subjects who performed the three basic psychological tasks for workload induction and two realistic tasks. The two realistic tasks were comprised of algebra and arithmetic tasks with different levels of difficulty. The tasks were designed to avoid perceptual motor confounds. More details can be found in (Walter, Schmidt, Rosenstiel, Gerjets, & Bogdan 2013).

When trying to classify the data, we found that a within-task classification (by cross-validation) works very well for the three basic tasks with an average accuracy >95%, while accuracies of 73% were reached for the arithmetic task and 67% for the algebra task. For the cross-task classification, in which the classifier was trained on data from one task, and tested on another task, we reached accuracies around 50% (chance level) regardless of the combination of the tasks. These results show that the EEG data contains workload-related activity that can be detected by machine learning methods within a task, but not across different tasks. To identify reasons why the cross-task classification failed, we also investigated the labels. We used the objective difficulty, which ranges from 1 (e.g., 0-back) to 3 (e.g., 2-back). This approach assumes that a 2-back task is as difficult as an arithmetic task of level 3. As this is a bold assumption that may not necessarily be true, we also investigated the subjective labels, as each subject rated the subjective difficulty of each trial. Unfortunately, the subjective ratings could not be used for all subjects, as some subjects had a strong bias in their response. However, for the remaining subjects, the cross-task classification clearly improved. When using two basic tasks for training, and testing on either the algebra or the arithmetic task, classification worked significantly above chance level for 16 out of 27 tests with accuracies $\geq$70%.

Pitfalls in Designing EEG-Based Learning Environments

As we have discussed in a previous paper (Gerjets, Walter, Rosenstiel, Bogdan, & Zander 2014), the results from the first two studies led to some insight, which can be summarized as four lessons that we have learned.

The first lesson was to avoid perceptual-motor confounds. If there are possible perceptual-motor confounds it is difficult to tell if classification occurs solely on the basis of workload-related changes in the brain activity or if the classifier is unintentionally trained to detect perceptual- or motor-related brain activity. By using simple material, which rules out any perceptual-motor confounds, one can ensure that classification occurs solely on workload-related brain responses. However, this raises another question, how well a classifier trained on workload data obtained from a simple task is able to predict workload in a realistic and complex scenario, where perceptual-motor confounds can not be eliminated.

The second lesson is about the task order in the context of learning. Randomizing the task order with regard to task difficulty is usually a good choice to avoid confounds. However, when it comes to learning, this is hard to implement due to learning effects. A task with medium difficulty might induce high workload in the beginning. With time, the subject learns how to perform the task and a medium difficulty will induce low workload after the subject has mastered the task. This shows not only that the labeling of the data during learning is problematic, but also that data from the same task and subject should not be used for prediction in a learning environment.

Lesson three is about using different tasks. As a cross-task classification avoids some problems (lesson 2), it has some different issues. Although perceptual-motor confounds should be avoided, there is always task-related brain activity present, which might be different when performing different tasks. Thereby, for a cross-task classification to work, it is advisable to use many different tasks. Thereby the classifier can be trained on the overall workload pattern, without overfitting to task-specific activity patterns.

Lesson four is also related to the labeling. As we have seen that a cross-task classification using multiple tasks for training is advisable, the questions of labeling arise. As the difficulty of a task is hard to compare between two tasks, and might also be perceived differently by different subjects, an objective label might not be the best choice. It would be better to use subjective labeling to assess the perceived difficulty of the task. While this might lead to a better quality of the data, it is more difficult to work with, as some subjects tend to give highly biased ratings (leading to imbalanced data) and additional time is needed to acquire the subject rating.

EEG-Based Workload Assessment During Arithmetic Learning

As a next step towards the development of an EEG-based learning environment, we performed a study in which the subjects performed arithmetic addition tasks. The EEG data collected during that experiment was then used to train a classifier to predict the difficulty of the task, which we used as measure for the expected mental workload.

Study Design

Ten students (six female; mean age: 24.9±5.3 years, 17–32) participated voluntarily in this study and received monetary compensation for participation. All participants reported normal or corrected-to-normal vision and no mathematical problems. Participants were chosen randomly. Nevertheless, all were university students (with different fields of study) and can thus be considered as having a high educational background.

The tasks consisted of simple arithmetic exercises in the form $x + y =?$, where the addends ranged from single-digit to four-digit numbers. Difficulty of a task was assessed with the Q-value (Thomas 1963), which ranged from 04 (e.g. 1+1) to 7.2 (addition of two four-digit numbers). The main parameters for problem difficulty in addition are the size of the summands involved (i.e., problem size Stanescu-Cosson et al. 2000) and whether or not a carry operation is needed (Kong et al. 2005). As Q-value takes into account both problem size and the need for a carry operation it is a more comprehensive measure of task difficulty as compared to using only one parameter such as problem size. Moreover, the Q-value also considers additional aspects, which should affect task difficulty (e.g., the reduced difficulty of specific problems such as $1000 + 1000$). Therefore, the Q-value is an adequate measure of task difficulty. In Spüler et al. (2016) more detail on how we calculated the Q-value is given.

Each participant had to solve 240 addition problems while EEG was measured. Due to possible learning effects, the tasks were presented with increasing difficulty. The time course of the experiment is illustrated in Fig. 8.1. The calculation task was shown for 5 s, in which the participants could solve the addition problem. After the 5 s, presentation of the task disappeared and the subjects had up to 3.5 s to enter their answer. The activation interval during the calculation phase was later used for analysis of the EEG signal. By separating the response from the calculation phase, we tried to avoid motor confounds in the EEG data during the calculation phase.

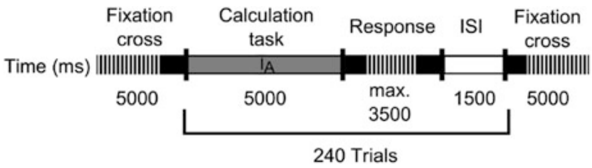

Fig. 8.1 Schematic illustration of the time course of the experiment. The *grey area* indicates the considered activation interval (I_A), whereas the *white area* represents response and inter-stimulus interval (ISI). At the beginning and the end of the experiment, a fixation cross was displayed

EEG Recording

A set of 28 active electrodes (actiCap, BrainProducts GmbH) attached to the scalp placed according to the extended International Electrode 10–20 Placement System (FPz, AFz, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, T7, C3, Cz, C4, T8, CPz, P7, P3, Pz, P4, P8, PO7, POz, PO8, O1, Oz, O2), was used to record EEG signals. Three additional electrodes were used to record an electrooculogram (EOG); two of them were placed horizontally at the outer canthus of either the left or the right eye, respectively, to measure horizontal eye movements, and one was placed in the middle of the forehead between the eyes to measure vertical eye movements. Ground and reference electrodes were placed on the left and right mastoids. EOG- and EEG-signals were amplified by two 16-channel biosignal amplifier systems (g.USBamp, g.tec) and sampled at a rate of 512 Hz. EEG data was high-pass filtered at 0.5 Hz and low-pass filtered at 60 Hz during the recording. Furthermore a notch-filter was applied at 50 Hz to filter out power line noise.

Analysis of EEG Data Regarding Workload-Related Effects

To remove the influence of eye movements we applied an EOG-based regression method (Schlögl et al. 2007). For analyzing the data we estimated the power spectrum during the calculation phase of each trial. We used Burg's maximum entropy method (Cover & Thomas 2006) with a model order of 32 to estimate the power spectrum from 1 to 40 Hz in 1 Hz bins.

As an indication of workload-related effects in the power spectrum we calculated R^2-values (squared Pearson's correlation coefficient) between the Q-value and the power at each electrode and frequency bin. In Fig. 8.2, the results for different frequency bands are shown. While workload-related changes can be observed in the delta and lower beta frequency band, changes are strongest in the theta and alpha band. Location of activation changes slightly between theta and alpha, with theta having a more centro-parietal distribution and alpha being stronger over the parieto-occipital region.

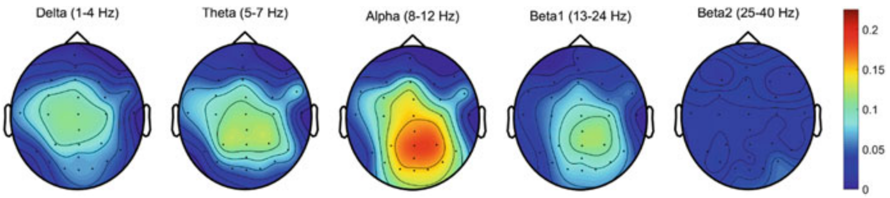

Fig. 8.2 Topographic display of R^2 values averaged over all participants, showing the influence of the Q-value for each electrode in different frequency bands. R^2 color coded with increasing *red reflecting* higher R^2 values

Prediction of Task Difficulty

To predict the task difficulty based on the EEG data, we processed the data similarly to that for the analysis in the previous section. In an earlier attempt to predict workload using this data (Walter, Wolter, Rosenstiel, Bogdan, & Spüler 2014), which is not presented in this section, we did not use an EOG-based artifact correction and found that the prediction performance is largely influenced by eye-movement artifacts. Thus, here we used an EOG-based regression method (Schlögl et al. 2007) to eliminate eye-movement artifacts from the data. We again used Burg's maximum entropy method (Cover & Thomas 2006) with a model order of 32 to estimate the power spectrum from 1 to 40 Hz in 1 Hz bins.

To correct for inter-participant variability in the participant's baseline EEG power, the first 30 trials (easy trials, with $Q < 1$) were used for normalization. The mean and standard deviation for each frequency bin at each electrode was calculated over the first 30 trials and the remaining 210 trials were scaled according to these means and standard deviations. The 30 trials used for normalization were not used further in the prediction process (either for training or for testing the model). Based on the normalized data, we trained a linear ridge regression model (regularization parameter $\lambda = 10^3$ determined by cross-validation on the training data). To reduce the number of features we used only every second frequency bin (1 Hz, 3 Hz, 5 Hz,...) to train the model. When not specified otherwise, we used the 17 inner electrodes (F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, CPz, P3, Pz, P4, POz, O1, Oz, O2) resulting in a total of 340 features used for prediction. However, other electrode subsets were also tested.

The prediction was tested in two different approaches. For the first approach, the within-subject prediction, data from the same subject were used for training and testing the prediction model. To ensure that data do not overlap, a 10-fold cross-validation was performed. The within-subject prediction was done only to get an estimate of the data quality and how well task difficulty can be predicted in an optimal scenario. The second approach, the cross-subject prediction, used data from nine subjects for training the model and was tested on the one remaining subject. The procedure was repeated so that data from each subject was used for testing once. With regard to a future application of the method in an online learning environment, the within-subject approach is less interesting, because it would mean that the learning environment needs to be calibrated to the user, which means that the user has to use the system for a while to collect a certain amount of data needed for training. The cross-subject prediction, on the other hand, allows the use of data from other subjects for training a prediction model. Using this approach in a learning environment would enable the subject to start using the system right away without the need for the collection of training data.

For evaluation of the prediction model, we used different metrics: correlation coefficient (CC) and root-mean-squared-error (RMSE). As both metrics assess different kinds of prediction error, it is important to look at both metrics (Spüler, Sarasola-Sanz, Birbaumer, Rosenstiel, & Ramos-Murguialday 2015).

As a regression predicts the Q-value, we assume that a precise prediction of task difficulty will not be necessary in an adaptive learning environment. It is more important to identify when the task gets too easy or too difficult to keep the learner in his optimal range of cognitive workload, which is why we also used the output of the regression model for a classification into three difficulty levels. For $Q < 2$ the trial was considered easy, for $Q > 4$ it was considered too difficult, while it had a medium difficulty level when Q was between 2 and 4. These ranges were estimated based on the relationship between Q and the behavioral task performance. To estimate the validity of our classification procedure, we applied this thresholding to the actual Q and the predicted Q and evaluated the classification accuracy (CA), which is given as the percentage of trials correctly classified.

Performance of Workload Prediction

The cross-subject prediction performance averaged over all subjects reached an average CC of 0.74, RMSE of 1.726, and a CA of 49.8% (chance level would be 33%). Regarding the classification accuracy, trials of normal difficulty ($2 < Q < 4$), were correctly classified with an accuracy of 74.2%, while it was difficult to differentiate easy 4.8% accuracy) and difficult (30.4%) trials. A closer look at the prediction revealed that trials with $Q > 6$ were mostly classified as easy, indicating that easy trials and trials with $Q > 6$ share similar patterns in EEG. To improve performance we used only trials with $Q < 6$ for training and testing the model, which resulted in a significantly better CC of 0.82 ($p < 0.01$, t-test), RMSE of 1.343 ($p < 0.001$, t-test), and an improved classification accuracy of 55.8% ($p < 0.05$, t-test) (Table 8.1).

As it would be beneficial for an EEG-based learning environment to use as few EEG electrodes as possible, we also investigated how performance changes if the number of electrodes is reduced. The results with different electrodes are summarized in Table 8.1. When using only one electrode (besides the ground and reference electrode), POz was the position that delivered best results. While using POz performs significantly worse than using 17 electrodes in terms of CA ($p < 0.05$, t-test), there was no significant difference in CC and RMSE ($p > 0.05$). There

Table 8.1 Prediction performance averaged over all participants for different electrode subsets using only trials with $Q < 6$

Number	Description	Positions	CC	RMSE	CA
28	All	See section "EEG Recording"	0.807	1.292	55.6%
17	Inner	See section "Prediction of Task Difficulty"	0.820	1.343	55.8%
7	Central	Fz,FCz,Cz,CPz,Pz,POz,Oz	0.839	1.352	54.7%
9	Frontocentral	F3,Fz,F4,FC3,FCz,FC4,C3,Cz,C4	0.725	1.465	45.5%
7	Parietooccipital	P3,Pz,P4,POz,O3,Oz,O4	0.831	1.383	54.1%
1	Best electrode	POz	0.788	1.479	50.5%

In each line one subset is described by the number of electrodes, a short description referring to the approximate location of the electrodes as well the position of all electrodes in the subset according to the 10–20 system. As performance metric we showed the correlation coefficient (CC), the root mean squared error (RMSE), and the average classification accuracy (CA)

was also no significant difference in prediction performance between using the 17 inner channels, all 28 channels, or a subset of frontocentral, central, or parietocentral electrodes ($p > 0.05$).

The within-subject prediction achieved an average CC of 0.90 and an RMSE of 0.95, which is significantly better for CC (<0.01) and RMSE ($p < 0.001$), but, as discussed earlier, the reduced calibration and training effort in a cross-subject prediction seems to outweigh the performance increase of the within-subject prediction.

A Passive BCI for Online Adaptation of a Learning Environment

As we have shown in the previous section, that it is possible to predict workload during solving mathematical exercises by using a cross-subject regression model, we wanted to apply this model in an online study. This study should serve as a proof-of-concept that a digital learning environment can adapt its content in real-time based on the EEG of the user. To evaluate the effect of the learning environment, we assessed the learning success of the subjects and compared it to a control group.

Study Design and Participants

The participants in this study were divided into two groups: an experimental group using the EEG-based learning environment and a control group using an error-adaptive learning environment, which can be considered as state-of-the-art. In both groups, the subjects learned arithmetic addition in the octal number system (e.g., $5 + 3 = 10$), which was a completely new task to all subjects.

To evaluate the learning success, each subject did a pre-test and a post-test, before and after using the learning environment for approximately 45 min (180 exercises). The tests consisted of 11 exercises, with varying difficulty. Although difficulty of the exercises were the same for the pre- and post-tests, the exercises themself were different.

The participants of both groups were university students of various disciplines, reported to have normal or corrected-to-normal vision, and participated voluntarily in the EEG experiment. Thirteen subjects (7 male and 6 female; mean age: 28.1 ± 4.3 years, 21–35) participated in the experimental group using the EEG-based learning environment. The control group consisted of 11 subjects (7 male and 4 female; mean age: 23.4 ± 1.4 years, 22–27), using an error-adaptive learning environment.

Cross-Subject Regression for Online Workload Prediction

Based on the EEG data obtained from the previous study (see section "EEG-Based Workload Assessment During Arithmetic Learning"), we created a prediction model that was used to predict the cognitive workload online and further was able to adjust the learning environment accordingly. In terms of usability, we wanted the EEG-based learning environment to be useable out-of-the-box without the need for a subject-specific calibration phase, which is why we used the cross-subject regression method.

Therefore, EEG data from the previous study were used for training a linear ridge regression model with a regularization parameter of $\lambda = 10^3$. The number of electrodes used for online adaptation was reduced to 16 inner electrodes (FPz, AFz, F3, Fz, FC3, FCz, FC4, C3, Cz, C4, CPz, P3, Pz, P4, Oz, POz). Furthermore, only trials with a Q-value smaller than 6 were used to train the regression model, since trials with a higher Q-value showed similar EEG patterns as very easy trials, which is most likely due to a disengagement of the subjects. While the trained regression model allows workload prediction on a single-trial level, the accuracy of the prediction was improved by averaging the prediction output using a moving average window of the last six trials. The moving average leads to a more robust prediction, but also makes the system react slower to sudden changes in workload, which is feasible since it is not recommended to adapt the difficulty of an online learning environment too rapidly.

The prediction model thus trained was then applied in the EEG-based learning environment, to predict the amount of cognitive workload for novice subjects in real-time.

Real-Time Adaptation of the Learning Environment

For the experimental group, the EEG data served as a workload indicator. Therefore, we used the output of the previously trained regression model to predict the current workload state of each learner and differentiated three difficulty levels. If the predicted workload was less than $Q = 0.8$, the presented task difficulty was assumed to be too easy. Thus the following Q-value was increased by 0.2. Vice versa, the target Q of the subsequent task decreased by 0.2 when the predicted workload was greater than $Q = 3.5$. In this case, the presented task difficulty was assumed to be too difficult. If the predicted workload was between $Q = 0.8$ and $Q = 3.5$, the Q-value for the next presented task remained the same and the difficulty level was kept constant. These thresholds were defined based on the relationship between error rate and Q-value (Spüler et al. 2016). Trials with $Q < 1$ were solved correctly in all cases, while none of the subjects were able to solve trials with a $Q > 6$. 50% of the trials with a $Q = 3.5$ were successfully solved on average.

For the control group, an error-adaptive learning environment was used. The number of wrong answers served as performance and adaptation measure. When subjects solved five consecutive tasks correctly, the difficulty level increased by 1. On the other hand, the difficulty level decreased by a factor of 1 when participants made three errors in a row. Otherwise, the Q-value did not change and the difficulty level was held constant. The adaptation scheme was kept similar in the control group, as in common tutoring systems. Because of the repetitions till the difficulty level was changed, the presented Q-values increased or decreased by steps of size 1 and the calculated Q-values were rounded to the next integer. The learning session for the experimental group as well as for the control group started with an exercise of difficulty level $Q = 2$.

To compare the learning success of the two groups, the learning effect after completing the learning phase serves as a performance measure and is used as an indicator of how successfully each subject was supported during learning. Hence, each subject had to perform a pre-test before the learning phase started. This was used to assess the prior knowledge of each user. After the learning session, each participant had to solve a post-test, and the difference in score between the two tests served as an indicator of the learning effect.

Task Performance Results

The behavioral results, how well the subjects performed the octal arithmetic task, are shown in Table 8.2 for the experimental group and in Table 8.3 for the control group. For the experimental group, 45.5% of the 180 exercises were solved correctly on average. Averaged over all subjects, a maximum Q-value of 5.85 was reached by using the EEG-based learning environment. Each subject achieved at least the difficulty level of $Q = 3.2$ (see Table 8.2).

Table 8.2 Task performance for the experimental group

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	Mean
Correct	38.3	88.9	60.6	10.0	28.9	25.6	31.1	17.8	66.1	33.3	16.7	92.2	82.2	45.5
max. Q	6.6	3.2	6.6	6.6	6.6	6.6	6.6	6.6	5.2	6.6	6.6	3.6	4.6	5.85

The relative amount of correctly solved trials in % and the maximum Q-value for each subject using the EEG-based learning environment is shown

Table 8.3 Task performance for the control group

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	Mean
Correct	64.4	62.2	57.2	66.7	63.3	61.1	63.3	67.8	65.0	66.1	66.7	64.0
max. Q	6	4	5	6	4	5	4	4	4	5	4	4.64

The relative amount of correctly solved trials in % and the maximum Q-value for each subject using the error-adaptive learning environment is shown

Table 8.4 Number of correctly solved trials in the pre-and post-test, as well as the learned factor for each subject of the experimental group and overall subjects

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	All
Pre-Test	3	0	6	0	2	1	0	1	0	5	1	1	0	1.54
Post-Test	7	0	9	8	2	3	3	2	8	7	7	2	8	5.08
Learned	4	0	3	8	0	2	3	1	8	2	6	1	8	3.54

The control group answered 64% of all 180 assignments correctly on average. Since the error-rate was used for adapting the difficulty level of the presented learning material, the number of correctly solved trials was similar across subjects. On average, a maximum Q-value of 4.64 was reached (see Table 8.3). The best subjects reached a maximum Q-value of 6, whereas each participant achieved at least the difficulty level of $Q = 4$.

Learning Effect

To evaluate if the EEG-based learning environment works and how it compares to an error-adaptive learning environment, we analyzed the learning effect of each subject by pre- and post-tests. Furthermore, the learning effects between the experimental and control group were compared.

Table 8.4 reports the learning effect of each individual subject of the experimental group, as well as the overall values. After using the EEG-based learning environment and thus learning how to calculate in an octal number system, a learning effect can be recognized for almost every subject, except for subject 2 and 5. On average, 5.08 assignments from 11 post-test tasks were solved correctly after completing the learning phase (see Table 8.4), 3.54 more assignments compared to the pre-test. On average a significant learning effect can be verified between the pre- and post-test ($p = 0.0026$, two-sided Wilcoxon test).

Table 8.5 Number of correctly solved trials in the pre- and post-tests, as well as the learned factor for each subject of the control group and overall subjects

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	Mean
Pre-Test	4	3	1	0	2	1	2	1	1	2	0	1.55
Post-Test	8	2	4	7	5	8	3	4	4	6	4	5
Learned	4	−1	3	7	3	7	1	3	3	4	4	3.45

The control group solved 1.55 tasks from the 11 pre-test assignments on average correctly. As the difference to the experimental group is not significant ($p > 0.05$, two-sided Wilcoxon test), equal prior knowledge can be implied for both groups. The best subject of the control group performed 36.36% of the pre-test tasks accurately, whereas the worst subjects gave no correct answers (see Table 8.5). For almost every subject, a learning effect for calculating in the octal number system is noticeable after the error-adaptive learning session, except for subject 2. On average, 3.45 more tasks were solved correctly in the post-test, compared to the pre-test (see Table 8.5), which also shows a significant learning effect between the pre- and post-tests ($p = 0.0016$, two-sided Wilcoxon test).

Although the learning effect for the group using the EEG-based learning environment is higher than for the control group, this difference is not significant ($p > 0.05$, two-sided Wilcoxon test). These results show that an EEG-based adaptation of a digital learning environment is possible and that results are similar to error-adaptive learning environments, which can be considered state-of-the-art.

Towards More Differentiated Workload Estimation

We have shown in the previous section that it is possible to use a passive BCI approach to adapt content of a digital learning environment in real-time, based on the user's EEG recordings. While the EEG-based learning environment worked, it showed results that were as good as an error-adaptive learning environment. This raises the question of how we can improve the EEG-based learning environments and what can we do with an EEG-based learning environment that we can not do with an error-adaptive learning environment. In the proposed EEG-adaptive learning environment we adapted according to the user's workload, which was initially measured by the Q-value. As the Q-value was shown to highly correlate with the error-rate (Spüler et al. 2016), it is not surprising that the results are similar to the ones obtained with an error-adaptive learning environment.

As we used a very simply concept of workload in the previous work, a more detailed definition of workload could help to improve the EEG-based learning environment. Miyake and colleagues (Miyake et al. 2000) defined three components of workload: Updating, Inhibition, and Shifting. Updating is described as a process of (un-)loading, keeping, and manipulating information in memory during a certain period of time (Ecker, Lewandowsky, Oberauer, & Chee 2010), whereas Shifting can be described as the adjustment of a current mind set to new circumstances or new rules of a task (Monsell 2003). Inhibition comprises blending out (irrelevant) information or withholding actions while being involved in a cognitive task (Diamond 2013). Miyake and colleagues describe those three core functions as sharing many properties but also having distinct diversities that characterize the individual functions. If the diversities between the functions lead to different patterns of EEG activity, those patterns could be detected and used to gain more detailed information about the user. Instead of only assessing the mental workload, it was also possible to assess how demanding a task is in terms of updating, inhibition, and shifting, and the content could be adapted more precisely to match the individual needs of the user.

With the aim of developing a prediction model that is able to differentiate between workload components, we performed a study to find evidence of two of the components (Updating and Inhibition) in EEG data and see how well machine learning methods can be used to differentiate those components based on the EEG data.

Task Design: n-Back Meets Flanker

The "n-back meets flanker" task (Scharinger, Soutschek, Schubert, & Gerjets 2015b) combines an n-back task with simultaneous presentation of flanker stimuli. The task was used to find and investigate interactions between the two executive functions Updating and Inhibition in terms of behavioral data as well as on a neurophysiological level. The n-back task is known to induce Updating demands in different intensities depending on the value of n (Miyake et al. 2000; Jonides et al. 1997). In this experiment n was altered between 0, 1, and 2 in a block design. In each trial, at each stimulus presentation, the subject had to react as accurately and quickly as possible to the following question: "Is the currently presented letter the same as the one I saw n before?" A yes or no decision was required, which was recorded via a button press. In addition to the given answer, the reaction time was recorded. The stimuli consisted of a set of four letters, S, H, C, and F. The different values of n indicated how many letters need to be constantly updated in memory throughout the specified block. For $n = 0$ the subjects had to remember one predefined stimulus letter, which remained the same throughout the block, which is why no updating of memory content was necessary.

The flanker task is known to induce Inhibition demands (Eriksen 1995; Sanders & Lamers 2002). A flanking stimulus can either be congruent or incongruent to the accompanied n-back task stimulus. An incongruent flanker usually introduces conflicts in the decision making, reactions eventually need to be inhibited even though the reflex urges the subject to react otherwise. The flanker stimuli in this experiment were presented simultaneously with the n-back stimulus and they were either congruent (identical) or incongruent (different) with the presented n-back stimulus. The central letter used for the n-back task was flanked by three letters on the right and three on the left (all identical). The subjects were instructed to not react to the flanker stimuli at all. Each n-back condition was presented twice, in which 154 trials where presented per block. Half of the stimuli of each block were targets, half were non-targets. The flanker items were chosen to be incongruent in one-third of all trials, congruent in the other two-thirds of all trials. By design an incongruent—incongruent sequence of stimuli was prevented, to avoid undesired effects. In addition to that in each block, a sequence of ten targets and non-targets was randomly chosen in which the central n-back stimulus was removed and only the flanker letters were presented. This was done to recreate awareness of the flanker items.

Data Recording and Preprocessing

Twenty-two subjects (10 male and 12 female; mean age: 28.5 $\pm$ 4.3 years, 21–35) participated in the study. All participants had normal or corrected-to-normal vision, none reported neurological disorders and all were right-handed (with one exception). Twenty-two electrodes (Acticap brain products) were used for the recording at a sampling rate of 500 Hz and placed according to the 10/20 system (Jasper 1958) with the reference at right mastoid and the ground electrode at AFz.

The data was bandpass filtered between 0.4 and 40 Hz and re-referenced to the common average. To remove artifacts, a threshold of 100 μV was chosen and all trials exceeding this level were discarded, whereas trials including EOG artifacts were corrected using independent component analysis (ICA), in which artefactual independent components were removed by visual inspection. After these steps, trials were sorted and more trials were discarded. The first four trials of each block, for example, as they were always congruent non-targets, were removed as well as all congruent trials following incongruent trials because the so-called Gratton effect (the effect caused by the incongruence is bigger if the incongruent trial follows a congruent trial compared to an incongruent trial) should not influence the analysis. Furthermore, all trials in which the subject did not respond correctly to the n-back task were removed from the dataset. On average this resulted in 215 trials per subject that could be used for the analysis. More details can be found in the publication (Scharinger et al. 2015b).

Neurophysiological Analysis

To investigate if Updating and Inhibition show different patterns in the EEG activity, the data were divided into trials containing Baseline (BL), Updating 1 (Up1), Updating 2 (Up2), and Inhibition (Inh). The first 0-1,000 ms after stimulus onset were used as epoch of interest. Updating demands are supposed to be induced for

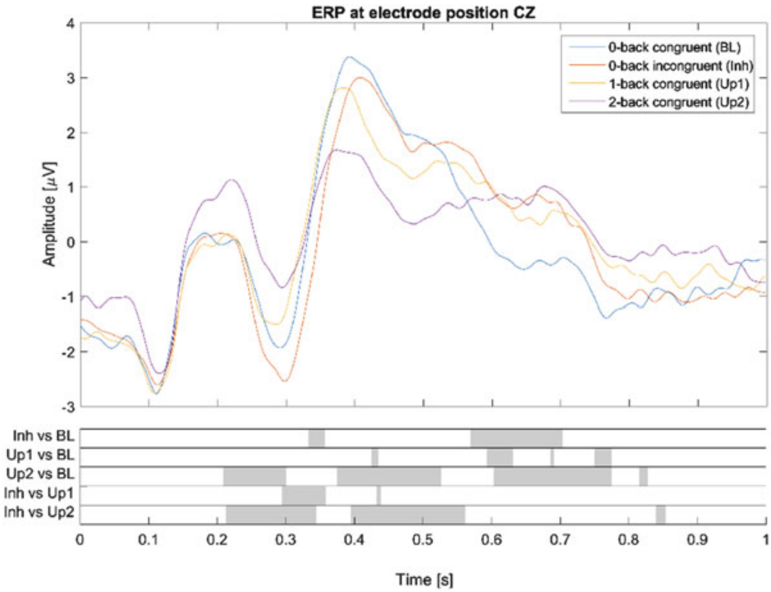

Fig. 8.3 ERPs from 0-1,000 ms from stimulus onset for the four conditions: BL, Inh, Up1, and Up2 at position Cz. The values represent the grand average over all subjects and trials. The *grey bars* (*bottom*) state that the amplitude differs significantly ($p < 0.05$, corrected for multiple comparisons) for the respective comparison

1-back and 2-back, while Inhibition demands are expected for incongruent flanker. Due to this, the data were divided into four conditions:

- Baseline (BL): 0-back with congruent flanker
- Inhibition (Inh): 0-back with incongruent flanker
- Updating1 (Up1): 1-back with congruent flanker
- Updating2 (Up2): 2-back with congruent flanker

To find evidence of diversity between Inhibition and Updating in the EEG, we first looked at the event-related potentials (ERPs) in the data. As the amplitude of the P300 at electrode positions Fz, Cz, Pz is supposed to be correlated with workload, we investigated the ERPs at those positions. While nearly no effects were found at electrode Fz, we found significant differences in amplitude between all conditions in the time range of 200–750 ms for electrodes Cz and Pz. The results for electrode Cz are shown in Fig. 8.3.

As this shows that Inhibition and Updating show different patterns in the ERPs, we further analyzed the data with regard to the power spectrum. Therefore, we used Burg's maximum entropy method (Burg 1972) with a model order of 32 to estimate the power spectrum between 4 and 13 Hz (theta and alpha band). Significant ($p < 0.05$) workload-related changes could be observed in the power spectrum in both frequency bands at Fz and Cz, while Pz showed only a significant effect in the alpha

Table 8.6 Classification accuracies averaged over all subjects using either evoked response potentials (ERPs) for classification or the power spectrum

Features	Inh vs BL	Up vs BL	Up2 vs BL	Up vs Inh	Up2 vs Inh
RP	61.12%***	61.20%***	73.38%***	68.23%***	74.73%***
Spectrum	52.48%*	52.68%*	63.17%***	55.46%***	64.13%***

Significance regarding the difference to chance level accuracy is indicated as follows: *$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$

band, which therefore was much stronger than at the other electrodes. However, the only significant effects ($p < 0.05$) between Updating (up1) and Inhibition were found in the theta band at electrode Cz.

Classification

As there is evidence that Inhibition and Updating show different patterns of EEG activity, we tried to use machine learning methods to classify the data in order to detect if a task induced Updating or Inhibition demands. For classification we used the ERP data, as well as the power spectral data. Data from 17 inner channels were used for classification by a support vector machine (SVM) with a linear kernel and a hyperparameter of $C = 1$. For the classification of the ERPs, we further used a CCA-based spatial filtering method (Spüler, Walter, Rosenstiel, & Bogdan 2014) to improve classification accuracy.

Classification results are shown in Table 8.6. While all results were significantly above chance level accuracy, classification using ERPs worked better than using the power spectrum. While Inhibition and Updating could individually be classified against baseline data, it was also possible to discriminate between Inhibition and Updating demands. While this is a clue that there might be differences between Inhibition and Updating, the results could also be explained by the different conditions inducing different amounts of workload and the classifier detecting that different amounts of workload. To rule out that hypothesis we performed a classification where the classifier was trained on one condition (e.g., BL vs. Inh) and tested on a different condition (e.g., Up1). If the patterns between the conditions were similar and only varied in strength due to different amounts of workload, we would expect that the conditions would still be classified correctly as a high-workload condition. However, most of the results were not significantly above chance ($p > 0.05$), showing that classification did not work. Only for one test, in which the classifier was trained on *Up2 vs. BL*, we obtained a classification accuracy significantly different from chance level ($p < 0.05$). As this classifier classified Inhibition trials mostly as BL, this indicates that Inhibition has a pattern more similar to Baseline than to the Up2 condition.

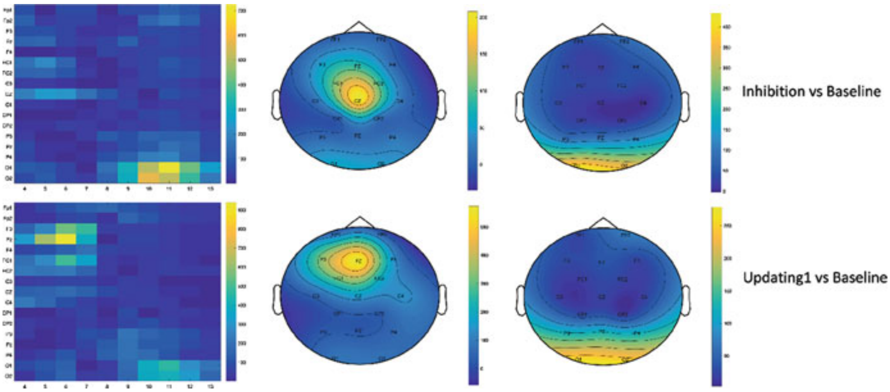

Fig. 8.4 Activation patterns obtained by SVM classification for Inhibition vs. Baseline (*top*) and Updating1 vs. Baseline (*bottom*). The activation patterns for all electrodes and frequencies are shown(*left*), a topographic display of the activation pattern for the theta-band (*middle*) and for the alpha-band (*right*)

In addition to the previous neurophysiological analysis regarding the changes in the power spectrum, we looked at the classifier weights in order to visualize the differences in the power spectra between Inhibition and Updating. As weights of a linear classifier should not be interpreted directly, we used the feature covariance matrix to calculate neurophysiologically interpretable activation patterns (Haufe et al. 2014) from the classifier weights. The activation patterns for Inh vs. BL and Up1 vs. BL are shown in Fig. 8.4. It can be seen that occipital alpha activity is a strong indicator of changes in workload for both components. In the theta-band, the pattern differed between Inhibition and Updating. While the theta activity in the central region changed with workload for the Inhibition component, the Updating component showed workload-related changes in the frontal area. Not only does the strength of the effect vary between the two components, also the location does; this is a further argument that the two components of the Miyake-model are also different on a neurophysiological level. With regard to an EEG-based learning environment, this opens up new possibilities as detecting the individual Inhibition and Updating demands of a task could lead to a more specific adaptation of the content than only detecting the overall workload.

Discussion and Outlook

In this chapter, we have outlined the progress of our project and our efforts to develop an EEG-based learning environment. While the lessons learned from our first studies were already outlined in the section "Pitfalls in Designing EEG-Based Learning Environments", the main message of these studies was that cross-task classification is not feasible. Although the cross-task classification could be improved by using subjective labels, a subjective labeling introduced other problems (difficulty to collect labels and subjective bias), which is why the cross-task classification was not investigated any further.

As an alternative, we employed a cross-subject classification, in which a prediction model is trained on a number of subjects, and used to predict workload in a different subject. Thereby, it is possible to use the workload prediction on a subject without the need for a time-consuming subject-specific calibration phase. As this method performed well when tested offline, it was tested online to adapt the content of a digital learning environment in real-time based on the EEG signals. This study was a successful proof-of-concept that an EEG-based adaptation of a digital learning environment is possible. However, we could not show any benefit of the system as it performed similarly to an error-adaptive learning environment. As this work also revealed new difficulties that arise when trying to evaluate a digital learning environment in an online study, these two issues need to be discussed in more detail.

Evaluating an EEG-Based Learning Environment

Commonly, if neural signals (like EEG) are used to estimate a user's mental state or a user's intention, the performance of the prediction method is assessed in terms of accuracy or other metrics that try to quantify how well the prediction model is working. While we have also evaluated the prediction performance of our model in the offline study (see section "EEG-Based Workload Assessment During Arithmetic Learning"), this kind of assessment is no more feasible with an online learning environment.

The reason for this is the lack of an objective measure for the user's workload. For the creation of the prediction model we used EEG data from a task that all subjects were carrying out fluently (addition in a decimal system). As no learning effects are expected in this case, the difficulty of the task (measured by Q) was used as a subjective measure of expected workload. But the sole purpose of an online learning environment is to induce learning effects, and we have shown that our system does so. Thereby, the relationship between task difficulty (Q) and workload is invalidated. At the beginning, when the task is unknown to the user, even easy exercises will induce a high workload. After using the learning environment, the user may have mastered arithmetic tasks in the octal number system, and even exercises with moderate difficulty will result in a low workload. This example shows that the relationship between task difficulty and workload is invalidated if learning effects are present.

As the task difficulty measured by the Q-value is the only objective measure we have, and the relationship to workload is invalid in the online scenario, there are no means to objectively assess the prediction performance of our model. Although we cannot assess the performance of the prediction model in the online scenario, we can evaluate the EEG-based learning environment with regard to its effect on learning success. Learning success was defined as the difference in score between the pre- and post-tests, which were performed by the subjects before and after using the learning environment. As it is not important for the learner how accurate the workload-prediction works, but it is important how much the learning success can be improved, learning success is also the most user-centered metric.

Due to these facts, that other usually used metrics are not applicable for the scenario of an online learning environment and that learning success is the most user-centered metric, we think that learning success should be used as prime outcome measure for this kind of study.

Benefits of an EEG-Based Learning Environment

As subjects using the EEG-based learning environment showed a significant learning effect, we could show that the EEG-based learning environment works. When comparing the experimental group using the EEG-based learning environment with the control group using the error-based learning environment, the learning success was higher for the experimental group, but the difference was not significant. This basically indicates that an EEG-based learning environment is an alternative to the state-of-the-art approach but offers no benefits in its current form. However, the presented system is meant as a proof-of-concept, and the parameters could be tweaked to optimize workload prediction, as well as the adaptation of the learning material.

Although we aimed at a high usability for the presented system by using a cross-subject prediction model to omit a subject-specific calibration phase, the usability of the current system is still rather low. The use of gel-based EEG needs some time and effort to prepare, thereby making it impractical to use an EEG-based learning environment on a wide basis. As we have shown in the section "EEG-Based Workload Assessment During Arithmetic Learning", the number of electrodes could be drastically reduced to reduce preparation time. Another approach would be the use of dry EEG electrodes, which would reduce the preparation time even more, as they don't require any gel. However, dry electrodes have a lower signal-to-noise ratio and would thereby lead to a lower prediction accuracy. Regardless of the preparation time, the costs of an EEG setup are high, which raises the question: what are the benefits of an EEG-based adaptive learning environment that would justify the additional cost and time investment compared to a normal digital learning environment?

It is likely that the answer is that there are no benefits when using an EEG-based learning environment in a broad population (e.g., in a class room). However, this technique could prove helpful for special cases in which the user suffers from learning disability or other problems. In the presented study, one subject of the experimental group suffered from test anxiety and reported feeling very comfortable using the EEG-based learning environment, as the system turned down the difficulty every time the subject was close to feeling overwhelmed, thereby providing a good learning experience. Depending on the individual case, the good learning experience can outweigh the disadvantages of the setup time and costs and thus render the usage feasible.

Improving an EEG-Based Learning Environment

As the EEG-based learning environment merely served as a proof-of-concept to show that the idea is feasible, it also inspires one to think about possible improvements that could be made to enhance the prediction performance. On one side, the system could be improved by using better classification and signal processing methods from the area of BCI research. While we used a simple normalization method of the EEG signal, which helped to improve results, results could be further improved by using more advanced methods to normalize EEG data like PCA-based covariate shift adaptation (Spüler, Rosenstiel, & Bogdan 2012b) or a classifier adaptation (Spüler, Rosenstiel, & Bogdan 2012a), which improves the classifier while using the learning environment and adapts it to changes in the EEG data. Regarding the cross-subject classification approach, there are also approaches from the field of transfer learning (Tu and Sun 2012; Samek, Meinecke, & Müller 2013) that could be used to improve the cross-subject (and possibly also a cross-task) classification, which should be evaluated in future studies.

Despite improving only the prediction performance, one could also try to increase the information about the user, by not only measuring workload, but also how demanding a task is regarding different subcomponents of workload. As each learner is different, one might have more problems with tasks that have a high Updating demand, while others might have more problems with high Inhibition demands. By being able to discriminate between those two, the learning content could be adapted more specifically to the individual needs of the user. Measuring the Inhibition demand could potentially also be used as an indicator of distractions outside of the learning environment and would thereby be an interesting measure for the application when applied in a classroom setting.

Measuring workload or differentiating workload into subcomponents are not the only things that could be done with an EEG-based learning environment. It was also shown that mental states like vigilance, alertness, or attention can be detected in EEG (Oken, Salinsky, & Elsas 2006; Berka et al. 2007). If those can be detected by an EEG-based learning environment this could lead to a more comprehensive monitoring of the user and not only allow for better adaptation of the learning content, but also introduce breaks or place motivating rewards based on the mental state of the user to improve learning effort and learning success of the user.

Lastly, such a BCI-based learning environment could also be improved by not only using EEG to measure neurocognitive processes, but also use other physiological signals, such as Eye-Tracking. As was shown by Scharinger, Kammerer, & Gerjets (2015a), cognitive workload influences pupil dilation, which can be measured in real-time and could be used as additional information to improve an adaptive learning environment. In the following chapter of this book (Scheiter et al.

2017), a learning environment is presented that measures eye-movements of the student and uses this additional information for adaptation. As the combination of eye-tracking information with neurophysiological signals (like EEG) would give a more comprehensive view of the current state of the learner, combining EEG with other physiological signals like eye tracking could allow for an even better adaptation of the system.

Acknowledgements This research was financed by the Leibniz ScienceCampus Tübingen "Informational Environments". It was further supported by the *Deutsche Forschungsgemeinschaft* (DFG; grant SP-1533\2-1) and the LEAD Graduate school at the Eberhard-Karls University Tübingen, which is funded by the Excellence Initiative of the German federal government.

References

Askew, M. (2015). Numeracy for the 21st century: A commentary. *ZDM: The International Journal on Mathematics Education, 47*(4), 707–712.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T.,. . ., Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space and Environmental Medicine, 78*(Supplement 1), B231–B244.

Brouwer, A.-M., Hogervorst, M. A., Van Erp, J. B., Heffelaar, T., Zimmerman, P. H., & Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of Neural Engineering, 9*(4), 045008.

Burg, J. P. (1972). The relationship between maximum entropy spectra and maximum likelihood spectra. *Geophysics, 37*(2), 375–376.

Calder, N. (2015). Student wonderings: Scaffolding student understanding within student-centred inquiry learning. *ZDM: The International Journal on Mathematics Education, 47*(7), 1121–1131.

Causse, M., Fabre, E., Giraudet, L., Gonzalez, M., & Peysakhovich, V. (2015). EEG/ERP as a measure of mental workload in a simple piloting task. *Procedia Manufacturing, 3*, 5230–5236.

Corbett, A. (2001). Cognitive computer tutors: Solving the two-sigma problem. In *Proceedings of the 8th International Conference on User Modeling* (pp. 137–147).

Cover, T., & Thomas, J. (2006). *Elements of information theory*. Hoboken, NJ: Wiley-Interscience.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135.

Dowker, A. (2004). *What works for children with mathematical difficulties?* (Vol. 554). Nottingham: DfES Publications.

Ecker, U. K., Lewandowsky, S., Oberauer, K., & Chee, A. E. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning Memory and Cognition, 36*(1), 170.

Eriksen, C. W. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition, 2*(2–3), 101–118.

Gerjets, P. H., & Hesse, F. W. (2004). When are powerful learning environments effective? the role of learner activities and of students conceptions of educational technology. *International Journal of Educational Research, 41*(6), 445–465.

Gerjets, P., Scheiter, K., & Cierniak, G. (2009). The scientific value of cognitive load theory: A research agenda based on the structuralist view of theories. *Educational Psychology Review, 21*(1), 43–54.

Gerjets, P., Walter, C., Rosenstiel, W., Bogdan, M., & Zander, T. O. (2014). Cognitive state monitoring and the design of adaptive instruction in digital environments: Lessons learned from cognitive workload assessment using a passive brain-computer interface approach. *Frontiers in Neuroscience, 8*, 385.

Gevins, A., Smith, M., McEvoy, L., & Yu, D. (1997, Jun). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty type of processing, and practice. *Cereb Cortex, 7*(4), 374–385.

Graesser, A., & McNamara, D. (2010). Self-regulated Learning in Learning Environments with Pedagogical Agents that Interact in Natural Language. *Educational Psychologist, 45*, 234–244.

Harmony, T., Ferna'ndez, T., Silva, J., Bosch, J., Valde's, P., Ferna'ndez-Bouzas, A.,. . ., Rodríguez, D. (1999). Do specific eeg frequencies indicate different processes during mental calculation? *Neuroscience Letters, 266*(1), 25–28.

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage, 87*, 96–110.

Jasper, H. (1958). The 10/20 international electrode system. *EEG and Clinical Neurophysiology, 10*, 371–375.

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber E. J., Awh, E., Minoshima, S., Koeppe, R.A. (1997). Verbal working memory load affects regional brain activation as measured by pet. *Journal of Cognitive Neuroscience, 9*(4), 462–475.

Karagiannakis, G. N., & Cooreman, A. (2014). Focused MLD intervention based on the classification of MLD subtypes. In *The Routledge International Handbook of Dyscalculia and Mathematical Learning Difficulties* (p. 265).

Käser, T., Baschera, G.-M., Busetto, A. G., Klingler, S., Solenthaler, B., Buhmann, J. M., Gross, M. (2013). Towards a framework for modelling engagement dynamics in multiple learning domains. *International Journal of Artificial Intelligence in Education, 22*(1–2), 59–83.

Kirschner, P., & Gerjets, P. (2006). Instructional design for effective and enjoyable computer-supported learning. *Computers in Human Behavior, 22*(1), 1–8.

Kohlmorgen, J., Dornhege, G., Braun, M., Blankertz, B., Müller K.-R., Curio, G.,. . ., Kincses, W. E. (2007). Improving human performance in a real operating environment through real-time mental workload detection. In *Toward Brain-Computer Interfacing* (pp. 409–422). Cambridge, MA: MIT Press.

Kong, J., Wang, C., Kwong, K., Vangel, M., Chua, E., & Gollub, R. (2005). The neural substrate of arithmetic operations and procedure complexity. *Cognitive Brain Research, 22*(3), 397–405.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134–140.

Murata, A. (2005). An attempt to evaluate mental workload using wavelet transform of EEG. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 47*(3), 498–508.

Oken, B., Salinsky M., & Elsas, S. (2006). Vigilance, alertness, or sustained attention: Physiological basis and measurement. *Clinical Neurophysiology, 117*(9), 1885–1901.

Richards, K. C., Enderlin, C. A., Beck, C., McSweeney J. C., Jones, T. C., & Rober son, P. K. (2007). Tailored biobehavioral interventions: A literature review and synthesis. *Research and Theory for Nursing Practice, 21*(4), 271–285.

Samek, W., Meinecke, F. C., & Müller, K.-R. (2013). Transferring subspaces between subjects in brain-computer interfacing. *IEEE Transactions on Biomedical Engineering, 60*(8), 2289–2298.

Sanders, A., & Lamers, J. (2002). The Eriksen flanker effect revisited. *Acta Psychologica, 109*(1), 41–56.

Scharinger, C., Kammerer, Y., & Gerjets, P. (2015a). Pupil dilation and eeg alpha frequency band power reveal load on executive functions for link-selection processes during text reading. *PloS One, 10*(6), e0130608.

Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2015b). When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating. *Psychophysiology, 52*(10), 1293–1304.

Scheiter, K., Fillisch, B., Krebs, M.-C., Leber, J., Ploetzner, R., Renkl, A., et al. (2017). How to design adaptive multimedia environments to support self-regulated learning. In *Informational Environments: Effects of Use Effective Designs* (Chap. 9).

Schlögl, A., Keinrath, C., Zimmermann, D., Scherer R., Leeb, R., & Pfurtscheller, G. (2007). A fully automated correction method of EOG artifacts in EEG recordings. *Clinical Neurophysiology, 118*(1), 98–104.

Soltanlou, M., Jung, S., Roesch, S., Ninaus, M., Brandelik, K., Heller, J., et al. (2017). Behavioral and neurocognitive evaluation of a web-based learning platform for orthography and arithmetic. In *Informational Environments: Effects of Use Effective Designs* (Chap. 7).

Spüler, M. (2015). A Brain-Computer Interface (BCI) system to use arbitrary Windows applications by directly controlling mouse and keyboard. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 4087–1090).

Spüler, M., Rosenstiel, W., & Bogdan, M. (2012a). Adaptive SVM-based classification increases performance of a MEG-based Brain-Computer Interface (BCI). In *International Conference on Artificial Neural Networks* (pp. 669–676).

Spüler, M., Rosenstiel, W., & Bogdan, M. (2012b). Principal component based covariate shift adaption to reduce non-stationarity in a MEG-based brain- computer interface. *EURASIP Journal on Advances in Signal Processing, 2012*(1), 1–7.

Spüler, M., Walter, A., Rosenstiel, W., & Bogdan, M. (2014). Spatial filtering based on canonical correlation analysis for classification of evoked or event- related potentials in EEG data. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 22*(6), 1097–1103.

Spüler, M., Sarasola-Sanz, A., Birbaumer, N., Rosenstiel, W., & Ramos-Murguialday, A. (2015). Comparing metrics to evaluate performance of regression methods for decoding of neural signals. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1083–1086).

Spüler, M., Walter, C., Rosenstiel, W., Gerjets, P., Moeller K., & Klein, E. (2016). EEG-based prediction of cognitive workload induced by arithmetic: A step towards online adaptation in numerical learning. *ZDM: The International Journal on Mathematics Education ZDM, 48*(3), 267–278.

Stanescu-Cosson, R., Pinel, P., van de Moortele, P.-F., Le Bihan, D., Cohen, L., & Dehaene, S. (2000). Understanding dissociations in dyscalculia. *Brain, 123*(11), 2240–2255.

Sweller, J., Van Merrinboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Thomas, H. B. G. (1963). Communication theory and the constellation hypothesis of calculation. *Quarterly Journal of Experimental Psychology, 15*(3), 173–191.

Tu, W., & Sun, S. (2012). A subject transfer framework for eeg classification. *Neurocomputing, 82*, 109–116.

Walter, C., Cierniak, G., Gerjets, P., Rosenstiel, W., & Bogdan, M. (2011). Classifying mental states with machine learning algorithms using alpha activity decline. In *European Symposium on Artificial Neural Networks*

Walter, C., Schmidt, S., Rosenstiel, W., Gerjets, P., & Bogdan, M. (2013). Using cross-task classification for classifying workload levels in complex learning tasks. In *Affective Computing and Intelligent Interaction (ACII), 2013* (pp. 876–881).

Walter, C., Wolter, P., Rosenstiel, W., Bogdan, M., & Spüler, M. (2014, 09). Towards cross-subject workload prediction. In *Proceedings of the 6th International Brain-Computer Interface Conference*, Graz, Austria.

Wang, Z., Hope, R. M., Wang, Z., Ji, Q., & Gray, W. D. (2012). Cross-subject workload classification with a hierarchical bayes model. *NeuroImage, 59*(1), 64–69.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology, 113*(6), 767–791.

Chapter 9
How to Design Adaptive Information Environments to Support Self-Regulated Learning with Multimedia

Katharina Scheiter, Benjamin Fillisch, Marie-Christin Krebs, Jasmin Leber, Rolf Ploetzner, Alexander Renkl, Holger Schmidt, Anne Schüler, and Gottfried Zimmermann

Learning with Multimedia Information Environments

The use of explanatory texts (spoken or written) combined with pictorial representations (e.g., diagrams, animations, or videos) as a format of instruction has become omnipresent in educational information environments. Such text-picture combinations are referred to as multimedia (Mayer, 2009). There is a large body of research showing that learning with multimedia is under most circumstances more effective than learning just from text alone, a finding commonly referred to as the "multimedia effect" (Anglin, Vaez, & Cunningham, 2004; Butcher, 2014). Different explanations for this multimedia effect have been proposed. According to the two most prominent theories in the field, the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2009) and the Integrated Model of Text and Picture Comprehension (ITPC; Schnotz, 2005), learning with text and pictures can result in richer and more elaborate mental representations of the to-be-learnt content than when studying only text. For this to happen, learners have to select relevant information from each representational format (i.e., the text or the picture) and mentally organize the information into meaningful mental structures. Most importantly, they need to establish

K. Scheiter (✉) • M.-C. Krebs • A. Schüler
Leibniz-Institut für Wissensmedien Tübingen, Schleichstraße 6, 72076, Tübingen, Germany
e-mail: k.scheiter@iwm-tuebingen.de

B. Fillisch • R. Ploetzner
Department of Psychology, University of Education Freiburg, Kunzenweg 21, 79117, Freiburg, Germany

J. Leber • A. Renkl
University of Freiburg, Engelbergerstr. 41, 79085, Freiburg, Germany

H. Schmidt • G. Zimmermann
Hochschule der Medien, Nobelstraße 10, 70569, Stuttgart, Germany

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_9

coherence between information from the text and picture by mapping elements from the text onto those of the picture and vice versa (Seufert, 2003). This process has been referred to as global coherence formation (as opposed to the organization of information within a single representation; Seufert, 2003) or text-picture integration (Mayer, 2009; Schnotz, 2005), respectively. Only if learners mentally integrate information from texts and pictures into coherent mental models, deeper learning as a prerequisite for being able to apply the learnt contents to novel situations (i.e., transfer) is assumed to occur. A number of studies have deployed process measures such as recordings of students' eye gazes as indicators for integration. These studies showed that more switches between text and picture processing as well as more frequent processing of pictorial elements that corresponded to text are related to better comprehension of multimedia materials (e.g., Mason, Tornatora, & Pluchino, 2013; Scheiter & Eitel, 2015).

In particular, the use of eye tracking as a research methodology has, however, also revealed that many learners fail to adequately carry out the cognitive processes necessary in order to benefit from multimedia (Renkl & Scheiter, 2015). Many learners do not pay sufficient attention to the pictures but rather focus on the text (e.g., Hannus & Hyönä, 1999; Schmidt-Weigand, Kohnert, & Glowalla, 2010a, 2010b); moreover, they fail to integrate information from texts and pictures if no proper instructional guidance is provided (Cromley et al., 2013; Mason et al., 2013; Scheiter & Eitel, 2015; Schwonke, Berthold, & Renkl, 2009). In early research on multimedia learning, this problem of inadequate processing of multimedia instruction has been addressed by developing various instructional design measures that are tailored towards supporting learners in integrating texts and pictures (Mayer & Moreno, 2002). For instance, spatial contiguity between texts and pictures, which is established by presenting texts close to their corresponding picture elements, was shown to be effective in fostering the integration process (Johnson & Mayer, 2012) and comprehension (for reviews and meta-analyses see Ayres & Sweller, 2014; Ginns, 2006). Similarly, cueing or signaling correspondences between texts and pictures, for instance by showing corresponding text and picture elements in the same color, also was shown to aid integration and comprehension (e.g., Ozcelik, Karakus, Kursun, & Cagiltay, 2009; Scheiter & Eitel, 2015; for a meta-analysis see Richter, Scheiter, & Eitel, in press).

Even though these design measures improve learning outcomes, relying on them as the only approach to support learners may be problematic for both practical and theoretical reasons. First, although a large body of research exists on them, knowledge on these design measures has not yet made it into broader educational practice. That is, by and large, existing educational resources such as printed and digital text books as well as online learning environments that have been produced for commercial reasons often violate the design measures that have been identified in educational research. Thus, it is rather exceptional that learners face well-designed multimedia material based on proven design measures. Second, partly as a consequence of the prior point, learners should not become overly reliant on the design of learning material; rather, they should be able to learn even with less optimally designed instructional material and make the most out of it. Third, there is evidence that applying design measures is helpful only for those with little prior knowledge of the considered domain, whereas those with more advanced knowledge tend to not benefit or even suffer when faced with allegedly "optimized" instruction. This finding has become known as the expertise-reversal effect in instructional design research (Kalyuga, 2007). For instance, with regard to signaling it has recently been shown that low prior knowledge learners benefitted from highlighting text-picture correspondences in the material, whereas learners with either medium or high prior knowledge showed even worse performance in the signaling condition if compared to a control condition (Richter, Scheiter, & Eitel, in press). Different explanations have been provided in the literature for why harmful instructional design effects for learners with advanced prior knowledge occur. First, the design of the material may nudge these learners into processing information that is redundant to what they already know (Kalyuga, Ayres, Chandler, & Sweller, 2003). Second, some designs may suppress meaningful learning activities that would have helped advanced learners to achieve deeper comprehension. For instance, it has been shown that highly coherent texts hamper understanding of these texts for advanced learners, because there is no more need for knowledge-based inferences to overcome the incoherence of the materials (McNamara, Kintsch, Songer, & Kintsch, 1996). As a consequence, learners with more domain-prior knowledge will learn better from incoherent compared with coherent texts, whereas the reverse is true for learners with less prior knowledge. To conclude, the expertise-reversal effect suggests that a one-size-fits-all for instructional design will not guarantee effective learning for all learners.

More recently, researchers have reframed the problem of inadequate cognitive processing of multimedia material into a challenge regarding learners' lack of or erroneous self-regulation of their learning processes (e.g., Kombartzky, Ploetzner, Schlag, & Metz, 2010; Stalbovs, Scheiter, & Gerjets, 2015). This reinterpretation, which will be explained in more detail in the next section, implies that rather than relying on optimally designed instructional materials, learners should be guided towards selecting and applying appropriate cognitive processing strategies. That is, learners should be supported in becoming "good information processors" (Pressley, Borkowski, & Schneider, 1989).

A Self-Regulated Learning Perspective on Multimedia Learning

Self-regulated learning can be characterized as including metacognitive, motivational, and behavioral processes that result in the active engagement of individuals in their own learning (Azevedo, 2005; Boekaerts, 1999; Winne & Hadwin, 1998; Zimmerman & Schunk, 2001). According to Boekaerts (1999), the heterogeneity of theories of self-regulated learning can be captured by analyzing these processes as a function of three different layers: The outer layer comprises the motivational and volitional regulation of the self (i.e., choice of goals and resources). The middle layer addresses the regulation of the learning process (i.e., use of metacognitive skills to direct one's learning), and the inner layer refers to the regulation of the information processes (i.e., choice of cognitive strategies). Here we focus on the two inner layers, since we are considering instructional settings, in which resources (i.e., the information environment) are given and instructional goals are predefined.

According to models of self-regulated learning, an inadequate cognitive processing of multimedia materials can be seen as a failure to regulate one's learning at the meta-cognitive and information-processing levels. Such a failure can occur for different reasons. First, learners may be unable to judge what they know and what they do not know, thereby failing to monitor their understanding of a domain (Bjork, Dunlosky, & Kornell, 2013). Inaccurate comprehension monitoring can result in either over- or underconfidence regarding one's level of understanding, with both biases resulting in different problems for a student's regulation of learning behavior. Overconfidence in one's knowledge may cause learners to terminate studying prematurely, whereas underconfidence will result in learners investing time in studying materials already well understood. Hence, both biases result in an inadequate allocation of study time (Son & Metcalfe, 2000). Recent research has shown that the use of multimedia materials (compared with text-only instruction) increases the likelihood of learners becoming overconfident in their knowledge (multimedia heuristic, Eitel, 2016; Serra & Dunlosky, 2010). Failure to regulate one's learning may also be caused by a student's lack of knowledge regarding the question of how to respond to, for instance, gaps in their understanding. That is, even when correctly detecting gaps, students may not know how to overcome these gaps. Veenman, Van Hout-Wolters, and Afflerbach (2006) have conceptualized this problem that occurs when students lack strategy knowledge. They suggest that ideally learners should know what to do (declarative strategy knowledge), when and why to do it (conditional strategy knowledge), and how to do it (procedural strategy knowledge).

Applying a self-regulated learning perspective to multimedia learning suggests that learners need to be supported in assessing what they (do not) know (monitoring support) and in regulating their learning behavior in a way that matches their current understanding. So far, support measures have focused on verbal or visual instructions that convey strategy knowledge and make its use in a given learning situation more likely (for an overview see Renkl & Scheiter, 2015). For instance, a number of studies have used prompts or prompt-like instructions that tell students to apply certain cognitive processes such as information integration (e.g., Bartholomé & Bromme, 2009; Kombartzky et al., 2010; Schlag & Ploetzner, 2010; Stalbovs et al., 2015). Moreover, visual instructions have been used where learners were shown eye movements that illustrated helpful visual behavior in advance to learning from multimedia materials (e.g., Mason, Pluchino, & Tornatora, 2015; Skuballa, Fortunski, & Renkl, 2015). These support measures have in common that they can be deployed irrespective of whether well-designed instructional materials are available; moreover, they can help students to become independent learners who are able to control their learning without having to rely on high-quality instructions. However, their effectiveness depends on a number of possible boundary conditions that are not yet fully known (Renkl & Scheiter, 2015). Moreover, they are deployed in a one-size-fits-all fashion to all learners regardless of whether they are already able to self-regulate their learning or not. Furthermore, not only for instructional design measures but also for self-regulation interventions, expertise-reversal effects (Kalyuga, 2007) have been revealed. For example, Nückles, Hübner, Dümer, and Renkl (2010) found that psychology students who self-regulated their learning during journal writing initially benefited from being supported by prompts that activated elaborative, organizational, and metacognitive learning strategies. When such prompting was continued over the course of a semester, however, it had detrimental effects on the motivational as well as on the cognitive level in the second half of the semester. Adaptive forms of support, which will be introduced next, offer a potential solution to this problem.

Adaptive (Multimedia) Learning Environments

Adaptivity is present when the instruction automatically changes in response to the learners' states and learning behaviors (Akbulut & Cardak, 2012; Park & Lee, 2004). Adaptive learning environments or response-sensitive systems (Park & Lee, 2004) require that relevant learner states and learning behaviors are assessed and evaluated online. Thus, the system takes over some of the monitoring that is required from learners during self-regulated study. Then the system will react towards the results of this diagnosis, thereby supporting (or even replacing) learners' regulation of the learning process. Adaptive learning environments respond to a learner's state and behavior on a moment-to-moment basis rather than to the results of a one-time assessment prior to learning. Their responses can take many forms: For instance, the system could reduce or increase the difficulty of the learning task, offer prompts that tell students how to proceed, or choose a different format of instruction (e.g., more or less elaborate explanations, another multimedia design variant). From a self-regulated learning perspective it is important to disinguish two different forms of adaptivity. Assistive adaptivity occurs when the system suggests to a learner how to proceed, while the learner maintains control over whether and how to follow the suggestion. For instance, a learner could decide to (not) follow the system's advice of restudying a previously encountered unit of instruction. Directive adaptivity, on the other hand, constrains learner control to a much stronger extent by offering a further choice to the learner. For instance, automatically displaying the previously encountered unit of instruction leaves no option to the learner other than to restudy it. Thus, whereas assistive adaptivity only scaffolds the process of self-regulated learning, directive adaptitivy imposes external control of the learning process. In both cases, the adaptivity mechanism presupposes that there is an unambigious mapping between a diagnosis and a system response (e.g., "if test item X is not answerred correctly, then display unit X again for restudy").

Adaptive learner environments differ in whether they select responses relative to single learner states or behaviors or relative to a learner model. Such a learning model captures all relevant states and behaviors and is continuously updated while the learner proceeds with the learning task. For adaptive systems relying on learner modeling, Shute and Zapata-Rivera (2008) have proposed an adaptive cycle that includes four components, namely capturing, analyzing, selecting, and presenting. First, the system captures data about a learner who is interacting with the system. These data are the foundation for the learner model, which is generated later in the course. Data collection is ongoing during the whole interaction process and aims at updating the learner model. Next, the captured data are analyzed in order to directly create a model of the learner according to content-specific information presented in the system's learning environment. Based on the learner model, the system can determine whether an intervention is necessary and, furthermore, identify the kind of intervention that should be presented to the learner. Thus, the third process refers to the selection of a system response such as a hint, a prompt, or an explanation. The main purpose of the system is to determine what kind of system response or information is appropriate. Although decision rules can be predefined, they continually need to be updated during the interaction between the learner and the system. Finally, the last component corresponds to the presentation of the selected adaptive intervention. Although such an adaptive cycle is linear at the beginning, recurrences and returns between the components become inevitable. The first cycles generate a rather coarse learner model, which becomes refined over time. Thus, the learner model is not static because new learning traces can be used to update, revise, or verify the learner model. Creating adaptive systems based on a comprehensive learner modeling can be very challenging, because it requires extensive a priori knowledge regarding an effective mapping between the various combinations of learner states/behaviors and the adequate responses of the system.

Moreover, adaptive systems differ in whether they rely on a diagnosis of a learner's processing behavior or on his/her current state of knowledge, skill, or motivation (e.g., disengagement).

Finally, adaptive systems can be distinguished according to the way that they gather information required for choosing an appropriate response from the system's repertoire. Relevant information can be gained explicitly by asking the learner to answer a questionnaire or test or implicitly by drawing inferences from the way how the learner interacts with the system. Explicit assessment methods provide—if well designed—a valid judgment of the current state of affairs. However, in particular the use of long assessments may disrupt the learning process; moreover, working on such assessments repeatedly can be demotivating for learners. To counteract these problems, rapid assessment tasks (RATs) have been suggested (Kalyuga, 2008) as an alternative to more comprehensive tests. RATs are short assignments that are interspersed into the learning material and that aim at assessing the current cognitive state of the learner. Verification tasks are a special version of RATs. They ask the learner to judge whether statements referring to the previously learnt content are correct or incorrect. This technique has been proven to be efficient and non-reactive in past research (Renkl, Skuballa, Schwonke, Harr, & Leber, 2015). Based on the results of such verification tasks the further path of the learning experience can be adjusted to the learners' needs: learners who provide correct answers may continue with their learning paths, while learners with insufficient knowledge receive specific instructions and interventions.

In contrast to explicit assessment methods, implicit assessment methods are characterized by their unobstrusiveness because data collection may take place even without the learner realizing it (Barab, Bowdish, Young, & Owen, 1996). However, it may be rather difficult to infer and interpret a learner's knowledge, and learning merely from his/her interactions with the learning environment. For instance, adaptive hypermedia systems (Brusilovsky, 2001), in which students' navigation behavior was used to inform the system's response, were not very successful in providing information concerning the deployment of information utilization strategies. This is because navigational behavior (e.g., selecting a link, browsing through a section) operates at a rather coarse level where it often remains unclear which navigation behavior corresponds to a certain cognitive or affective process. Moreover, rather than looking at a single click on a link, it is often necessary to analyze longer navigational sequences to derive meaningful behavioral patterns, making unambigious data interpretation rather difficult.

More recently there have been attempts to use more fine-grained assessment methods that are assumed to be more closely linked to the actual learning behavior and learner states (cf. Azevedo et al., 2017; Spüler et al., 2017; Winne et al., 2017). In the present chapter we focus on the use of eye tracking as a way of assessing learning behavior as well as inferring learners' cognitive (and motivational) states from it (cf. Conati & Merten, 2007; D'Mello, Olney, Williams, & Hays, 2012; Roda & Thomas, 2006; Toet, 2006). Eye tracking provides information on where a person is looking, for how long, and in which order. According to Just and Carpenter (1980), this information can be taken as an indication of a person's processing of information at the cognitive level (eye-mind hypothesis). In particular, it is assumed that elements that are fixated will be processed in the mind without any considerable delay (immediacy assumption). The duration of a fixation (i.e., the time when the eye is positioned on one spot and when information intake occurs) thus can be interpreted as the intensity with which some information element is processed. Longer fixation times can thus indicate more interest from a learner or more relevance of that information for the learning task. Saccades (i.e., rapid eye movements in between fixations during which information intake is suppressed) are indicative of the order of information processing. In multimedia research, saccades between text and pictures are seen as indicators for the process of text-picture integration (Johnson & Mayer, 2012; Scheiter & Eitel, 2016).

Conati and Merten (2007) investigated the usefulness of gaze data and compared three different probabilistic models describing a learner's behavior. They demonstrated that the inclusion of on-line eye movement data improved sensitivity and specificity in predicting when learners were implicitly self-explaining learning content. Gaze data hence helped to model the mental state of a learner more accurately. In addition, gaze data can also provide information about the motivational states of learners such as boredom and disengagement (D'Mello et al., 2012). Eye tracking has also yielded important insights into how students learn from multimedia materials (cf. Scheiter & Eitel, 2016; Scheiter & van Gog, 2009; van Gog & Scheiter, 2010). Thus, this method seems highly suitable as a diagnostic instrument that could also be used for the design of adaptive learning environments.

Until recently, the use of eye tracking was both cumbersome and expensive; however, this has changed to some extent with the development of customized systems. Importantly, at the moment these systems provide relatively easy-to-use methods for analyzing eye-tracking data offline after the learning took place. However, online analysis methods have not yet been implemented in the software packages. Therefore, one important aspect of the development of the multimedia learning system described in the next section was to allow for an online analysis of the learners' eye movements as a prerequisite for incorporating adaptivity.

The Adaptable and Adaptive Multimedia System (AAMMS)

The Adaptable and Adaptive Multimedia System (AAMMS) is a multimedia learning environment that was developed in an interdisciplinary project with researchers from computer science, psychology, and educational technology. It is based on the ILIAS 4 Open Source Framework (www.ilias.de) and uses the infrastructure of an established ILIAS installation to present multimedia content as well as instructional interventions such as prompts. The AAMMS allows for different modes of regulation: On the one hand, learners can adapt the instruction to their own information needs by choosing which content should be displayed in which format and which types of support they would like to receive (adaptable mode). On the other hand, the AAMS offers an Adaptive Learning Module (ALM) that allows for implementing adaptive instruction based on rapid assessments and eye tracking (adaptive mode). Both modes can also be switched off, in which case the multimedia content is displayed as pre-determined by an instructor or researcher (fixed mode).

The User Interface

Figure 9.1 shows the user interface with its five main areas. The navigation tree (1) provides access to the learning units by means of a hierarchical menu. The navigation bar (2) offers page-by-page browsing. Each learning unit is made up of several representations such as written and spoken texts, schematic and realistic images, videos and animations. In adaptable learning scenarios, where learners can decide upon the learning content themselves, different representation formats can be accessed via the media shelf (3). The content area (4) displays the representations

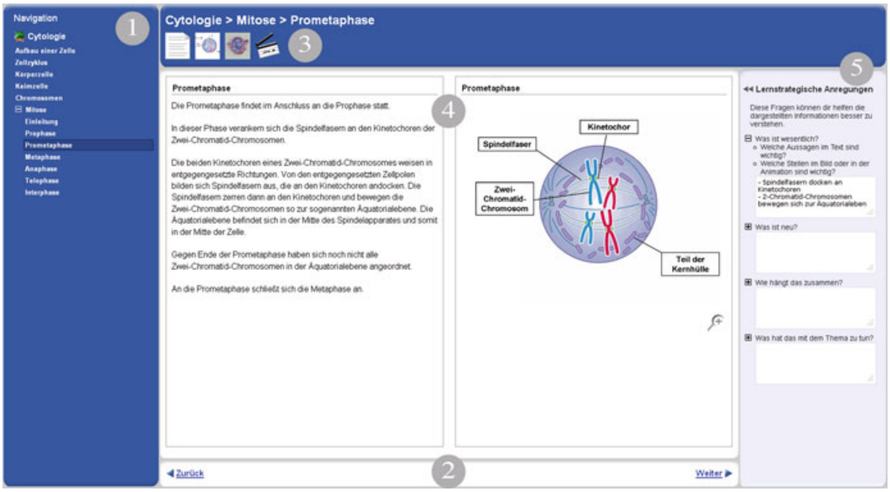

Fig. 9.1 The user interface of the Adaptable and Adaptive Multimedia System (AAMMS)

that were selected by a learner (adaptable mode), selected by the system (adaptive mode), or pre-determined by an instructor or researcher (fixed mode). The support area (5) can be used to offer prompts or other interventions to foster specific cognitive learning activities.

When learning in the adaptable mode of the AAMMS, users can either follow a linear pathway through the learning units or they can use the navigation tree to navigate freely in accordance with their individual learning goals. For each learning unit, the system suggests an initial combination of representations and displays them in the content area. The learners may customize this combination by dragging their preferred representations from the media shelf into the content area. A preview of each representation is displayed in the media shelf to assist in the selection process. The learners' combinations of representations are recorded by the system. If the learners re-visit a learning unit, the stored combinations are displayed again.

In each learning unit prompts can be presented to the learners that ask them to engage more deeply with the instructional materials (cf. Ruf & Ploetzner, 2014). They are stated as questions, for example: (a) Which information is essential? (b) Which relations can be identified? (c) How is the information related to the overall topic? Learners who need additional information can obtain more specific questions by clicking on the questions. Each question has a textbox below it to take notes. The notes are automatically stored so that learners can review or change them when they re-visit the learning unit.

Moreover, it is possible to employ different types of assessment during learning. Before the learners start a new learning unit, they may be asked questions about the unit just completed. These questions can be used either by the learners to self-assess and monitor their current level of understanding or by the system to trigger adaptations of the learning material. Moreover, in conjunction with an eye tracker, the ALM included in the AAMMS allows analyzing students' eye movements online and adapting the instruction based on this analysis.

System Architecture of the Adaptive Learning Module (ALM)

The ALM offers predefined adaptivity functions that authors can implement in their learning environments. These adaptivity functions allow adaptive behavior of the learning platform in response to a user's behavioral data (Schmidt, Wassermann, & Zimmermann, 2014). The adaptivity functions are mainly based on learner data from two sources: a learner's eye movements and answers to rapid assessment verification tasks.

The eye-tracking application of ALM connects the eye-tracking hardware and the web-based learning environment (cf. Wassermann, Hardt, & Zimmermann, 2012). In particular, the ALM allows receiving, analyzing, and responding to eye-tracking data via a web-socket interface in real-time. The application's adaptivity is based on the capture of gaze fixations on pre-defined areas of interest (AOIs). When a user fixates an AOI for a predefined time, a fixation is registered and recorded. The system allows counting the number of fixations, their overall duration (dwell time), and the frequency of transitions between AOIs, that is, when learners move their visual attention from one AOI (e.g., a text) to another (e.g., a picture) and vice versa. The data are recorded for each learning unit separately and are analyzed once a learner indicates that he or she wants to proceed to the next learning unit. For each learning unit, threshold values can be defined that need to be reached in order to be considered as adequate learning behavior. If the learning behavior remains below the threshold values, a system response is generated and the learner is prevented from proceeding to the next learning unit. For instance, when the ALM registers a too short reading time for a text in a given learning unit, it can prompt the learner to re-read the text or the text can be highlighted to nudge the learner into rereading it (for details on the adaptivity of ALM and its technical implementation, see Schmidt et al., 2014).

Adaptivity functions can also be based on a learner's performance in interspersed rapid assessment tasks (e.g., Kalyuga, 2008; Renkl et al., 2015). ALM supports overlay prompts on the screen for asking, for example, multiple-choice questions. This ALM function can be used to easily integrate rapid assessment tasks into the learning environment. When learners answer a rapid assessment task, the system assesses the predefined answers and reacts to it. For example, the system can limit the learner's ability to navigate through the learning content by temporarily disabling the "continue" button and asking the learner to re-read highlighted areas of the learning material again. The system can also present other learning aids such as prompts asking the learners to rethink particular aspects of the learning content or to write down a self-explanation.

Empirical Evidence

Two sets of studies have been carried out evaluating the effectiveness of the ALM for supporting learning from multimedia. In the first set of studies, we tested ways to best close knowledge gaps that remain during learning with multimedia learning materials. In the second set of studies, the ALM was used to detect potential learning problems in real-time using eye tracking and to change the design of the instruction adaptively. Both sets of studies will be sketched next.

Closing Knowledge Gaps During Learning with Multimedia

This section exemplarily presents two studies on how to best close learners' knowledge gaps by an adaptive procedure based on rapid assessment tasks. An initial study addressed the question of whether eye-tracking indicators can be used to reduce the number of presented rapid assessment tasks and, thereby, learning time without the drawback of overlooking many of the learners' knowledge gaps. The second study evaluated variants of restudy prompts that can be presented when a learner fails to correctly answer a rapid assessment task. Hence, this second study implemented assistive adaptivity.

In the first study (Skuballa, Leber, Schmidt, Zimmermann, & Renkl, 2016; $N = 60$ university students), we tested whether eye-tracking data can be used to select and, thereby, reduce the number of rapid assessment tasks without the disadvantage of missing potential knowledge gaps. More specifically, we compared two conditions: (a) In the full-presentation condition, we provided all available rapid assessment tasks; and (b) in the adaptive-presentation condition, we provided a rapid assessment task whenever the eye-tracking data hinted towards a potential knowledge gap (e.g., very short dwell time at a certain picture). We found that the selection of rapid assessment tasks increased their hit rates (i.e., enhanced diagnostic sensitivity) in that these tasks were answered more often incorrectly as compared to tasks in the full-presentation condition (43% vs. 32%). The adaptive presentation also reduced the learning time by about 17% without compromising learning outcomes; the latter were comparable in both conditions. Overall, our findings suggest presenting rapid assessment tasks based on eye-tracking indicators hinting towards potential knowledge gaps. The main advantage is that the learners need less study time.

Beyond the question about which rapid assessment tasks should be presented, it is an open question which type of restudy prompt is best provided in the case of a knowledge gap. A prompt may encourage a learner to close the very specific knowledge gap that has been detected by a rapid assessment task (e.g., the fact that it is the nucleus where DNA doubles during mitosis). Such a specific prompt might be a parsimonious intervention but it may fail to address a potentially "bigger problem": The learner may miss not merely a specific piece of knowledge, but his or her knowledge representation may be incomplete with respect to a broader sub-area of the learning contents (e.g., what happens in general in the nucleus during mitosis). In the latter case, a prompt that encourages not just looking up the specific missing knowledge piece, but considering the "field" of related knowledge pieces as well might have broader effects on learning outcomes. Such broader and more unspecific prompts, however, have the potential disadvantage of being less efficient than specific prompts when just such a specific piece of knowledge is missing. Furthermore, unspecific prompts can induce redundant (i.e., unnecessary) processing of already understood materials (cf. the redundancy effect; Sweller, Ayres, & Kalyuga, 2011).

Renkl, Skuballa, Schwonke, Harr, and Leber (2015; Exp. 2) compared the effects of specific and unspecific restudy prompts (i.e., focusing on a very specific piece of knowledge or the "field" of related knowledge pieces; $N = 41$ university students). In the specific prompts condition, the relevant text passages were highlighted by darkening the less relevant information on the page. The prompt requested learners to restudy the relevant passage in order to solve the task correctly, and the task was repeated (Fig. 9.2). In the unspecific prompts condition, the learners were asked to restudy and figure out both the direct answer to the question and to explore the broader context (Fig. 9.3).

We assumed that specific prompts were superior in repairing the specific knowledge gaps identified by the rapid assessment tasks and in acquiring knowledge about the central issues of the mitotic process (as these issues were covered by the rapid assessment tasks). We expected that unspecific prompts were more effective

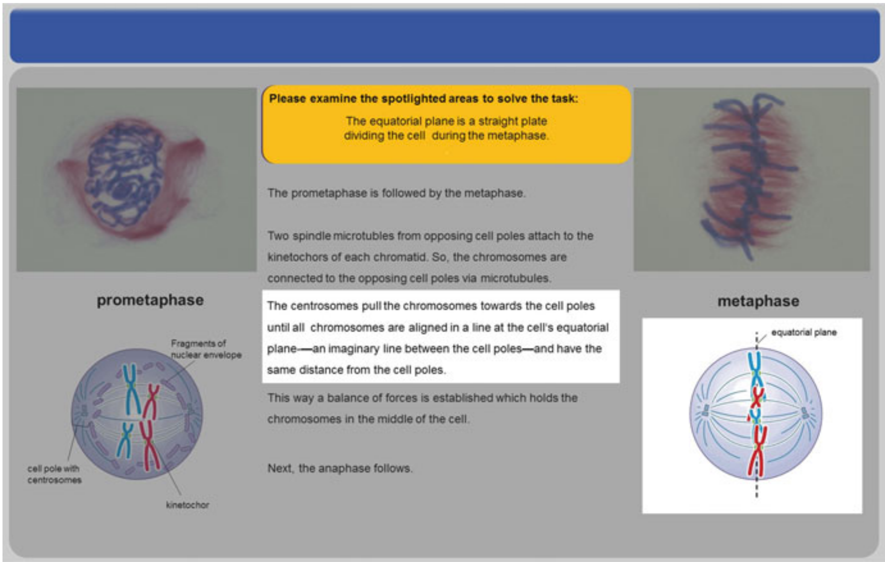

Fig. 9.2 Screenshot of a specific prompt (taken from Renkl et al., 2015)

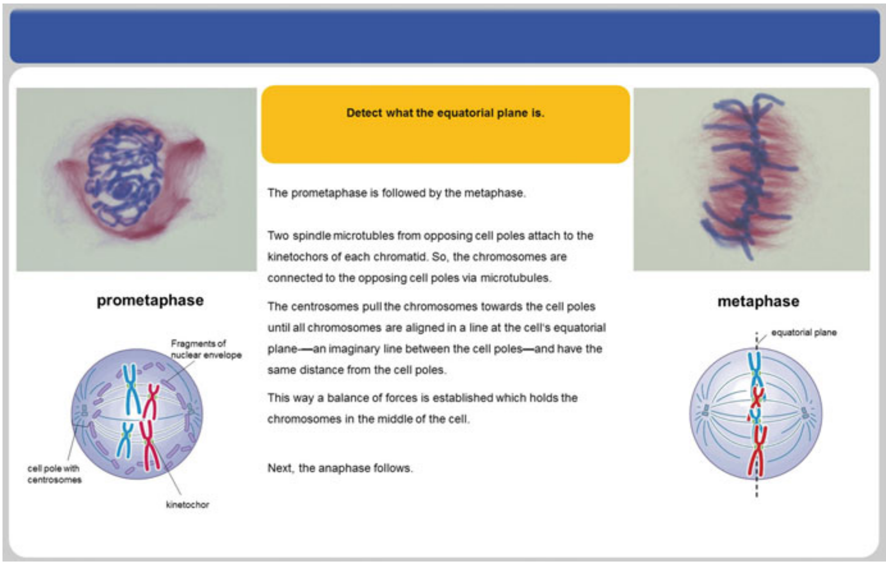

Fig. 9.3 Screenshot of an unspecific prompt (taken from Renkl et al., 2015)

in fostering knowledge about more general issues related to mitosis. The results showed that both types of prompts repaired the specific knowledge gaps in most cases (in over 80% of the cases). Moreover, we found a general superiority of unspecific prompts, thereby suggesting that knowledge gaps should be closed by unspecific restudy prompts.

Overall, adaptive systems based on a rapid-assessment procedure should provide rapid assessment tasks only if eye-tracking indicators hint towards a potential knowledge gap. If the learners cannot answer correctly a rapid assessment task, they should receive prompts that ask them to restudy the corresponding field of related knowledge pieces.

Adapting the Multimedia Design in Response to Learners' Eye Movements

Previous research has shown that some learners have difficulties in adequately using effective cognitive processes like selection, organization, and integration while processing multimedia materials (e.g., Mason et al., 2013). To support learners by providing them with personalized, just-in-time instructional support, the ALM was used to monitor, analyze, and modify the learners' individual processing behavior online based on the learners' eye movements in two studies described below. In contrast to the aforementioned set of studies, the adaptivity mechanism relied on directive adaptivity in that the system's response led to a change in the instructional format of the material, which learners were forced to make use of.

In order to tailor the ALM to the population under study, Schubert et al. (n.d.) first determined threshold values in a pre-study. In this study ($N = 32$ students), patterns of eye movements were identified that were suited to distinguish successful versus less successful learners in a non-adaptive multimedia learning session on mitosis. Results showed that successful learners had longer fixations times and higher fixation counts on text and pictures as well as more transitions between text and pictures than less successful learners. These findings were then used to implement the gaze-based adaptive system that analyzed learners' eye movements during learning and altered the presentation of the materials according to learners' viewing behavior. Whenever learners showed a viewing behavior similar to that of the unsuccessful learner group in the pre-study (i.e., too short fixation times on either text or pictures or too few text-picture transitions), the system presented the same content in an instructional design that should prompt adequate processing. In particular, whenever either text or picture on a given page were processed for too short a time (i.e., below the threshold values derived from the first study), the text or the picture was enlarged, thereby covering most of the screen (Fig. 9.4, left panel). In case of too few text-picture transitions, the page design was altered in that corresponding elements from the text and the picture were then highlighted using the same colors (Fig. 9.4, right panel), thereby signaling the conceptual relations between the verbal and pictorial information (cf. Richter et al., 2016, for the effectiveness of multimedia integration signals). The system enforced processing of the redesigned multimedia materials in that these were presented for a fixed amount of time before the students were allowed to proceed to the next page of the learning materials. This page was again presented in the standard layout without enlargements of color coding.

The first study with the adaptive system ($N = 79$ students) investigated whether the adaptive multimedia learning system would have any beneficial effects on learning compared to a non-adaptive, fixed presentation of the same materials.

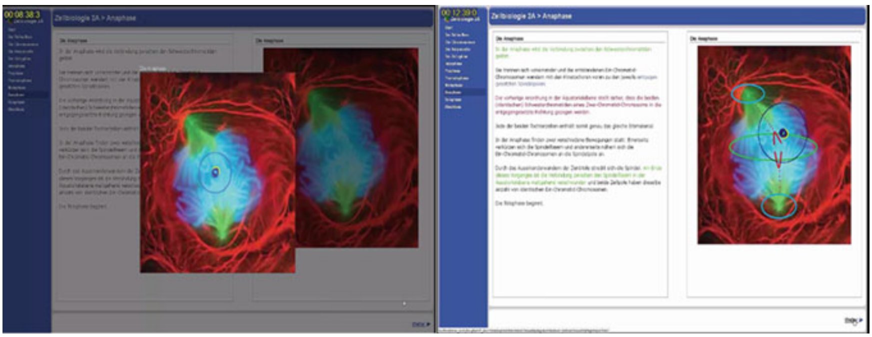

Fig. 9.4 Exemplary pages of the adaption: Zooming-out of the picture (*left panel*) and presentation of the color-coded version (*right panel*)

Students learned with either the adaptive or the non-adaptive multimedia instruction about mitosis while their eye movements were recorded. After learning, their recall and comprehension of the materials was assessed. As thresholds, we used the mean fixation times on either text or picture and the mean number of transitions of the group of non-successful learners identified in the pre-study and added one standard deviation to it. Results showed that irrespective of learners' prior knowledge, the gaze-based adaptive system had no effect on the effectiveness of multimedia learning. A possible reason for these findings was that the thresholds might have been set too high so that successful learners were also falsely identified as poor learners.

The aim of second study ($N = 58$ students) conducted with the gaze-based adaptive system was hence to improve the adaptivity mechanism of the system. To this end, we adjusted the threshold values by choosing the mean values of the group of unsuccessful learners from the pre-study (rather than adding one standard deviation to it). This way we expected only learners with inadequate processing behavior to receive personalized instructional support. Again, students either learned with the adaptive or the non-adaptive multimedia instruction about mitosis. Results showed no effects for recall performance. For comprehension, there was a significant interaction between experimental condition and students' prior knowledge: stronger students scored marginally higher with than without adaptive instructional support, whereas weaker students scored significantly worse with adaptive instructional support. These results can be interpreted in at least two ways: First it may have been the case that the adjusted thresholds were now too restrictive so that learners with inadequate processing behavior were falsely identified as successful learners. Second, there might have been too little adaptive instructional support especially for learners with lower prior knowledge. Further studies are required that look more closely at both the assessment as well as the support component of the adaptive multimedia learning system, before conclusions regarding the effectiveness of adjusting instruction to a learner's gaze behavior can be drawn.

Conclusions

Adaptive learning environments are supposed to enhance learning by providing personalized support for every individual student. The multimedia information environment described in this chapter shows how challenging the design of such a system can be. From a learning sciences perspective, the challenges pertain to at least three aspects that match well onto the description of adaptive systems provided by Shute and Zapata-Rivera (2008).

The first challenge is to decide what constitutes a learner state that requires adaptive instruction and how this state should be diagnosed, thereby pertaining to the *assessment* component of adaptive systems. In the sample studies sketched in this chapter, two diagnostic approaches have been implemented: rapid verification tasks and eye movements. Rapid verification tasks address students' current knowledge in an explicit fashion, hence their interpretation is rather straightforward. Eye movements, on the other hand, are more ambiguous in this respect. For instance, the relation between a given eye movement behavior and learning outcomes appears to vary across different material and learners (cf. Scheiter & Eitel, 2016; Schwonke et al., 2009), requiring extensive pretesting to calibrate these relations before an adaptive system can be deployed. Moreover, eye movements are not just indications of students' cognitive processes while learning, but they may also point towards students' motivational states such as boredom and disengagement (D'Mello et al., 2012), thereby further complicating the state of affairs. Against this background, Skuballa et al. (2015) used an adaptation procedure in which eye-tracking data and rapid assessment tasks were combined. First evidence suggested that such a combination is promising.

Despite these difficulties, there is a current trend of considering more rather than fewer physiological indicators as possible candidates for assessing learner states as a prerequisite for adapting instruction—including emotional expressions, skin responses, and brain activity parameters (see also Azevedo et al., 2017; Spüler et al., 2017). The promise of this multivariate approach is that triangulation of these different data sources will allow better disambiguation of a student's current state of learning. Moreover, these advanced learning technologies often rely on more complex methods of data analyses such as machine learning algorithms to determine patterns of parameters related to successful learning (see also Azevedo et al., 2017; Spüler et al., 2017; Winne et al., 2017). Using these methods thus allows describing learner behavior and states in a multidimensional space that comprises emotional, cognitive, and metacognitive components. Moreover, there is increasing interest in learning analytics based on large data sets to account for the complexity of (self-regulated) learning that results from interindividual as well as intraindividual differences in learning processes (Winne et al., 2017).

The second challenge pertains to the question of which instructional support measures promote learning best, thereby addressing the *response* component of adaptive systems. While instructional design research has made enormous progress in determining how instruction should be delivered in order to be effective, there still remains quite a bit of ambiguity in this regard. Instructional design variants do not prove effective in all situations; rather, their effectiveness seems to be tied to certain boundary conditions that are not yet fully understood (cf. Renkl & Scheiter, 2015). For instance, the effectiveness of prompts appears to depend on their focus (specific vs. general), whether they require externalizing of knowledge, or reproduction of knowledge versus generating new information via inferencing, to name just a few dimensions.

The third challenge results from the first two and refers to the question of matching a learner state with the most adequate system response. Thus, it addresses the *rules that should be used to link assessment and response* to each other. It is yet unclear whether we will ever be able to come up even with heuristics telling instructional designers which variant of an instructional material will be most suited for which learner. What has become very evident in the past years of educational research is that a student's level of prior knowledge plays an important role in this regard. That is, for many multimedia designs it has been shown that design variants that clearly improve learning for students with less prior knowledge have no or even detrimental effects for students with high prior knowledge (cf. expertise reversal effect, Kalyuga et al., 2003; Kalyuga & Renkl, 2010). Hence, this research suggests that adaptivity mechanisms need to consider a person's prior knowledge. However, it is still an open question whether there are other variables that have similar effects on the effectiveness of different instructional designs.

In the present chapter, we described different versions of a fully adaptive system that diagnoses a learner's behavior and responds accordingly. Based on the results of the two sets of studies reported, it looks as if assistive adaptivity is more promising than directive adaptivity. However, even though we used similar learning materials and assessments, there are still a number of differences between the sets of studies beyond the adaptivity mechanism. Most importantly, the studies also differ in the assessments (eye tracking plus rapid assessment tasks to diagnose a learner's knowledge state vs. eye tracking to diagnose learning behavior) based on which a response was given. Thus, further studies are needed that systematically address the question of how much learner control should/can be offered during regulation.

Irrespective of the adaptivity mechanism that had been implemented, we faced all three of the aforementioned challenges in the design of our adaptive system. An alternative to this approach is offered by systems that provide an assessment a learner's current state and/or his/her learning processes, which is then communicated to the learner (e.g., via a visualization of his/her state that is described based on multichannel data, see Azevedo et al., 2017; nStudy, Winne et al., 2017). Importantly, these systems then leave regulation to the learner, thereby implementing the extreme of assistive adaptivity. They thus rely on the implicit assumption that while learners may experience difficulties in accurately monitoring their current state, they are very capable of regulating their learning behavior relative to their states. These systems thereby circumvent the second and third challenge mentioned above. That is, they do not require any insight into possible system responses and how these match with learner states.

References

Akbulut, Y., & Cardak, C. S. (2012). Adaptive educational hypermedia accommodating learning styles: A content analysis of publications from 2000 to 2011. *Computers & Education*, *58*, 835–842. doi:10.1016/j.compedu.2011.10.008

Anglin, G. J., Vaez, H., & Cunningham, K. L. (2004). Visual representation and learning: The role of static and animated graphics. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd ed., pp. 865–916). Mahwah, NJ: Erlbaum.

Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 206–226). New York, NY: Cambridge University Press.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, *40*, 199–209.

Azevedo, R., Millar, G. C., Taub, M., Mudrick, N. V., Bradbury, A. E., & Price, M. J. (2017). Using data visualizations to foster emotion regulation during self-regulated learning with advanced learning technologies. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 225–247). New York, NY: Springer.

Barab, S. A., Bowdish, B. E., Young, M. F., & Owen, S. V. (1996). Understanding kiosk navigation: Using log files to capture hypermedia searches. *Instructional Science*, *24*, 377–395.

Bartholomé, T., & Bromme, R. (2009). Coherence formation when learning from text and pictures: What kind of support for whom? *Journal of Educational Psychology*, *101*, 282–293. doi:10.1037/a0014312

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444. 10.1146/annurev-psych-113011-143823

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, *31*, 445–457.

Brusilovsky, P. (2001). Adaptive hypermedia. *User Modeling and User-Adapted Interaction*, *11*, 87–110.

Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 174–205). New York, NY: Cambridge University Press.

Conati, C., & Merten, C. (2007). Eye-tracking for user modeling in exploratory learning environments: An empirical evaluation. *Knowledge-Based Systems*, *20*, 557–574. doi:10.1016/j.knosys.2007.04.010

Cromley, J. G., Bergey, B. W., Fitzhugh, S. L., Newcombe, N., Wills, T. W., Shipley, T. F., & Tanaka, J. C. (2013). Effectiveness of student-constructed diagrams and self-explanation instruction. *Learning & Instruction*, *26*, 45–58. doi:10.1016/j.learninstruc.2013.01.003

D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies*, *70*, 377–398. doi:10.1016/j.ijhcs.2012.01.004

Eitel, A. (2016). How repeated studying and testing affects multimedia learning: Evidence for adaptation to task demands. *Learning and Instruction*, *41*, 70–84. doi:10.1016/j.learninstruc.2015.10.003

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning & Instruction*, *16*, 511–525. doi:10.1016/j.learninstruc.2006.10.001

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low-and high-ability children. *Contemporary Educational Psychology*, *24*, 95–123. doi:10.1006/ceps.1998.0987

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology. Applied*, *18*, 178–191. doi:10.1037/a0026923

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–355.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, *19*, 509–539. doi:10.1007/s10648-007-9054-3

Kalyuga, S. (2008). When less is more in cognitive diagnosis: A rapid online method for diagnosing learner task-specific expertise. *Journal of Educational Psychology*, *100*, 603–612. doi:10.1037/0022-0663.100.3.603

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*, 23–31. doi:10.1207/s15326985ep3801_4

Kalyuga, S., & Renkl, A. (2010). Expertise reversal effect and its instructional implications: Introduction to the special issue. *Instructional Science*, *38*, 209–215. doi:10.1007/s11251-009-9102-0

Kombartzky, U., Ploetzner, R., Schlag, S., & Metz, B. (2010). Developing and evaluating a strategy for learning from animations. *Learning & Instruction*, *20*, 424–433. doi:10.1016/j.learninstruc.2009.05.002

Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology*, *41*, 172–187. doi:10.1016/j.cedpsych.2015.01.004

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education*, *60*, 95–109. doi:10.1016/j.compedu.2012.07.011

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.

Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*, 107–120.

McNamara, D., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*, 1–43. doi:10.1207/s1532690xci1401_1

Nückles, M., Hübner, S., Dümer, S., & Renkl, A. (2010). Expertise reversal effects in writing-to-learn. *Instructional Science*, *38*, 237–258. doi:10.1007/s11251-009-9106-9

Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. *Computers & Education*, *53*, 445–453. doi:10.1016/j.compedu.2009.03.002

Park, O.-C., & Lee, J. (2004). Adaptive instructional systems. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 651–684). Mahwah, NJ: Erlbaum.

Pressley, M., Borkowski, J. G., & Schneider, W. (1989). Good information processing: What it is and how education can promote it. *International Journal of Educational Research*, *13*, 857–867.

Renkl, A., & Scheiter, K. (2015). Studying visual displays: How to instructionally support learning. *Educational Psychology Review*, 1–23. doi:10.1007/s10648-015-9340-4

Renkl, A., Skuballa, I. T., Schwonke, R., Harr, N., & Leber, J. (2015). The effects of rapid assessments and adaptive restudy prompts in multimedia learning. *Educational Technology & Society*, *18*, 185–199.

Richter, J., Scheiter, K., & Eitel, A. (in press). Signaling text–picture relations in multimedia learning: The influence of prior knowledge. *Journal of Educational Psychology.* https://doi.org/10.1037/edu0000220

Roda, C., & Thomas, J. (2006). Attention aware systems: Theories, applications, and research agenda. *Computers in Human Behavior*, *22*, 557–587. doi:10.1016/j.chb.2005.12.005

Ruf, T., & Ploetzner, R. (2014). One click is too far! How the presentation of cognitive learning aids influences their use in multimedia learning environments. *Computers in Human Behavior*, *38*, 229–239. doi:10.1016/j.chb.2014.06.002

Scheiter, K., & Van Gog, T. (2009). Using eye tracking in applied research to study and stimulate the processing of information from multi-representational sources. *Applied Cognitive Psychology*, *23*, 1209–1214. https://doi.org/10.1002/acp.1524

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction*, *36*, 11–26. doi:10.1016/j.learninstruc.2014.11.002

Scheiter, K., & Eitel, A. (2016). The use of eye tracking as a research and instructional tool in multimedia learning. In C. Was, F. Sansosti, & B. Morris (Eds.), *Eye-tracking technology applications in educational research* (pp. 143–164). Hershey, PA: IGI Global.

Schlag, S., & Ploetzner, R. (2010). Supporting learning from illustrated texts: Conceptualizing and evaluating a learning strategy. *Instructional Science*, *39*, 921–937. doi:10.1007/s11251-010-9160-3

Schmidt, H., Wassermann, B., & Zimmermann, G. (2014). An adaptive and adaptable learning platform with real- time eye-tracking support: Lessons learned. In S. Trahash, R. Ploetzner, G.

Schneider, C. Gayer, D. Sassiat, & N. Wöhrle (Eds.), *Tagungsband DeLFI 2014* (pp. 241–252). Bonn, Germany: Köölen Druck & Verlag GmbH.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010a). A closer look at split visual attention in system-and self-paced instruction in multimedia learning. *Learning and Instruction*, *20*, 100–110. doi:10.1016/j.learninstruc.2009.02.011

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010b). Explaining the modality and contiguity effects: New insights from investigating students' viewing behavior. *Applied Cognitive Psychology*, *24*, 226–237. doi:10.1002/acp.1554

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 49–69). Cambridge: Cambridge University Press.

Schubert, C., Scheiter, K., Schüler, A., Schmidt, H., Zimmermann, G., Wassermann, B., … Eder, T. (n.d.). Adaptive multimedia: Using gaze-contingent instructional guidance to provide personalized processing support.

Schwonke, R., Berthold, K., & Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. *Applied Cognitive Psychology*, *23*, 1227–1243. doi:10.1002/acp.1526

Serra, M. J., & Dunlosky, J. (2010). Metacomprehension judgements reflect the belief that diagrams improve learning from text. *Memory*, *18*, 698–711. doi:10.1080/09658211.2010.506441

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning & Instruction*, *13*, 227–237. doi:10.1016/S0959-4752(02)00022-1

Shute, V. J., & Zapata-Rivera, D. (2008). Adaptive technologies. In J. M. Spector, D. Merrill, J. Van Merriënboer, & M. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 277–294). New York, NY: Erlbaum.

Skuballa, I. T., Fortunski, C., & Renkl, A. (2015). An eye movement pre-training fosters the comprehension of processes and functions in technical systems. *Frontiers in Psychology*, *6*, 598. doi:10.3389/fpsyg.2015.00598

Skuballa, I. T., Leber, J., Schmidt, H., Zimmermann, G., & Renkl, A. (2016). Using online eye-movement analyses in an adaptive learning environment. In L. Lin & R. K. Atkinson (Eds.), *Educational technologies: Challenges, applications, and learning outcomes* (pp. 115–142). Hauppauge, NY: Nova Science.

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.

Spüler, M., Krumpe, T., Walter, C., Scharinger, C., Rosenstiel, W., & Gerjets, P. (2017). Brain-computer interfaces for educational applications. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 177–201). New York, NY: Springer.

Stalbovs, K., Scheiter, K., & Gerjets, P. (2015). Implementation intentions during multimedia learning: Using if-then plans to facilitate cognitive processing. *Learning & Instruction*, *35*, 1–15. doi:10.1016/j.learninstruc.2014.09.002

Sweller, J., Ayres, P. L., & Kalyuga, S. (2011). *Cognitive load theory*. New York, NY: Springer.

Toet, A. (2006). Gaze directed displays as an enabling technology for attention aware systems. *Computers in Human Behavior*, *22*, 615–647. doi:10.1016/j.chb.2005.12.010

Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, *20*, 95–99. https://doi.org/10.1016/j.learninstruc.2009.02.009

Veenman, M. J. V., Van Hout-Wolters, B., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition & Learning*, *1*, 3–14. doi:10.1007/s11409-006-6893-0

Wassermann, B., Hardt, A., & Zimmermann, G. (2012). Generic gaze interaction events for web browsers: Using the eye tracker as input device. In *WWW2012 Workshop: Emerging web technologies, facing the future of education* (p. 6). Retrieved from http://www2012.wwwconference.org/proceedings/nocompanion/EWFE2012_006.pdf

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Erlbaum.

Winne, P. H., Vytasek, J. M., Patzak, A., Rakovic, M., Marzouk, Z., Pakdaman-Savoji, A., … Nesbit, J. C. (2017). Designs for learning analytics to support information problem solving. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 249–272). New York, NY: Springer.

Zimmerman, B., & Schunk, D. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.). Mahwah, NJ: Erlbaum.

Chapter 10
Using Data Visualizations to Foster Emotion Regulation During Self-Regulated Learning with Advanced Learning Technologies

Roger Azevedo, Michelle Taub, Nicholas V. Mudrick, Garrett C. Millar, Amanda E. Bradbury, and Megan J. Price

Introduction

Emotions play a critical role during learning and problem solving with advanced learning technologies. However, learners typically do not accurately monitor and regulate their emotions and may therefore not learn as much, disengage from the task, and not optimize their learning of the instructional material. Despite their importance, relatively few attempts have been made to understand learners' emotional monitoring and regulation during learning with advanced learning technologies by using data visualizations of their own (and others') cognitive, affective, metacognitive, and motivational (CAMM) self-regulated learning (SRL) processes to potentially foster their emotion regulation during learning with advanced learning technologies. We present a theoretically-based and empirically-driven conceptual framework that addresses emotion regulation by proposing the use of visualizations of one's and others' CAMM-SRL multichannel data (e.g., cognitive strategy use, metacognitive monitoring accuracy, facial expressions of emotions, physiological arousal, eye-movement behaviors, etc.) to facilitate learners' monitoring and regulation of their emotions during learning with advanced learning technologies. We

An earlier version of this chapter was presented at the 7th International Learning Analytics and Knowledge (LAK 2017) conference, Vancouver, Canada and published as Azevedo, R., Millar, G. C., Taub, M., Mudrick, N. V., Bradbury, A. E., & Price, M. J. (2017). *Using data visualizations to foster emotion regulation during self-regulated learning with advanced learning technologies: a conceptual framework.* In X. Ochoa, I. Molenaar, & S. Dawson (Eds.), *Proceedings of the 7th International Conference on Learning Analytics & Knowledge*: *Understanding, Informing, and Improving Learning* (pp. 444–448). New York, NY: ACM.

R. Azevedo (✉) • M. Taub • N.V. Mudrick • G.C. Millar • A.E. Bradbury • M.J. Price
North Carolina State University, 2310 Stinson Drive, 640 Poe Hall, Raleigh, NC, 27695, USA
e-mail: razeved@ncsu.edu

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_10

use examples from several of our laboratory and classroom studies to illustrate a possible mapping between theoretical assumptions, emotion-regulation strategies, and the types of data visualizations that can be used to enhance and scaffold learners' emotion regulation, including key processes such as emotion flexibility, emotion adaptivity, and emotion efficacy. We conclude with future directions that can lead to a systematic interdisciplinary research agenda that addresses outstanding emotion regulation-related issues by integrating models, theories, methods, and analytical techniques for the areas of cognitive, learning, and affective sciences, human-computer interaction, data visualization, big data, data mining, open learner models, and SRL.

Understanding and reasoning about cognitive, affective, metacognitive, and motivational (CAMM) self-regulated learning (SRL) processes to foster emotion regulation during learning with advanced learning technologies by using data visualizations is key to advancing conceptual, theoretical, methodological, analytical, and educational issues currently plaguing the fields of the learning, cognitive, educational, and computational sciences. For example, data visualizations (e.g., static histogram of the frequency of use of learning strategies across time, videos of eye-gaze behavior during learning with a hypermedia system, etc.) allow researchers, teachers, and learners to visualize, illustrate, conceptualize, model, and understand complex mechanisms and processes, such as metacognitive monitoring and control, accuracy of metacognitive judgments, learning trajectories, cyclical nature of SRL, emotion flexibility, emotion adaptivity, and emotion efficacy.

Data visualizations of CAMM SRL processes are important because they can be used by researchers, learners, teachers, trainers, designers, administrators, and policy-makers for various purposes (Scheiter et al., 2017). These purposes include: (a) articulating complex conceptual issues (e.g., emotion flexibility), (b) illustrating the dynamics of monitoring and control processes, (c) reasoning about the importance of key metacognitive processes and inferring subsequent behaviors, (d) generating hypotheses about underlying SRL mechanisms and their impact on learning, (e) developing teaching and training tools to enhance learners' emotions, self-regulation, and emotion regulation and teachers' ability to monitor and regulate learners' emotions and SRL, and (f) developing sophisticated teacher dashboards to trigger and support instructional decision-making that include interface elements (e.g., open learner models representing metacognitive accuracy) to provide learner data that may foster changes in SRL behaviors.

Despite these and other potential uses of data visualizations, many conceptual, theoretical, methodological, and analytical questions remain unanswered (see Buder & Hesse, 2017). For example, while the majority of research has focused on data visualizations designed for learners to keep track of their task- or domain-specific knowledge and skills by using open learner models (Bull & Kay, 2016), teacher dashboards have focused on presenting student assessments (Verbert, Duval, Klerkx, Govaerts, & Santos, 2013; Verbert, Govaerts, et al., 2013) with little attention paid to learners' emotions experienced during learning with advanced learning technologies (ALTs) such as intelligent tutoring systems, multimedia, hypermedia, simulations, and serious games. Emotions are critical for learning, problem solving, and performing tasks across domains, and if left unchecked can have deleterious impacts on learning (e.g., frustration transitioning to boredom and disengagement; D'Mello & Graesser, 2012). As such, our chapter focuses on the neglected area of emotion regulation during learning with the goal of presenting, discussing, and debating the *significance* of visualizing CAMM SRL process data such as metacognitive judgments, deployment of cognitive strategies, fluctuations in affective states, changes in self-efficacy, nature of temporality, granularity of time scales, levels of abstraction, static versus dynamic representations, converging evidence from multimodal multichannel data, etc. This is accomplished by using various *methodological* tools including eye-tracking data, concurrent and retrospective think alouds, classroom discourse, facial expressions of emotions, log files, recordings from physiological sensors, screen recordings, verbal and nonverbal human-artificial agent interactions, and classroom videos to foster learners' emotion regulation with advanced learning technologies (see Azevedo, 2015). In sum, by addressing emotion regulation with these visualizations we argue that they can advance the fields of emotion and emotion regulation, SRL, human-computer interaction (HCI), data visualizations, big data, data mining, personalized instruction, and advanced learning technologies.

Self-Regulated Learning (SRL), Emotions, and Advanced Learning Technologies (ALTs)

Fundamentally, SRL involves the temporal deployment of CAMM processes during complex learning with advanced learning technologies, such as MetaTutor, Betty's Brain, Crystal Island, and nSTUDY (Azevedo, Taub, & Mudrick, 2015; Biswas, Segedy, & Bunchongchit, 2016; Sabourin & Lester, 2014; Winne & Hadwin, 1998). SRL plays a central role in learning, reasoning, and problem solving in complex domains with advanced learning technologies (Azevedo & Aleven, 2013a, 2013b; Biswas et al., 2016; Santos, Govaerts, Verbert, & Duval, 2012; Winne & Azevedo, 2014; Winne et al., 2017). CAMM processes can be measured using trace data including log files, eye tracking, physiological sensors, and facial expressions of emotions (Azevedo, 2014a, 2014b; Azevedo et al., 2013; Azevedo, Martin, et al., 2016; Azevedo, Taub, Mudrick, Farnsworth, & Martin, 2016; Azevedo et al., 2015; Biswas et al., 2016; D'Mello & Graesser, 2012; Lafford, 2004; Winne & Hadwin, 1998). Despite the emerging evidence on the effectiveness of multimodal multichannel trace data to measure emotions (Azevedo & Aleven, 2013a, 2013b; Calvo, D'Mello, Gratch, & Kappas, 2015), there is no systematic program of research that focuses on emotion regulation.

Emotion regulation is vital to successful emotions and can significantly impact learning and SRL about complex topics across domains with advanced learning technologies such as intelligent tutoring systems, serious games, virtual environments, simulations, augmented reality, and tangible user interfaces. Contrary to the wealth of empirical research about the fundamental role of emotion regulation found in the affective, clinical, social, and neurosciences, there is a paucity of research on emotion regulation in the educational, learning, and cognitive sciences (see Gross, 2015a, 2015b; Pekrun & Linnenbrink-Garcia, 2014; Scherer, 2009). Emotions play a critical role in learning and without a systematic, empirically-driven program of research aimed at understanding the monitoring and regulatory processes underlying emotion regulation, there is a lack of a fundamental understanding of the affective SRL processes that can impact learning, problem solving, and conceptual understanding, especially in STEM domains with advanced learning technologies, where ineffective learning can lead to failure in STEM careers and diminish the impact and educational effectiveness of advanced learning technologies.

We argue that advances in emotion regulation are needed and that researchers should focus on two specific areas. First, collecting rich traces of temporally unfolding emotions along with other SRL processes (e.g., metacognitive judgments, use of effective cognitive strategies) during learning with advanced learning technologies to understand the nature of these processes and how their characteristics (e.g., antecedents, duration, sequence, parallel/serial deployment) are related to changing internal and external demands, and are predictive of performance (e.g., on embedded science assessments), emotion regulation (e.g., emotion flexibility, emotion efficacy, and emotion adaptivity), learning, and transfer (e.g., immediate and delayed pre-post learning gains). Second, scaffolding emotion-regulation strategies in real-time to provide students with optimal scaffolding that supports emotions, emotion regulation, as well as other SRL processes (e.g., feedback from pedagogical agent → confusion → prolonged fixation on pedagogical agent → frustration → pedagogical agent deploys situation modification strategy by suggesting learner selects a new learning goal → . . .). In terms of our emerging conceptual framework, we begin with the presentation of Table 10.1, which lists several multimodal multichannel affect data sources, including both online trace and self-report data, typically collected by researchers. We argue that the collection of these data sources are critical in ascribing data visualizations of self- and other CAMM-SRL processes to foster learners' emotion regulation.

Despite the potential benefits of using data visualizations of CAMM SRL to foster learners' emotion regulation, humans find themselves with the growing pressure of communicating or simply making sense out of increasingly more abstract data. Moreover, humans struggle with abstracting meaning from large amounts of complex data, especially when the physiological sensors and instruments used to collect the process data grow more advanced. Visualization techniques that are validated as effective for one particular type of problem or context may not be well suited for another. As such, it is difficult to design a data visualization framework that can both keep up with the technology used to collect online trace data as well as apply design components across multiple domains and different types of learners. For example, presenting physiological data (e.g., high frequency of skin conductance responses from electrodermal activity; see Table 10.1) in

Table 10.1 Sample multimodal multichannel affect data used to detect emotions and provide evidence for scaffolding emotion regulation

Data sources	Affective indicator/component
Screen recordings (video and audio)	• Context (mouse movements, facial expressions, interface elements, learner-agent dialogue, etc.) • Emotion eliciting event (pedagogical agent scaffolding, complexity of multimedia content)
Concurrent and retrospective think-alouds	• Self-reports of emotions verbalized either *during* or *following* learning, problem solving, and performance
Eye tracking	• Repeated number of fixations on areas of interest (AOIs) • Sequential patterns of fixations • Revisits to AOIs • Saccades, smooth pursuit eye movements (e.g., overall gaze behavior)
Log files	• Context (mouse movements, facial expressions, interface elements, learner-agent dialogue, etc.) • Emotion eliciting event (pedagogical agent feedback) • Task performance • Scores on embedded assessments • Time spent on various aspects of the system (e.g., time spent reading text, inspecting diagram, planning) • Sequence of events during interaction with advanced learning technologies (e.g., planning → setting goals → reading → using strategies → taking quiz → regulating emotions → complying with pedagogical agent . . .)
Facial expressions of emotions	• Automated evidence scores of learner-centered emotions (e.g., frustration, confusion, boredom) • Automated evidence scores of basic emotions (e.g., joy, fear, anger, sadness, disgust, surprise) • Automated evidence scores of action units (AUs) (e.g., AUs 1, 4, 14, 28, 30) • Automated evidence scores of valence • Manually coded/Observer coded scores of basic and learner-centered emotions, and AUs • Micro-expressions
Physiological sensors	• Skin conductance responses • Electrodermal activity over time (tonic vs. phasic)
Self-report questionnaires (e.g., AEQ, EV, ERQ, PAUSe)	• Perceptions of academic achievement emotions (AEQ) • Perceptions of current emotions (EV) • Perceptions of ability to cognitively reappraise and suppress expressions (ERQ) • Perceptions of the utility of affect and emotions (PAUSe)

real time poses serious challenges for learners' interpretation, comprehension, and inference generation (critical for emotion monitoring and regulation) by potentially imposing extraneous cognitive load and inducing negative emotions (e.g., frustration). Cognitive load theory (Paas & Sweller, 2012) describes the relationship between the capacity of working memory and the cognitive demands of a particular task. The core of the theory asserts that people have a limited cognitive capacity, which can lead to decreased interpretation, comprehension, and inferences made during visualization-centered learning and problem solving tasks. In addition, representations depicting processes (e.g., facial expressions of confusion, high frequencies of skin conductance responses from electrodermal activity) in real-time call for accurate and rapid detection, discrimination, and comprehension of relevant information. As such, it is necessary to design an interface that minimizes extraneous cognitive load experienced by learners when attempting to comprehend and reason about visualizations presented in real-time during the process of emotion regulation (e.g., attending to the CAMM SRL data visualizations, making inferences about the antecedents that are portrayed in the data, selecting the relevant emotion-regulation strategy necessary to address the experienced emotions, using the emotion-regulation strategy, assessing the adequacy of strategy use, etc.). In the next section, we present a succinct synthesis of current advanced learning technologies' detection and fostering of emotions during learning and problem solving.

Detection and Fostering of Emotions During Learning and Problem Solving by ALTs

Over the past decade, advanced learning technologies have been developed to detect and respond to student emotions (Calvo et al., 2015; D'Mello, Lehman, Pekrun, & Graesser, 2014). Foundational studies have shown that learner-centered emotions (e.g., boredom, confusion, frustration; D'Mello, 2013) occur far more frequently during interactions with advanced learning technologies than basic emotions (e.g., anger, joy, sadness; Ekman, 1977). These empirical findings have been driven by human labeling of emotions, through self-reports or trained observers (Porayska-Pomsta, Mavrikis, D'Mello, Conati, & Baker, 2013). Further, these efforts have resulted in a theoretically-based and empirically-refined model of affective dynamics that reveals likely antecedent-consequent associations among learning-centered emotions (D'Mello & Graesser, 2012). Results to date have shown that emotionally adaptive advanced learning technologies (e.g., Affective AutoTutor) may improve learning over non-emotionally adaptive advanced learning technologies for specific subsets of students, such as students with low prior knowledge. Emotionally adaptive advanced learning technologies have specifically addressed the problem of detecting transient emotional states and providing immediate responses. In the case of spoken dialogue advanced learning technologies, the system may produce an utterance in order to change students' current emotions. For instance, students who display boredom may be told that more exciting material will come soon (Nye, Graesser, & Hu, 2014). Alternatively, an advanced learning technology may alter interface elements to provide a visual cue to allow students to receive emotionally adaptive feedback at their own pace (Grawemeyer et al., 2016).

AutoTutor, a spoken dialogue tutoring system, has been shown to be consistently adaptive to students' emotions across learning domains (Graesser, 2016; Graesser, Lippert, & Hampton, 2017). One specific version of AutoTutor, Supportive AutoTutor, provides supportive dialogue using an animated virtual tutor that externally attributes the causes of students' emotions to the learning material by adapting to negative emotions (e.g., boredom or frustration) by empathizing that the current learning material is boring or difficult while encouraging task completion so that students can move on to the next task (D'Mello & Graesser, 2012). Supportive AutoTutor detects learner-centered emotions (i.e., boredom, confusion, frustration) through a combination of dialogue features, body posture movement, and facial expressions. In contrast, ITSPOKE uses spoken dialogue to handle uncertainty and disengagement during physics problem solving (Forbes-Riley & Litman, 2011; Litman & Forbes-Riley, 2014). The most recent version of ITSPOKE included both uncertainty and disengagement adaptations (Litman & Forbes-Riley, 2014), which produced higher learning gains with both uncertainty and disengagement adaptations compared to regular ITSPOKE, but this was not statistically significant, perhaps due to the small sample size used across three conditions. Another system, Wayang Outpost (MathSpring), is an emotionally adaptive advanced learning technology for mathematics learning (Muldner et al., 2015; Woolf et al., 2009). It includes a cartoon-style animated pedagogical agent (male or female, as selected by students) that provides emotional interventions and supportive feedback to students. The agents adapt to students' emotions using a variety of strategies, including empathetic nonverbal behavior displays, providing supportive messages that emphasize students' efforts, and attributing negative states to the learning material. Lastly, the most recent emotionally adaptive advanced learning technology is iTalk2Learn, developed for mathematics learning (Grawemeyer et al., 2016), where students solve fraction problems and interact with the advanced learning technology. iTalk2Learn provides support through affect-boosting messages, task-oriented messages and prompts, self-explanation prompts, and talk-aloud prompts. Recent findings reveal that students were found to engage in less off-task behaviors during interactions with emotionally adaptive iTalk2Learn, but with learning gains equal to a non-emotionally adaptive version.

Despite substantial evidence indicating that students dynamically transition between emotions during learning, existing advanced learning technologies respond to isolated instances of emotions. Understanding *what*, *why*, *how*, and *when* these transitions occur is crucial for effective intervention and the basis for emotional monitoring and regulation, and emotion regulation scaffolding. More specifically, the *what* refers to the exact emotional state (e.g., anger, frustration) assuming it can be accurately detected; the *why* relates to the antecedent or event that caused or induced the specific emotional state; the *how* refers to the multi-componential expression (e.g., facial expression, physiological arousal) of the emotion. Lastly, *when* refers to the timing (e.g., duration) of the occurrence of the emotional state and needs to be contextualized for interpretation (e.g., a pedagogical agent's negative feedback following an incorrect inference induced confusion in a learner). As such, our framework focuses on the next generation of emotionally adaptive advanced learning technologies that aim to understand the history, sequence, and dynamics of students' emotions in real-time to provide targeted support for emotion adaptation and regulation. Given the current state of existing advanced learning technologies' emotion regulation capabilities, our framework aims to use data visualizations to foster emotion regulation during SRL with advanced learning technologies that is theoretically (e.g., Bonanno & Burton, 2013; Bosse, Gerritsen, Man, & Treur, 2013; D'Mello & Graesser, 2015; Doré, Silvers, & Ochsner, 2016; Gratch & Marsella, 2015; Gross, 2015a, 2015b; Reeck, Ames, & Ochsner, 2016; Scherer, 2009; van Kleef, 2009) and empirically-based, as well as based on existing research on dashboards and open learner models (see Bull & Kay, 2016; Verbert, Duval, et al., 2013; Verbert, Govaerts, et al., 2013), as earlier described in this chapter. Next we describe the several key theories and models of emotions and emotion regulation that serve as the foundation for our emerging conceptual framework.

Theoretical Background: Emotion Regulation During Learning with ALTs

Despite the major emphasis on cognitive and metacognitive SRL processes, relatively little attention has been paid to understanding the complex underlying processes related to learners' accurate monitoring and regulation of their emotions (Azevedo et al., 2015, Azevedo, Taub, et al., 2016). While we acknowledge the importance of motivation, it is not the focus of this chapter. We argue that neglecting emotion regulation is detrimental to learning and that researchers need to focus on measuring, detecting, modeling, and fostering emotion regulation processes to sustain engagement and foster a positive affective experience during learning with advanced learning technologies across disciplines (e.g., solving a biological outbreak with a serious game, understanding how different organs support human biological functioning with a hypermedia-based intelligent tutoring system, planning resource allocation with tangible landscape models, etc.). As such, this section provides a brief overview of the existing models of affective dynamics (used by advanced learning technology researchers) followed by two leading models of emotion and emotion regulation that incorporate emotion regulation strategies that can be modeled using data visualizations to support their use during learning.

A major theoretical framework guiding much of the work in affective computing with advanced learning technologies is D'Mello and Graesser's (2012) model of affective dynamics, which suggests that affective states (e.g., boredom, confusion, frustration) that are experienced during complex learning can be expressed in various ways (i.e., based on the multi-componential assumption of emotions) from subjective feelings (collected with self-report measures), motor expressions (e.g., facial expressions of emotions), action tendencies (e.g., gaze behaviors when coordinating multiple representations of information), physiological measures (e.g., frequency of skin conductance responses), etc. that can be measured using multichannel trace data (see Azevedo, Taub, et al., 2016; Harley, Bouchet, Hussain, Azevedo, & Calvo, 2015; Taub et al., 2016). As such, our focus (as proposed by major models by D'Mello & Graesser, 2012; Gratch & Marsella, 2015; Gross, 2015a, 2015b; Kappas, 2013; Scherer, 2009; van Kleef, 2009, 2016) will be discussing data visualizations of multimodal multichannel CAMM SRL data that can assist learners in the regulation of their emotions given a myriad of factors. These factors include the repertoire of emotion-regulation strategies, efficacy in the use of emotion regulation strategies, emotion regulation flexibility based on temporal adaptation to different *internal* (e.g., lack of cognitive strategies, standards for metacognitive monitoring) and *contextual* (e.g., compliance with pedagogical agent feedback and scaffolding, perceived utility of pedagogical agent's feedback and scaffolding, temporality of competing goals, frequency of pedagogical agent's pedagogical interventions) conditions and demands.

Emotion regulation is complex, therefore it necessitates the integration of models of cognition, metacognition, and affect proposed in the cognitive, learning, social, and computational sciences (e.g., see Barrett, Lewis, & Haviland-Jones, 2016; Calvo et al., 2015). Overall, there is an assumption that emotions experienced during complex learning will impact learning and problem solving. D'Mello and Graesser's (2012) model fits well into learning with advanced learning technologies to foster complex learning. However, the model does not account for individual differences in the ability to regulate emotions. Some learners may make many attempts at overcoming obstacles before transitioning between emotions (e.g., from confusion → frustration). Conversely, some learners may only persist briefly before becoming confused or frustrated. Their model shows the likely paths through emotional states, but cannot determine when learners will transition between states, and why they transition. To date, their model has been used to collect affect data during complex learning with Affective AutoTutor, but it has not been used to provide emotion regulation scaffolding during learning. This and other work on advanced learning technologies (e.g., MetaTutor, CRYSTAL ISLAND) have documented the types of affective states learners experience, and the predictive adequacy of emotions linked to several learning and performance measures (e.g., Taub et al., 2017; Harley, Poitras, Jarrell, Duffy, & Lajoie, 2016; Harley et al., 2015). However, none have used multichannel data to examine and provide emotion-regulation scaffolding during learning. Below, we provide a succinct overview of two major models—(a) Scherer's (2009) component process model (CPM) and (b) Gross' (2015a, 2015b) process model of emotion regulation and subsequently describe our conceptual framework for supporting emotion regulation during SRL with ALTs.

A comprehensive model by Scherer (2009) describes emotions as dynamic processes that unfold over time, and includes interactions between multiple components and processing on several levels. Accordingly, emotions arise from different *appraisals* of a relevant event (e.g., setting learning goals, relevance of a diagram, compliance with pedagogical agent's feedback and scaffolding). The appraisal sequence includes four steps: (a) relevance, (b) implications, (c) coping, and (d) normative significance. These appraisal steps consist of discrete sets of evaluations that define different appraisal outcomes. For example, evaluation of the relevance of an event includes assessments of its novelty, intrinsic pleasantness, and goal conduciveness. The appraisals can occur in a fixed or recursive manner, which means that some appraisal outcomes might be revisited and updated during an emotional episode and that some of those processes even occur simultaneously. The appraisal process both influences and is influenced by other components of the emotional process including the subjective feeling component, motor expressions, action tendencies, and physiological changes, as well as other psychological processes such as metacognition and SRL.

When applying Scherer's (2009) CPM to learning with ALTs, one of the core assumptions is that the emotion a learner experiences is defined by the appraisal of a significant event. The emotion-eliciting event can be either internal (e.g., remembering an unpleasant interaction with a pedagogical agent) or external (e.g., receiving feedback from a pedagogical agent during learning). The appraisal criteria can be processed on different levels, which can range from effortless, unconscious, low-level neural mechanisms (e.g., pattern matching for attention/perception) to complex, conscious considerations involving metacognitive control. Even though emotional episodes are defined by the interaction of several components, the CPM postulates that the nature of an emotion is exclusively characterized by the pattern of appraisal outcomes and their development over time (Scherer, 2009). Given the extensive research by Scherer and colleagues, we use CPM to theoretically guide our discussion on emotions, SRL, and emotion regulation during learning and problem solving with ALTs with multimodal multichannel CAMM SRL data.

Additionally, our conceptual framework incorporates Gross' Process Model of Emotion Regulation (Gross, 2015a, 2015b; Suri & Gross, 2016) that posits that an emotion may be regulated at five points in the emotion generation process: (a) selection of the situation, (b) modification of the situation, (c) deployment of attention, (d) cognitive change, and (e) modulation of responses. Although regulation strategies can be—and often are—used in combination, the heuristic value of this framework arises from its ability to simplify a complex problem space and direct attention to each of the separate families of emotion regulation (Gross, 2015a). The model makes several assumptions: (a) emotions involve loosely coupled changes in subjective experience, behavior, and peripheral physiology; (b) emotions unfold over time; (c) emotions can be either helpful or harmful, depending on the context; and (d) emotions differ based on duration, intensity, and quality. In terms of the four emotion-regulation strategies, next we describe each and illustrate their importance in learning. Situation selection is not included in our taxonomy since we assume that using an advanced learning technology to learn and solve problems precludes one from being able to select a (new) situation.

Table 10.2 Taxonomy of emotion-regulation strategies and corresponding scaffolding strategies potentially enacted by advanced learning technologies

Situation modification	Attentional deployment	Cognitive change	Response modulation
Postpone or create a sub-goal	Divert attention and make metacognitive judgment to assess relevancy of content (instead of only reading)	Make a metacognitive judgment (e.g., negative judgment of learning [JOL-]), get positive feedback from agent → do understand it	Self-initiate prompts when frustration arises (thinking about prompts)
Select different content pages or multiple representations of information	Ignore prompt to summarize after repeated prompts to avoid frustration	Spend more time reading content → not confused, just need more time	Smile/laugh after receiving a low quiz score (frustration → laughing)
Do not write a summary when prompted by the system	Inspect diagram (read less) to ensure understanding	Make a summary when prompted → not frustration, it can be helpful	Postpone a sub-goal to decrease frustration with failing sub-goal quizzes
Add content and learning strategies to a Planner	Focus on Planner at the top of the interface to assess if following it will resolve current negative affective state	Try to understand what feelings are → not failing to use Skill Diary, just need explanations of terms	Watch a movie that describes pollution to be more engaged (boredom → engagement)
Modify entries in a Skill Diary	Listen to agent's explanation of emotions without reading content	Read through table of contents (or similar content structure) to find relevant material → not confusion, just reading wrong material	Ask the pedagogical agent what learning strategies are so they can be used effectively (confusion → engagement)

Situation selection is not included in our taxonomy since we assume that using an advanced learning technology to learn, solve problems, etc. precludes one from being able to select a (new) situation

In Table 10.2 we illustrate our initial taxonomy addressing four emotion-regulation strategies and corresponding scaffolding strategies that can be enacted by ALTs to foster effective emotion regulation. This table addresses the four main emotion-regulation strategies delineated in Gross's model and are based on Scherer's CPM model, D'Mello and Grasser's model of affective dynamics, as well as on extensive research on emotions in ALTs including theoretical assumptions of SRL (see Schunk & Greene, in press.; Zimmerman & Schunk, 2011) and related empirical evidence on CAMM SRL and ALTs (see Azevedo & Aleven, 2013a, 2013b).

First, **situation modification** refers to taking actions that directly alter a situation in order to change its emotional impact. Examples include not complying with an agent's feedback, selecting a new learning goal, and electing to take an embedded assessment to control cognition following an accurate metacognitive judgment, decrease an emerging negative emotion (e.g., frustration), and increase persistence and interest in the overall learning activity (motivational processes not covered in our taxonomy). Second, **attentional deployment** refers to directing one's attention with the goal of influencing one's emotional response. A common strategy involves redirecting attention within a given situation (e.g., from an emotion-eliciting interaction with complex representations of information) or shifts attention away from the present situation altogether (e.g., averting gaze behavior from a pedagogical agent to search for new informational sources). This strategy involves change in one's gaze and/or shifts in one's internal focus on different interface elements in ALTs and can be elicited either by the individual learner, proposed by the pedagogical agent, or scaffolded by visualizations. The third emotion-regulation strategy is **cognitive change** and refers to modifying one's appraisal of a situation in order to alter its emotional impact. Cognitive change can be applied to an external situation (e.g., "my performance in this laboratory study is not directly relevant to my major but it is a chance for me to learn more about science"), internal situation (e.g., "My racing heart is not a sign of anxiety; it means my body is preparing for understanding this complex diagram"), or to alter how one thinks about one's capacity to manage situational demands using self-regulatory skills (e.g., "Although learning about this content feels overwhelming, I know I can handle it using my existing knowledge, skills, and the support provided by the system"). The last emotion-regulation strategy discussed is **response modulation,** which refers to directly influencing experiential, behavioral, or physiological components of the emotional response after the emotion is well developed. For example, when learning with ALTs we envision learners' self-initiating prompts when frustration arises as smiling when receiving negative feedback or a low score on a quiz or managing boredom by watching a video of some related topic.

In summary, we argue for a theoretical amalgamation of the model of affective dynamics, componential process model of affect, and a process model of emotion regulation to provide a comprehensive foundation for a conceptual framework that aims to foster emotion regulation using data visualization from CAMM SRL process data. Our approach integrates these leading models, theories, and frameworks along with contemporary research on dashboards and open learner models.

Conceptual Framework: Supporting Emotion Regulation and SRL Through CAMM Data Visualizations

Supporting learners' emotion regulation during SRL with ALTs by providing theirs and others' CAMM SRL data visualizations is innovative and has the potential to significantly impact learners' SRL and the educational effectiveness of ALTs. In addition to the theoretical and empirical literature already presented, we make a few additional assumptions when proposing our conceptual framework.

Core assumptions. First, multimodal multichannel CAMM SRL data are *relevant* and therefore allow learners to accurately monitor and regulate their emotions. We emphasize that the relevancy of CAMM SRL data will vary based on individual differences, one's ability to accurately monitor and regulate CAMM processes, and internal and contextual conditions. Second, the relevance of CAMM SRL data visualizations has to be *synchronized* accurately and *coordinated* with other visualizations in order for maximum potential for facilitating monitoring and regulation of one's emotions. Third, presentation of multimodal multichannel data should not exceed human cognitive architecture constraints (e.g., limited attention, not exceed WM capacity, cognitive load), must adhere to SRL assumptions (e.g., learners have the *potential* to monitor and regulate their cognitive, affective, and metacognitive processes), not compete for learners' attention, and be easy to process (i.e., not become a dual task) and actionable (i.e., providing data with which learners can control their emotions). Fourth, the coordination and timely presentation of multimodal multichannel data needs to be carefully considered based on the learner's immediate "needs" (e.g., deal with negative affect that exceeds a time threshold and is negatively impacting the learner's level of cognitive engagement) as well as their learning history (typically modeled in a student or user model). This assumption is based on several types of data collected from the learner (e.g., self-reports, online trace data; see Table 10.1 for a sample) at different times (e.g., prior to learning, during learning, and following learning) and across time spans (e.g., several minutes, one hour, one day, one month, one semester, etc.). Fifth, CAMM processes are represented in ways that include *number of representations* for each and all CAMM SRL processes (e.g., histogram of frequency of use of learning strategies vs. histogram of frequency of facial expressions of emotions and corresponding histogram of the learner's effectiveness of use of specific emotion-regulation strategies), *representation type* (e.g., static vs. dynamic or both), *ease of understanding and interpretability* (e.g., balance between simple to understand and abstract representation that may require further cognitive processing and reflection of one's affective reactions), and *configuration* (e.g., if more visualizations where each one focuses on a different CAMM process, then how will it be presented on the interface to maximize selective attention, processing, interpretation, and understanding vis-à-vis SRL behaviors, learning, and problem solving. Sixth, it is critical to account for the roles of CAMM SRL processes both individually and together when considering the abovementioned assumptions; however, in this chapter we focus on the emotional aspects. Seventh, we assume that all learners (of all ages, levels of expertise, levels of prior knowledge) have the potential to accurately monitor and effectively regulate their CAMM processes. However, there may be individual-, task-, and context-specific variables, factors, and processes that may impede the successful monitoring and regulation of CAMM SRL processes. Eighth, our framework emphasizes the strategic use and presentation of CAMM SRL visualizations designed to facilitate learners' accurate monitoring and regulation of their emotions. Lastly, data visualizations of CAMM processes are important because emotion monitoring and regulation can stem (i.e., the antecedents) from cognitive, metacognitive, and affective processes. We acknowledge the importance of motivational processes; however, we do not discuss it due to space limitations. Below we describe the conceptual framework based on the type of SRL process, data sources, data representations, and emotion monitoring and regulation, as well as the scaffolding that can be applied to each situation.

We present an initial conceptual framework (see Table 10.3) that includes the CAMM SRL processes, multimodal multichannel data sources, and sample data representations to be used as the data visualizations to foster emotion regulation, and a preliminary description of some scaffolding techniques that will allow learners to engage in accurate monitoring and effective regulation of their emotions during learning. We highlight the following about Table 10.3: (a) it is not a complete list of CAMM SRL processes (e.g., we did not include motivation); (b) there are many other data sources from sensors and instruments (e.g., posture sensor) that can be used to collect CAMM SRL data that are not included in Table 10.3; (c) the sample of data representations is not comprehensive but includes the most notable examples for optimally supporting emotion regulation; (d) we list several scaffolding[1] techniques for emotion regulation monitoring and scaffolding that can augment those presented in Table 10.2; and, (e) we do not discuss how the core assumptions as well as the other abovementioned theoretical assumptions relate to critical issues related to scaffolding (e.g., timing, type, duration, delivery method, assessing effectiveness; based on Azevedo & Hadwin, 2005; Azevedo et al., 2015; Belland, Walker, Kim, & Lefler, 2017) nor do we make strong assumptions about how to deal with more complex issues such as emotion flexibility, emotion efficacy, and emotion adaptivity (Aldao, Nolen-Hoeksema, & Schweizer, 2015).

Sample Data Visualizations to Foster Emotion Regulation from CAMM SRL Data. In Table 10.3 we list a few examples of data representations that can be used to foster and scaffold learners' emotion regulation during SRL with ALTs. Figure 10.1a illustrates the use of a heat map to show a learner that a source of frustration is their lack of attention to the intelligent virtual human as the learner tended to allocate most of their attention to the last paragraph followed by the first and second paragraphs and the diagram.

A second example, presented in Fig. 10.1b, is a learner's gaze behavior exemplifying how the static or dynamic representation (video clip of gaze behavior) can be used by a pedagogical agent (during scaffolding or emotion regulation training) to explain (or verbally walk through) how a learner's attentional deployment is indicative of an ineffective strategy when trying to learn from multimedia materials because it is not indicative of effective selection, organization, and integration of the text and diagram (see Mayer, 2014). Additionally, the speed of attentional deployment is related to poor metacognitive monitoring (Antonietti et al., 2015) needed to accurately assess the relevancy of the materials. Additionally, relatively few fixations on the human's face failed to capitalize on the agent's nonverbal facial

[1]Up to this point in our chapter we have strategically not used the word scaffolding. We now emphasize scaffolding because there is an implicit assumption that learners will require scaffolding to accurately monitor and regulate their emotions.

Table 10.3 Conceptual framework including the CAMM SRL processes, multimodal multichannel data sources, and sample data representations as visualizations to foster emotion regulation

CAMM SRL processes	Data sources	Sample data representations	Scaffolding techniques for ER monitoring and scaffolding
COGNITIVE Cognitive Load	Eye tracking	Heat map and overall gaze behavior (e.g. revisits to areas of interest [AOIs], sequential patterns of fixations, pupil dilation, etc.). (See example illustrated in Fig. 10.1a, b)	Example of advanced learning technology with overlaid summary statistic of eye-tracking data • Redirection of attention to specific component of interface (e.g., multimedia content, agent, overall learning goal) • Scaffolding indicating that "you should be focusing on element X of the interface instead of element Y, and this is why . . . " [provide explanation and articulate reasons] • Use eye movement modeling examples (EMMEs) (e.g. presentation of expert and/or peer's eye movements) compared to one's gaze behavior to facilitate underlying cause of frustration and confusion
COGNITIVE Cognitive Strategy Use	Log files, screen recording, verbalizations, eye tracking	Comparison of frequency and duration of strategy use (at both macro and micro level) and valence	Overall strategy is to have learners compare their strategy use with others (e.g., peers) • Redirect user to take notes instead of only reading • Scaffolding indicating that "you are focusing too much on strategy X when you should be using strategy Y", and this is why . . . " [provide explanation and articulate reasons] • Modeling of declarative and procedural knowledge through instructional video, pedagogical agent, interface elements, etc.
META-COGNITIVE Metacognitive Judgments	Log files, verbalizations, eye tracking, screen recording	Comparison of metacognitive judgments made during critical points of learning (e.g., completion of a learning goal) compared to task performance (e.g., quiz score)	Overall strategy is to have learners compare their metacognitive judgments relative to learning and SRL behaviors at specific moments during learning and problem solving compared to ideal metacognitive monitoring (e.g., perfect calibration), their history of metacognitive judgments (during learning), compared to peers' metacognitive accuracy (e.g., everyone else is under confident and this is the source of frustration)

(continued)

Table 10.3 (continued)

CAMM SRL processes	Data sources	Sample data representations	Scaffolding techniques for ER monitoring and scaffolding
			• Redirect learner to activate their prior knowledge to determine their understanding of the current multimedia materials • Modeling of declarative, procedural, and conditional knowledge through instructional video, agent, interface elements, etc. and link to emotion regulation strategies.
AFFECTIVE Physiological Data	Galvanic skin response and electrodermal activity (GSR and EDA)	Skin conductance responses and EDA over time (See example illustrated in Fig. 10.1c)	Overall strategy involves using a learner's GSR in comparison to baseline • Highlight specific time points of low/high arousal and how they are related to other expressions of emotions (e.g., facial expressions of negative emotions) • Scaffolding indicating, "You seem to be getting bored, try selecting a new learning goal and then summarize the content instead of re-reading" • Emphasize and illustrate emerging accurate emotion regulation flexibility and efficacy by graphically illustrating how negative affective reactions to agent's scaffolding have now transitioned to positive reactions accompanied by complying to scaffolding and feedback
AFFECTIVE Facial Expressions	Facial expressions & verbalizations (basic & learner-centered emotions)	Video data streams of learners' facial expressions	Overall strategy involves displaying face video alongside associated summary statistics of experienced emotions • Redirect attention from an emotion-eliciting feature of interface to a more positive feature • Scaffolding indicating, "Here is what these emotions mean, see if you can understand more about each" • Focus on explicit training of emotion-regulation strategies (e.g., see Table 10.2) that will lead to successful emotion regulation (e.g. situation modification, attentional deployment, cognitive change, response modulation)

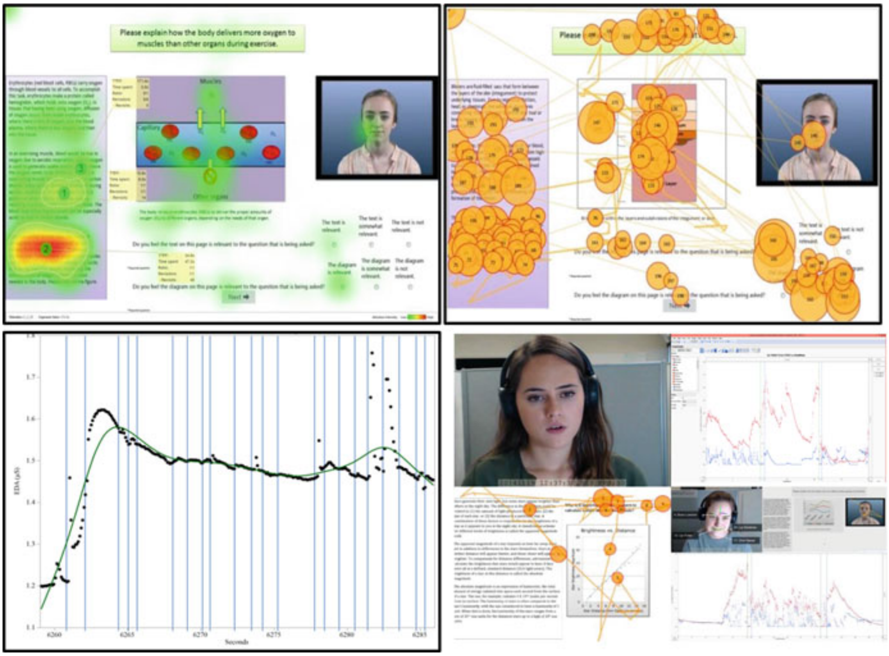

Fig. 10.1 Sample data visualizations based on CAMM SRL data that can be used to foster and scaffold emotion regulation, including a heat map from eye movements (**a**), gaze behavior (**b**), electrodermal activity (**c**), and multimodal multichannel data (**d**).

cues related to the relevancy of the materials vis-à-vis the science question (green box at the top of the image). It is important to highlight that these two examples relate to the *attentional deployment* emotion regulation strategy proposed by Gross (2015a, 2015b) (see also Table 10.2).

Physiological data about arousal levels (from electrodermal activity sensors) are also important when considering data visualizations to foster emotion regulation. In Fig. 10.1c we present a learner's arousal level (black line compared to baseline) that can be presented as an effective data visualization to facilitate a learner's emotion regulation. The visualization can indicate specific episodes during learning when a learner's levels of arousal exceed their pre-established baseline. As such, emotion regulation scaffolding would focus on *response modulation* (Gross, 2015a, 2015b) by highlighting that modulating their response to a pedagogical agent's scaffolding to use an effective sophisticated strategy (e.g., making inferences) is related to their lack of self-efficacy in using the strategy and this created a high level of arousal accompanied by facial expressions of confusion and frustration that persist and may lead to cognitive disengagement if a response modulation strategy is not used by the learner.

Our fourth example includes the use of multimodal multichannel data (see Fig. 10.1d) whereby the advanced learning technology will present several types of

CAMM SRL data (i.e., leaners' facial expressions, gaze behaviors, context, coded facial expressions, and several physiological indicators of affect) to scaffold a learner through explicit emotion regulation training. This emotion regulation scaffolding includes, but is not limited to, articulating the conditional knowledge related to emotion monitoring and regulation and how underlying cognitive, metacognitive, and motivational SRL processes influence the learner's emotions during learning and problem solving and how different emotion-regulation strategies can be deployed to address each of the SRL processes (in Table 10.2).

Lastly, we present a more complex visualization (see Fig. 10.2) based on the integration of several CAMM SRL data (similarly presented in Fig. 10.1) that can be used to provide elaborate adaptive scaffolding to a learner by illustrating how cognitive and metacognitive processes are the basis for this particular emotion-regulation scaffolding. This adaptive emotion regulation scaffolding technique explicitly illustrates critical variables such as time spent on different (relevant and irrelevant) areas of interest as well as the time spent on and accuracy of use of

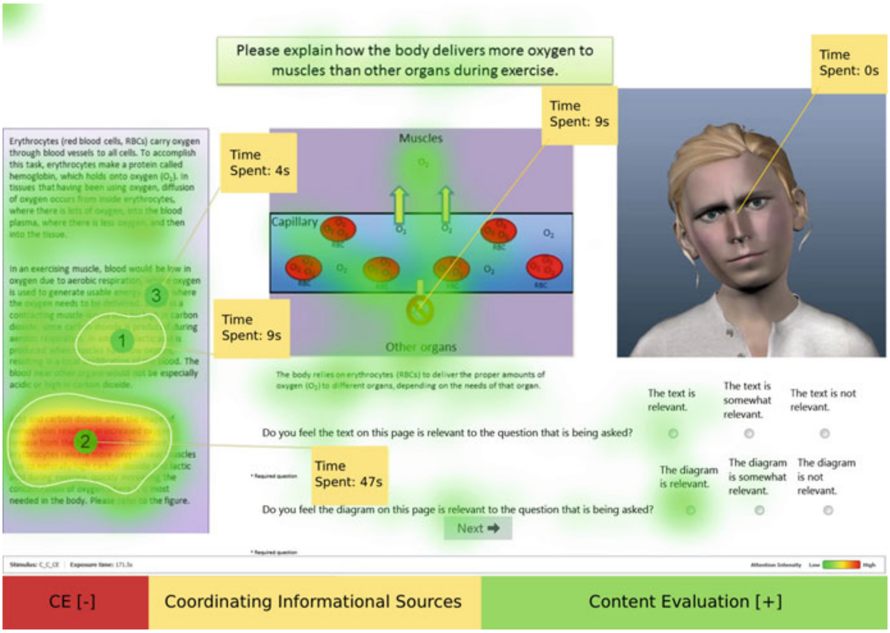

Fig. 10.2 A more complex visualization (similar to Fig. 10.1) of CAMM SRL data that can be used to foster and scaffold emotion regulation, including elements presented in Fig. 10.1 in addition to time spent on different (relevant and irrelevant) areas of interest as well as the time spent on cognitive strategy use and accuracy of strategy use and metacognitive judgments during learning. *Note.* CE = Content Evaluation, a metacognitive judgement related to a representation's relevancy given a learning goal; size of CE bars is representative of the accuracy of metacognitive judgments; size of coordinating informational sources bar (a learning strategy) is representative of the amount of time spent using multiple sources of information (e.g. text, diagram, PA expression)

specific cognitive strategies (e.g., coordinating informational sources) as well as the accuracy of specific metacognitive judgments (e.g., content evaluation) during learning. Implicit connections between cognitive strategies and metacognitive judgments (e.g., coordinating informational sources is associated with content evaluation) are symbolically represented on the same colored bar. This approach allows for ease of interpretation and integration of cognitive and metacognitive monitoring processes, and can be made explicit by having a human or artificial agent do a verbal (cognitive) walkthrough of this entire figure at some point during emotion-regulation scaffolding. Lastly, we emphasize that these CAMM SRL processes and emotion-regulation scaffolding techniques presented in our framework need to be experimentally tested prior to adoption in ALTs to foster successful emotion regulation.

Conclusion and Future Directions

Emotions play a critical role during learning and problem solving with ALTs. We argue that despite their importance, relatively few attempts have been made to understand and foster learners' emotional monitoring and regulation by using data visualizations of their own (and others') CAMM SRL processes since they can potentially foster emotion regulation during learning with ALTs. In this chapter, we presented a theoretically-based and empirically-driven conceptual framework that addresses emotion regulation by proposing the use of one's and others' CAMM-SRL data visualizations (e.g., eye-movement behaviors, facial expressions of emotions, physiological arousal) to facilitate learners' monitoring and regulation of their emotions during learning with ALTs. We proposed several examples of emotion-regulation strategies based on the presentation of multimodal multichannel data to illustrate the mapping between theoretical assumptions, emotion regulation strategies, and the types of data visualizations that can be used to enhance learners' emotion regulation. We expect this line of work to lead to interdisciplinary efforts to address the use of CAMM SRL data visualizations to foster emotion regulation and lead to optimal and successful learning with ALTs. In doing so, we expect advances in educational, learning, cognitive, affective, social, engineering, and computational sciences to contribute immensely with the myriad of issues presented in our chapter. The plethora of conceptual, theoretical, methodological, and analytical challenges (e.g., identifying robust behavioral signatures of emotion flexibility, emotion adaptivity, and emotion efficacy) from multimodal multichannel data will lead to the impact on contemporary research in the areas of cognitive, learning, and affective sciences, human-computer interaction, data visualization, big data, data mining, data science, learning analytics, open learner models, and SRL.

Acknowledgements This chapter was supported by funding from the National Science Foundation (DRL#1431552, DRL#1660878, CMMI#1744351) and the Social Sciences and Humanities Research Council of Canada (SSHRC 895-2011-1006). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or Social Sciences and Humanities Research Council of Canada.

The authors would like to thank Carina Tudela, Mitchell Moravec, Alex Haikonen, and Pooja Ganatra from the SMART Lab at NCSU for their assistance.

References

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2015). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review*, *30*, 217–237.

Antonietti, A., Casellato, C., Garrido, J. A., Luque, N. R., Naveros, F., Ros, E., … Pedrocchi, A. (2015). Spiking neural network with distributed plasticity reproduces cerebellar learning in eye blink conditioning paradigms. *IEEE Transactions on Biomedical Engineering*, *63*, 210–219.

Azevedo, R. (2014a). Issues in dealing with sequential and temporal characteristics of self- and socially-regulated learning. *Metacognition and Learning*, *9*, 217–228.

Azevedo, R. (2014b). Multimedia learning of metacognitive strategies. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 647–672). Cambridge, MA: Cambridge University Press.

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, *50*, 84–94.

Azevedo, R., & Aleven, V. (2013a). Metacognition and learning technologies: An overview of the current interdisciplinary research. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 1–16). Amsterdam, The Netherlands: Springer.

Azevedo, R., & Aleven, V. (Eds.). (2013b). *International handbook of metacognition and learning technologies*. Amsterdam, The Netherlands: Springer.

Azevedo, R., & Hadwin, A. F. (2005). Scaffolding self-regulated learning and metacognition—Implications for the design of computer-based scaffolds. *Instructional Science*, *33*, 367–379.

Azevedo, R., Harley, J., Trevors, G., Feyzi-Behnagh, R., Duffy, M., Bouchet, F., & Landis, R. S. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 427–449). Amsterdam, The Netherlands: Springer.

Azevedo, R., Martin, S. A., Taub, M., Mudrick, N., Millar, G., & Grafsgaard, J. (2016). Are pedagogical agents' external regulation effective in fostering learning with intelligent tutoring systems? In A. Micarelli, J. Stamper, & K. Panourgia (Eds.), *Proceedings of the 13th International Conference on Intelligent Tutoring Systems—Lecture Notes in Computer Science 9684* (pp. 197–207). Amsterdam, The Netherlands: Springer.

Azevedo, R., Taub, M., & Mudrick, N. (2015). Technologies supporting self-regulated learning. In M. Spector, C. Kim, T. Johnson, W. Savenye, D. Ifenthaler, & G. Del Rio (Eds.), *The SAGE Encyclopedia of educational technology* (pp. 731–734). Thousand Oaks, CA: SAGE.

Azevedo, R., Taub, M., Mudrick, N., Farnsworth, J., & Martin, S. A. (2016). Interdisciplinary research methods used to investigate emotions with advanced learning technologies. In M. Zembylas & P. Schutz (Eds.), *Methodological advances in research on emotion and education* (pp. 231–243). Amsterdam, The Netherlands: Springer.

Barrett, L. F., Lewis, M., & Haviland-Jones, J. M. (Eds.). (2016). *Handbook of emotions*. New York, NY: The Guildford Press.

Belland, B., Walker, A., Kim, N., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research*, *87*(2), 309–344.

Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From design to implementation to practice—A learning by teaching system: Betty's brain. *International Journal of Artificial Intelligence in Education*, *26*, 350–364.

Bonanno, G. A., & Burton, C. L. (2013). Regulatory flexibility an individual differences perspective on coping and emotion regulation. *Perspectives on Psychological Science*, *8*, 591–612.

Bosse, T., Gerritsen, C., Man, J. D., & Treur, J. (2013). Learning emotion regulation strategies: A cognitive agent model. In *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)* (Vol. Vol 2, pp. 245–252). Washington, DC: IEEE Computer Society.

Buder, J., & Hesse, F. W. (2017). Informational environments: Cognitive, motivational-affective, and social-affective forays into the digital transformation. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 1–25). New York: Springer.

Bull, S., & Kay, J. (2016). SMILI: A framework for interfaces to learning data in open learner models, learning analytics and related fields. *International Journal of Artificial Intelligence in Education*, *26*, 293–331.

Calvo, R. A., D'Mello, S., Gratch, J., & Kappas, A. (Eds.). (2015). *The Oxford handbook of affective computing*. Oxford, England: Oxford University Press.

D'Mello, S. K. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology*, *105*, 1082–1099.

D'Mello, S., & Graesser, A. (2012). AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems*, *2*, 1–39.

D'Mello, S. K., & Graesser, A. C. (2015). Feeling, thinking, and computing with affect-aware learning technologies. In R. A. Calvo, S. K. D'Mello, J. Gratch, & A. Kappas (Eds.), *The Oxford handbook of affective computing* (pp. 419–434). New York, NY: Oxford University Press.

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170.

Doré, B. P., Silvers, J. A., & Ochsner, K. N. (2016). Toward a personalized science of emotion regulation. *Social and Personality Psychology Compass*, *10*, 171–187.

Ekman, P. (1977). Facial expression. In A. Siegman, & S. Feldstein (Eds.), Nonverbal Communication and Behavior (pp. 97–126). New Jersey: Lawrence Erlbaum Association.

Forbes-Riley, K., & Litman, D. (2011). Benefits and challenges of real-time uncertainty detection and adaptation in a spoken dialogue computer tutor. *Speech Communication*, *53*, 1115–1136.

Graesser, A. C. (2016). Conversations with AutoTutor help students learn. *IJAIED*, *26*, 124–132.

Graesser, A. C., Lippert, A., & Hampton, D. (2017). Successes and failures in building learning environments to promote deep learning: The value of conversational agents. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 273–298). New York, NY: Springer.

Gratch, J., & Marsella, S. (2015). Appraisal models. In R. A. Calvo, S. K. D'Mello, J. Gratch, & A. Kappas (Eds.), *Handbook of affective computing* (pp. 57–67). Oxford, UK: Oxford University Press.

Grawemeyer, B., Mavrikis, M., Holmes, W., Gutierrez-Santos, S., Wiedmann, M., & Rummel, N. (2016). Affecting off-task behaviour: How affect-aware feedback can improve student learning. In *Proceedings of the 6th Int. Conference on Learning Analytics & Knowledge* (pp. 104–113). New York, NY: Association for Computing Machinery (ACM).

Gross, J. J. (2015a). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, *26*, 1–26.

Gross, J. J. (2015b). The extended process model of emotion regulation: Elaborations, applications, and future prospects. *Psychological Inquiry*, *26*, 1–26.

Harley, J. M., Bouchet, F., Hussain, S., Azevedo, R., & Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior*, *48*, 615–625.

Harley, J. M., Poitras, E. G., Jarrell, A., Duffy, M. C., & Lajoie, S. P. (2016). Comparing virtual and location-based augmented reality mobile learning: emotions and learning outcomes. *Educational Technology Research and Development*, *64*(3), 359–388.

Kappas, A. (2013). Social regulation of emotion: Messy layers. *Frontiers in psychology*, *4*, 1–11.

Lafford, B. A. (2004). Review of tell me more Spanish. *Journal on Language Learning and Technology*, *8*, 21–34.

Litman, D., & Forbes-Riley, K. (2014). Evaluating a spoken dialogue system that detects and adapts to user affective states. In *Proceedings of the 15th Meeting of the Special Interest Group on Discourse and Dialogue* (pp. 181–185).

Mayer, R. (2014). Multimedia instruction. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (4th ed., pp. 385–400). Amsterdam, The Netherlands: Springer.

Muldner, K., Wixon, M., Rai, D., Burleson, W., Woolf, B., & Arroyo, I. (2015). Exploring the impact of a learning dashboard on student affect. In C. Conati, N. Heffernan, A. Mitrovic, & M. Verdejo (Eds.), *Artificial intelligence in education* (pp. 307–317). Heidelberg, Germany: Springer.

Nye, B. D., Graesser, A. C., & Hu, X. (2014). Multimedia learning with intelligent tutoring systems. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (3rd ed., pp. 705–728). Cambridge, MA: Cambridge University Press.

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review*, *24*, 27–45.

Pekrun, R., & Linnenbrink-Garcia, L. (Eds.). (2014). *International handbook of emotions in education*. New York, NY: Routledge.

Porayska-Pomsta, K., Mavrikis, M., D'Mello, S., Conati, C., & Baker, R. S. (2013). Knowledge elicitation methods for affect modelling in education. *International Journal of Artificial Intelligence in Education*, *22*, 107–140.

Reeck, C., Ames, D. R., & Ochsner, K. N. (2016). The social regulation of emotion: An integrative, cross-disciplinary model. *Trends in Cognitive Sciences*, *20*, 47–63.

Sabourin, J., & Lester, J. (2014). Affect and engagement in game-based learning environments. *IEE Transactions on Affective Computing*, *5*, 45–56.

Santos, J. L., Govaerts, S., Verbert, K., & Duval, E. (2012). Goal-oriented visualizations of activity tool tracking: A case study with engineering students. In S. Buckingham Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge LAK'12* (pp. 143–152). New York, NY: ACM.

Scheiter, K., Fillisch, B., Krebs, M., Leber, J., Ploetzner, R., Renkl, A., … Zimmermann, G. (2017). How to design adaptive multimedia environments to support self-regulated learning. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 203–233). New York, NY: Springer.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotions*, *7*, 1307–1351.

Schunk, D. H., & Greene, J. A. (Eds.). (in press). *Handbook of self-regulation of learning and performance* (2nd ed.). New York, NY: Routledge.

Suri, G., & Gross, J. J. (2016). Emotion regulation: A valuation perspective. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (4th ed.). New York, NY: The Guildford Press.

Taub, M., Mudrick, N., Azevedo, R., Millar, G., Rowe, J., & Lester, J. (2016). Using multi-level modeling with eye-tracking data to predict metacognitive monitoring and self-regulated learning with Crystal Island. In A. Micarelli, J. Stamper, & K. Panourgia (Eds.), *Proceedings of the 13th ITS Conference. Lecture Notes in Computer Science 9684* (pp. 240–246). The Netherlands: Springer.

Taub, M., Mudrick, N. V., Azevedo, R., Millar, G. C., Rowe, J., & Lester, J. (2017). Using multi-channel data with multi-level modeling to assess in-game performance during gameplay with CRYSTAL ISLAND. *Computers in Human Behavior*.

Van Kleef, G. A. (2009). How emotions regulate social life the emotions as social information (EASI) model. *Current Directions in Psychological Science*, *18*, 184–188.

van Kleef, G. A. (2016). *Toward an integrative theory of emotions as social information*. Cambridge, MA: Cambridge University Press.

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, *57*, 1500–1509.

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Assche, F., Parra, G., & Klerkx, J. (2013). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing*, *18*, 1499–1514.

Winne, P. H., & Azevedo, R. (2014). Metacognition. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 63–87). Cambridge, England: Cambridge University Press.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Hillsdale, NJ: Erlbaum.

Winne, P. H., Vytaske, J., Patzak, A., Rakovic, M., Marzouk, A., Pakdaman-Sajovi, A., … Nesbit, J. (2017). Designs for learning analytics to support information problem solving. In J. Buder & F. W. Hesse (Eds.), *Informational environments: Effects of use, effective designs* (pp. 249–272). New York, NY: Springer.

Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: Recognizing and responding to student affect. *International Journal of Learning Technology*, *4*, 129–164.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2011). *Handbook of self-regulation of learning and performance*. New York, NY: Routledge.

Chapter 11
Designs for Learning Analytics to Support Information Problem Solving

Philip H. Winne, Jovita M. Vytasek, Alexandra Patzak, Mladen Rakovic, Zahia Marzouk, Azar Pakdaman-Savoji, Ilana Ram, Donya Samadi, Michael P.C. Lin, Arita Liu, Amna Liaqat, Nashwa Nashaat-Sobhy, Zahra Mozaffari, Jason Stewart-Alonso, and John C. Nesbit

In recent years, learning science researchers have turned with growing interest toward investigating information problem solving (IPS). Problems are situations in which people identify a goal, although not always fully and precisely, and are uncertain how to achieve it. Information problems are a class of problems in which people are challenged to identify, find, examine, use, and communicate information to learn about a new topic, create a product, or solve an overarching problem. IPS can be carried out alone or in collaboration.

In the post-secondary setting, a nearly universal assignment that calls for IPS is the term project. It has several common forms: a major paper (10 pages or more) assigned in a course, a business plan, or a proposal for scholarly research or a thesis. Students almost universally engage in information problem solving when they tackle term projects throughout their academic career. Professors and mentors intend these exercises in information problem solving to substantially extend students' knowledge and deepen understanding about the topic of the project, e.g., how polling affects modern political elections or whether a carbon tax effectively controls pollution. To denote this purpose explicitly, we label these large, complex assignments as learning projects.

P.H. Winne (✉) • J.M. Vytasek • A. Patzak • M. Rakovic • Z. Marzouk • A. Pakdaman-Savoji
I. Ram • D. Samadi • M.P.C. Lin • A. Liu • Z. Mozaffari • J. Stewart-Alonso • J.C. Nesbit
Faculty of Education, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada
V5A 1S6
e-mail: winne@sfu.ca

A. Liaqat
University of Toronto, Toronto, ON, Canada

N. Nashaat-Sobhy
Universidad San Jorge, 50830 Villanueva de Gállego, Zaragoza, Spain

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_11

By synthesizing Brand-Gruwel, Wopereis, and Walraven's IPS-1 model (Brand-Gruwel, Wopereis, & Walraven, 2009) and Winne and Hadwin's model of self-regulated learning (Winne, 2017a; Winne & Hadwin, 1998), we describe round-trip IPS in a learning project as including a variety of specialized problem solving activities (cf. Eisenberg, 2008):

- Setting an overall goal that identifies key attributes of the product to be generated by engaging successfully in a learning project,
- Defining and setting parameter values for kinds of information that contribute to the learning project,
- Searching for and filtering information sources (websites, documents, videos) that are judged to contribute high-quality information needed to meet goals,
- Analyzing and extracting information from filtered sources and organizing that content,
- Designing and drafting a product (e.g., the term paper or research proposal), and
- Evaluating and revising the draft(s) of that product to produce a polished final version.

Beyond learning about a topic, we believe mentors intend learning projects to provide their students opportunities to develop generalizable knowledge and skills for solving information problems. Many label these as lifelong learning skills. Alongside "cold" knowledge and skills used to work directly on solving information problems and managing workflow, "hot" features are infused in IPS (e.g, Schwarz, 2002; Wyatt et al., 1993). These include strategies students use to motivate themselves and cope with the demands of IPS tasks, and strategies for managing emotions, particularly obstacles or setbacks arise. We are quick to add an assumption: Students and their mentors sometimes, perhaps often, do not explicitly or deliberately attend to warmer features of learning projects.

At the start of the first author's academic career, nothing like today's array of tools was available to support practically instant search for sources of topical information or search for particular information within sources. It was tedious rather than simple to edit drafts. The upshot of a round of editing rarely preserved changes that could be easily undone. After editing, the manuscript was left in a state that hardly allowed studying it to discern how changes affected its architecture, clarity, and cogency. Sharing drafts with peers to gather their constructive critiques required meeting them, then meeting again to retrieve their commented copies. Coordinating their commentaries and editorial suggestions across multiple paper copies was unduly time consuming and, with more than two peers providing reviews, a trial of organizational skills. This environment was poorly structured to promote learning. "Getting to work" demanded effort and robbed time that might have been allocated to learning.

Today, students' opportunities to engage in all the multiple facets comprising round-trip IPS are dramatically different. The advent of modern computing technologies, the Internet, and search engines significantly facilitate locating sources, workflow, information sharing and organization, and revisions of draft products.

These are steps toward an effective design for learning. At the same time, however, these affordances generate new information problems. Examples are a surfeit of available information that can appear relevant, tools that allow for tailored and potentially excessive cataloging and tagging of a very large volume of snippets of information, and systems that gather but don't help organize peers' reviews of draft reports. In this newer world of digital "info-wealth," learners need to develop and apply skills to cope in new ways. Successful learners are productive self-regulated learners. Can technologies provide assistive adaptivity to help?

A Framework for Learning Analytics about Information Problem Solving

Alongside the upsurge in research on IPS and its constituents, recent years have seen skyrocketing interest in and research on search for meaningful patterns in data—analytics—and the specialization of analytics about learning—learning analytics. As is common in emerging fields, there is in the field of learning analytics abundant variety in kinds of data gathered, techniques applied to analyze data, and proposals for assembling and formatting reports to learners about patterns in data that describe learning and work constituting IPS.

What distinguishes learning analytics from conventional feedback intended to guide learning? We propose three major but not exclusive characteristics (see also Winne, 2017b). First, compared to an achievement test score or the average of responses to items on a self-report survey gauging motivation, learning analytics are commonly constructed using data about learning activities in conjunction with conventional data. Second, while by no means universal, learning analytics commonly are generated across a timeline of activities rather than data gathered at just one point in time. Third, these two features lend emphasis to the purpose of learning analytics as being formative. That is, learning analytics are intended to guide change(s) in a process to bring about a desired change in a product, such as a learning project. In these three ways, learning analytics are well positioned to help learners become better at solving information problems.

In this chapter, we frame learning analytics within the context of a modern approach to solving information problems, namely, when the student uses a state-of-the-art software system that can capture very fine-grained, time-stamped information about the learner's work on a learning project within an environment of practically limitless information. We leverage affordances of this technology as a tool for supporting learning-controlled assistive adaptivity by gathering, analyzing, and delivering reports about students' work as they engage with common information problems (e.g., see Marzouk et al., 2016). We adopt a perspective of learners as self-interested, self-regulated learners who seek better and minimally taxing ways to solve information problems (Winne, 2017b; Roll & Winne, 2015).

The learning analytics we sketch in this chapter are designed to serve two purposes. The first is to identify particular features of students' work in a learning project that, if the learner adapted them, would enhance learning about the topic the learner researches. The second is providing process feedback that guides self-regulated learning about skills for successfully addressing information problems. In this vein, we extend common frameworks for a learning analytic by adding a requirement that data gathered and interpretations of analytics generated with that data align to one or several empirically-supported principles in learning science. We believe mining empirical research in learning science will increase the likelihood learning analytics can help learners build topic knowledge and improve skills for addressing information problems.

Context: Researching a Term Paper Using nStudy

Suppose a learner has been assigned a learning project: Research and develop a 10-page paper arguing for or against this contentious proposition: "A national consumption (e.g., sales) tax is a better option for funding a nation's priorities than a tax on its citizen's wages (e.g., an income tax)." Our learner—we name her Mia—works on this learning project using state-of-the-art software called nStudy. nStudy is an extension to the Chrome web browser. It records data about every bit of information Mia operates on when she uses nStudy's features, how Mia operates on that information and the time at which each operation occurs accurate to approximately 1/100th s. Metaphorically, these data form a script that, if "played back," completely and exactly reveals what a person hovering over Mia's shoulder could observe about how Mia works on the learning project except for Mia's self-talk that she elects not to enter in nStudy as a "note to self." As Mia works, these data are transmitted in approximately real time to a remote server where they are stored in a database. Modules on the server combine a database query that extracts particular data that are then processed by scripts that analyze those data using tools within the statistical/analytic framework R. Output generated by those modules is then formatted using other R packages to create analytics as a web page. URLs pointing to those web pages are delivered to Mia when she asks for them or by the system when attributes describing the state of Mia's work match conditions in rules of a production system. In the latter case, the production system plays the role of a dynamic help system that monitors in nearly real time conditions that signal Mia might benefit from learning analytics reports about how she is working and what information is processed by her work activities.

nStudy

When learners use nStudy's tools, they create artifacts that nStudy stores for them and trace data that describe how information in an artifact was processed. Traces are operational definitions of cognition, metacognition, and motivation. Typing a URL into the browser's address field or clicking an existing bookmark in nStudy's library of artifacts logs the URL of the sought information in nStudy's database. This trace datum identifies Mia is seeking information, she predicts information she seeks is available in a particular artifact, and is motivated to inspect the web page (or pdf file) that will be displayed in her web browser. Metacognitively, the trace indicates Mia judges information she cannot recall well enough or judges she does not know exists in the particular web page that will be displayed in the web browser.

Dragging the cursor across a string of text in a web page selects it and automatically pops up a menu where Mia can create a quote, a note, or a term. The text Mia selects is automatically copied by nStudy into a sidebar for future access; hence, we label it a quote. If Mia chooses, she can tag any instance of quoted text by typing a tag (or several tags) into a text box. As she types, a dropdown list of existing tags is displayed, reminding Mia of tags already created. The dropdown list of tags is successively filtered as Mia types additional letters into the text box.

A note elaborates a quote. Each note has a text field where Mia identifies the topic of the note, a second text box where she can tag the note independently of the quote she may have applied to the quote associated with a note, and at least one labeled field for the note itself. Notes can have more complex designs called a note template. A note template can be basic, e.g., just a single text field labeled "Comment." Or a note template can represent a larger schema, e.g., four text fields providing a note that represents a causal relation: "Cause," "Effect," "Context," and "Reason." A variety of types of fields can be configured to design multifaceted note templates that include text fields, lists of checkbox items (select one or many), lists of radio button items (select only one), sliders, date, image, and a link field where Mia can create a hyperlink joining the note she is currently making to another artifact, e.g., a bookmark representing a different web page. A note template has a name, e.g., "quality of expression" or "reasoning." nStudy automatically tags each note Mia creates with the name of the note template she chose for annotating text she quoted.

A term note is a special note template for creating entries in a glossary of key concepts. The topic field of a term note contains the concept being described. A "Description" text box holds the concept's description or definition. A link field lists other concepts associated with this term based on a relationship. One kind of relationship is the in-terms-of relation. It is automatically created by nStudy when one term's description includes another term. Key concepts and terms in sources relate in ways beyond definitions. Under the assumption that authors of sources construct conceptually coherent and valid sentences, a second relation is created when terms co-occur in a sentence.

A display of key concepts and their relationships to one another can be pictured in a node-link graph called a termnet. Terms are vertices in this graph and links (edges) are defined by either or both the in-terms-of relation and sentence co-occurrence relations. As a mathematical graph, the termnet allows indexes to be calculated that describe, e.g., the "importance" of a concept or the complexity of a neighborhood of concepts.

Quotes, notes, and terms are copied into a sidebar associated with the web page in which these artifacts were created. Artifacts in the sidebar can be filtered by typing in a text field; e.g., typing the tag "vague" filters out all artifacts not tagged "vague." Each artifact is also represented by a colored nub adjacent to the scroll bar at the right edge of the window. Clicking an artifact in the sidebar or a nub scrolls the web page to the location of the quote associated with that artifact and opens a window over the web page showing the artifact's contents in context.

nStudy also provides an HTML editor Mia can use to draft her essay. The essay tool provides a toolbar with common formatting features (e.g., italic, bold, levels of headings, bullet list, sequential list) to format the essay. As well as typing directly to draft her essay, Mia can copy an artifact, e.g., a quote or note, and paste its information into the essay. Because the essay is actually a web page, Mia (and peers, see below) can annotate it by identifying text (quoting it), and tagging and annotating it.

Each time Mia selects text, types text, or clicks a "live" feature associated with an artifact, nStudy records a 3-tuple: text, operation, time. For example, if Mia filters the sidebar by typing the tag "vague," the data recorded are a complete representation of Mia's engagement.

nStudy's collaborative tool, called the Hub, is a channel for Mia to communicate with peers, share learning artifacts and give and receive feedback on her learning project in the form of peers' reviews and learning analytics. Discussions are supported by a dropdown list that suggests roles she and peers can adopt, e.g., critic, analyst, or manager. To guide Mia's participation in discussions, nStudy offers prompts keyed to each role. For example, in the critic role, sample prompts might be: "Is the evidence for ___ reliable?" or "Is the view that ___ corroborated?" Clicking a prompt inserts it into the text box where Mia can fill in blanks and add additional information before sending her contribution to the discussion. Mia can also drag and drop artifacts, such as a tagged bookmark, a note, or her draft essay into the message box to share with peers.

Empirically-Grounded Learning Analytics About Learning Projects

Students like Mia face many challenges as they work on a learning project, such as the example we introduce here of researching and developing an argument about a relatively new topic. Here, we elaborate just three: planning and learning how to plan more accurately, monitoring whether a draft essay has appropriate coverage of a topic, and developing new knowledge about a topic. For each challenge, we describe: the challenge itself, data nStudy gathers as the learner works, aims served by a learning analytic developed with that data, a learning analytic report for presentation to the learner, and a rationale from learning science that provides grounds for predicting the learner can benefit by maintaining or adapting features of work on the learning project as described by the learning analytic.

Challenge 1: Planning and Learning How to Plan More Accurately

In researching material for a learning project, students need to search for relevant information, study sources they locate, extract and synthesize content, and develop an understanding of a topic to form a conceptual foundation for generating a product. In the context of these IPS tasks, students engage in various metacognitive tasks, e.g., estimating time required for subtasks, planning a path through the space of the information problem, monitoring progress toward goals, and adjusting time and foci for planned efforts as work on the learning project proceeds. These are known challenges for many learners (Buehler, Griffin, & Ross, 1994; Kruger & Dunning, 1999; Winne, 2001; Zimmerman, Bonner, & Kovach, 1996). We describe an nStudy feature in development, the PlanNet, to help learners address these challenges.

nStudy Data

As Mia searches the Internet for raw materials she will mine for her essay, nStudy traces her search queries. Among sources she finds, she bookmarks selections in her nStudy library. Mia can tag a bookmarked source, and content in a source can be annotated by quotes, tagged or not, and notes, tagged or not. In addition, Mia can create terms. As described earlier, throughout all her work on the learning project, nStudy logs data about Mia's exposure to terms in sources she views and her use of terms in the text of artifacts she creates: searches, quotes, notes, terms, and her essay.

The PlanNet (see Fig. 11.1.) doubles as a tool for developing plans and an interactive analytic. As Mia uses its features to customize a PlanNet display, nStudy logs data identifying bookmarked sources she cites in the essay she is developing, values that mark time points in time within plans and for reviewing plans, and metadata about terms displayed in a termnet or that are targeted for planned action. The metadata about terms identify whether Mia selects a term (a node) and how Mia chooses to operate on it (e.g., to review it, to search for it). As the state of the PlanNet changes, metadata about terms are updated to identify whether, at each state of Mia's work, a term is a member of the subset of terms defined by Mia's configuration of PlanNet features. For example, if Mia selects a single bookmark as the scope for the PlanNet display, nStudy logs data that identifies whether each term in that selected source appears in the draft of Mia's essay as of a date (version) she specifies for the essay's state.

Conceptual Description of Learning Analytics

The PlanNet window shown in Fig. 11.1. has several features. In the upper left is a list of artifacts. The first is a dropdown list identifying versions of Mia's essay.

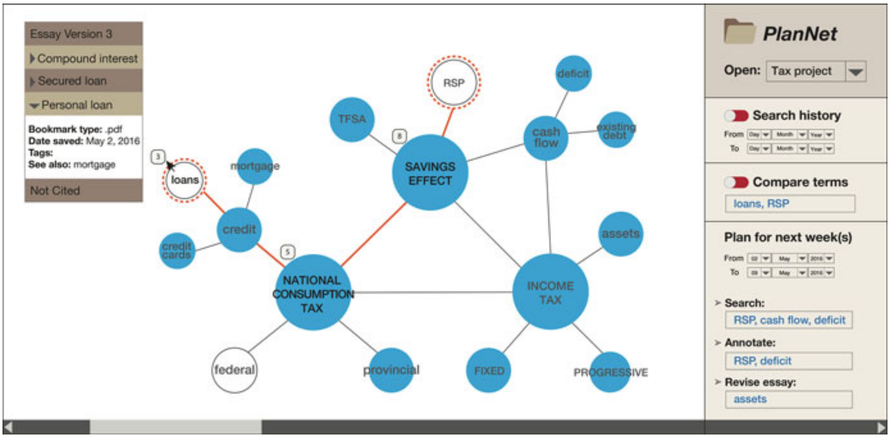

Fig. 11.1 PlanNet for Mia's Project

Mia can select a particular version to examine in the PlanNet. Next listed is the subset of sources Mia bookmarked while researching her topic, which she has cited in the selected version of her essay. The final entry in the list, "Not Cited," is a folder. Clicking it lists sources Mia bookmarked while researching her topic but are not cited in the selected version of her essay. Selecting an artifact in this list (or multiselecting several sources) marks terms in the termnet shown in the main area of the window by changing color codes identifying how a term is used, as discussed next.

In the center of the PlanNet window is a termnet. It is a union of the "seed" termnet Mia's instructor provided to help students begin their learning project with the complete set of terms Mia created while researching her topic. Several features are found in the termnet:

- Node diameter is proportional to a term's importance relative to all sources Mia bookmarked, i.e., the greatest scope of information Mia judged relevant in her research about the topic. Importance of a term is measured by an index of that term's centrality relative to all other terms across Mia's bookmarked sources (Dehmer & Emmert-Streib, 2009).
- Distances separating nodes (the lengths of nondirectional edges in the graph) are inversely proportional to linked terms' co-occurrences within sources and across them (Manning, Raghavan & Schütze, 2008a, 2008b). While simple, representing text this way highlights useful properties of information (Hopkins & King, 2010). To generate a readable termnet within a confined space, the ForceAtlas2 algorithm (Jacomy, Venturini, Heymann, & Bastian, 2014) is applied. This pushes highly connected nodes (hubs) away from each other and arrays nodes associated with a hub in clusters around hubs. If Mia selects any

two terms, nStudy applies coloring to links that form the shortest path between them.

- Terms “seeded” by Mia’s instructor appear in all uppercase font.
- Font is black for terms appearing in bookmarked sources Mia cites in the current version of her essay. Font color is gray for other terms unique to sources Mia is not citing in the current version of her essay.
- When Mia selects her Essay in the list of artifacts, nodes representing terms included in the current version of her essay are filled with light blue.
- When a source other than the Essay is selected, the weight of the perimeter stroke for nodes representing terms in the source increases +2.

Nodes in the termnet are decorated according to metadata. Numbered bubbles adjacent to nodes mark the number of bookmarked sources Mia used in her essay in which the term appears. Clicking a bubble lists the titles Mia assigned when she created those bookmarks. Decorations in the termnet help Mia quickly identify important qualities of information in her essay relative to one or several sources cited in it. Selecting different sources redecorates the termnet to highlight differences in concepts the source(s) describe.

Mia can quickly review information about a term by double-clicking a node. This opens a term window showing the term, the definition/description of that key concept, and a list of bookmarks in which the term appears. Clicking a bookmark opens it for quick review to support a scan for specific information related to that concept or extended restudy.

To support Mia’s exploration of “what if” scenarios, i.e., plans, Mia can use PlanNet tools in the right-hand sidebar to adjust the termnet display. As she does, the termnet is redecorated to reflect her plan cast as a configuration of key concepts. First, Mia identifies an interval when she will work to implement a plan selecting start and due dates. She then might drag-and-drop a node from the termnet into the Search box. This changes the fill color of that node and its radiating neighbors to a shade of blue using heatmap shading. Intensity of the fill color fades as terms become more distant from the focal term. This invites Mia to (re)consider information that might be added to or elaborated through her search and subsequently in her essay. Mia can operationalize a plan to expand her understanding of a term by dragging it into the Annotate box, thus marking a term for further study. If the redecorated termnet is sufficient to influence Mia to revise her essay in relation to a term, she can drag it into the Revise Essay box. When “done” with work relating to a term, Mia can place it in the Completed box.

Mia’s interactions with the PlanNet are fully traced. Data record how she explores conceptual configurations represented by redecorated termnets, plans she devises for revising essays and how this work unfolds in relation to the time marks she sets for work on a plan.

Mia also needs to check she understands the instructor’s seed terms, and plan how to integrate those concepts in her essay as she searches and reviews sources. When uppercase font terms are filled in blue, this goal is met. For these terms, and others she decides are important to her essay, a plan can be developed using the PlanNet's features. nStudy can remind her which terms in her plan have not yet received attention as she planned, and whether she is on schedule according to the plan interval she set. Displays contrasting a simple timeline of work planned to work accomplished, e.g., expanding annotations about a term, inform Mia about whether she might increase or reduce workload in a plan interval. By toggling the drop-down time range to "on" and adjusting dates in the Search history section of the PlanNet sidebar, Mia can compare progress on her project as reflected by work planned and work accomplished. As an analytic, data collected in the PlanNet and elsewhere in nStudy explicitly represent conceptual development as a function of goals, type of work, and time. Mia can review how earlier plans were defined, track what she accomplished relative to plans, and compare the state of her project now to its forecasted forms.

Learning Analytics Reports and Learning Science Principles

Based on Mia's interaction with the PlanNet, nStudy generates five learning analytics reports that are embedded within the PlanNet. This design builds on a rich history of two-dimensional graphic organizers as visual representations that support students' use of knowledge (Alvermann, 1981; Ives & Hoy, 2003; Winn, 1991). A meta-analysis by Nesbit and Adesope (2006) compared concept maps to other learning activities (e.g., reading texts, attending lectures, and class discussion participation). They reported concepts maps were more beneficial for promoting retention and transfer of information across various learning settings, topics, and learners' educational level. They conjectured these benefits resulted from greater opportunities for engaging with information in a concept map compared to other media and learning activities.

Analytic Report 1: Use of Terms. Decorations in the PlanNet are analytics that afford Mia opportunities to metacognitively monitor features of key concepts (terms) related to the learning project. Numbers in bubbles adjacent to terms reflect the scope of sources Mia has bookmarked about key concepts. Low numbers or the absence of bubbles signal limited or no search, a nudge to plan additional research into these concepts. Terms in black font (included in Mia's essay) relative to terms in gray font (not included) indicate how well Mia's essay spans the complete set of concepts, including those her instructor marked as critical (uppercase).

These analytics are grounded on the coherence effect (25 learning principles, 2015): learning benefits from a connected and articulated representation of main ideas in material (Mayer, 2001). The termnet is such a representation. Also, vocabulary knowledge influences strategies learners choose to study, inferences they form from source materials and, ultimately, achievement (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010). For students pursuing learning projects where content and language learning are integrated (CLIL), word choice and phrasing affect the register and quality of writing in the product of a learning project (Whittaker, Llinares, & McCabe, 2011).

As Mia extends her knowledge of key concepts, support provided by this interactive analytic may increase. Gurlitt and Renkl (2008) reported university students benefited more from low map coherence where they had to create and label links. In the termnet, as Mia adds concepts to those her instructor provided, she creates and can subsequently edit descriptions of how they relate conceptually. As her prior knowledge grows over the timeline of her learning project and as she actively engages with the termnet of concepts, she successively records elaborations of her knowledge via the termnet.

Analytic Report 2: Mapping Term Relationships. The PlanNet's interactive analytics mark relationships among terms Mia has not used yet in her essay. To explore how she might introduce new content given what Mia has included in her essay, she can select two or more terms in the termnet and PlanNet will map the shortest path between (among) them. This analytic is based on Mayer's (2001) notion of actively engaging with information to form a coherent mental representation or mental model. Here, terms are key, conceptually structured items that Mia is modeling in her essay.

If Mia elects to preserve a particular set of conceptually related terms, she can save the artifact in the sidebar under "Compare Terms." Clicking it at a later time shows Mia how terms in the set relate by decorating the paths that link them. Like other artifacts, Mia can annotate a conceptual relationship, e.g., creating a "superordinate" term or a note that elaborates the conceptual structure. According to Bransford and Schwartz (1999), comparing conceptual structures, represented via the PlanNet as different conceptual structures, can help novices notice deep features that otherwise may be unnoticed. Here, Mia may observe new relationships involving a particular term or neighborhoods of terms.

Mia controls how she uses the termnet, how many sets of terms she explores, and the span of each conceptual structure. This aligns to the Goldilocks principle (25 learning principles, 2015): learning improves when the scope of information is tailored to current abilities.

The termnet also implements the Organization Effect, which emphasizes the importance of integrating and synthesizing information (25 learning principles, 2015). As Mia adds terms and annotates them, and when she traces coverage of her searches and her draft essay, she is actively and constructively (Chi, 2009) engaged with content.

Analytics Report 3: Comparing a Draft Term Paper and PlanNet. Mia can select the current or a past version of her essay in the dropdown list at the upper left corner of the PlanNet. This draws a termnet representing the content in that version of the essay over a backdrop formed by the termnet of all terms across all of Mia's instructor's terms, terms Mia added, and her bookmarked sources. This may help Mia monitor deeper features of information in the drafts of her essay relative to neighborhoods in the domain and the entire domain of her topic (Bransford & Schwartz, 1999).

Several studies reported mapping and graphical planning strategies are an effective pre-writing activity to promote students' understanding of text organization (Cronin, Sinatra & Barkley, 1992; Pieronek, 1994; Saddler, Moran, Graham &

Harris, 2004; Schultz, 1991; Washington, 1988). These activities correlate with beliefs of control over concepts and the writing task (Bascones & Novak, 1985; Novak, 1991, 1998). Reynolds and Hart (1990) reported concept maps were stronger supports for planning writing than traditional outlining or brainstorming.

As well, planning by mapping is also reported to help learners manage memory (Novak, 1990; Novak & Wandersee, 1991), increase vocabulary (Harley, Howard, & Roberge, 1996; Johnson & Steele, 1996), and promote reading comprehension (Baumann & Bergeron, 1993; Lipson, 1995) and comprehension in the subject area (Patterson, 2001; Roth, 1994). All of these effects contribute to meaningful learning (Ojima, 2006).

Analytics Report 4: Feedback About Planning. Productively self-regulated learners forge plans about how they will learn (Winne, 2017b; Winne & Hadwin, 1998). They monitor progress and adjust the strategy or the goal as needed, and set a new subgoal when the current one is attained (Schunk, 2001). When a work session is complete, they analyze them to generate forward-reaching transfer that will improve plans they will make in the future (Winne, 2017a).

Learners find it difficult to determine progress toward a distant goal (Schunk, 1995); proximal, short-term goals are achieved more quickly, with greater motivation and better self-regulation compared to long-term goals (Bandura, 1997; Boekaerts et al., 2000; Locke & Latham, 1990). Thus, the PlanNet is designed to invite students to divide large learning projects—the full termnet—into subgoals expressed as self-set, short- to moderate-term, concept-explicit goals about searching for key concepts, developing understanding about them through annotations, and displaying what has been understood by using terms in the product of their learning project.

In the Plans sector of the PlanNet sidebar, Mia can enter or drag-and-drop terms from the termnet to identify a focus for new searches, new annotations, and revisions to her essay. She also can define a time interval within which this work should occur. Each configuration of these features operationally defines a subgoal relating to Mia's learning project.

Using data nStudy gathers in the PlanNet and as Mia uses its other features, analytics can be generated that describe characteristics of Mia's plans as well as how well she matches actual work to her plans. For example, students often underestimate time to complete a project (e.g., the planning fallacy; Buehler et al., 1994; Kahneman & Tversky, 1979). nStudy can track and report about whether Mia searches for and annotates terms identified in the plans sector of PlanNet. In the essay tool, nStudy can track whether Mia edits or adds sentences or paragraphs that contain terms she targeted for revision. Comparing plans to actual work provides Mia with raw data she needs to consider in how to set goals and plan in the future (Anderman & Wolters, 2006).

Regulating work along a timeline is also an important component of effective planning. When Mia sets a time interval for a plan regarding specific concepts and particular actions applied to them (e.g., developing annotations about the concept of deficit), nStudy can post alerts when deadlines approach or are missed. A timeline contrasting deadlines to when work of particular kinds on particular concepts was begun can invite Mia to consider whether and why she procrastinates (Patzak & Winne, 2016).

Analytics Report 5: Planning History. If Mia sets dates in the Plan sector of PlanNet's sidebar to values in the past, earlier plans are retrieved and attributes of terms in the termnet that relate to that plan are set to their retrospective values. By "paging through" a timeline of plans, Mia can inspect their temporal and conceptual scope to develop an evidence-based perception about how well she plans. This sets a stage for more accurate planning and time management (Kruger & Dunning, 1999) that characterize self-regulating learners (Winne, 2017a; Winne & Hadwin, 1998).

Challenge 2: Learning about the Topic

When learners research a topic about which they have limited or only basic knowledge, we predict they are challenged to accurately characterize what they have learned about the topic and how well the product they generate—Mia's essay—represents the topic.

With respect to a learner's judgment of what has been learned about a topic, much research suggests learners overestimate how much they know and how well they know (can recall) it, and, moreover, may not benefit from repeated feedback about their overconfidence (e.g., Foster, Was, Dunlosky, & Isaacson, 2016).

Regarding how aptly a product represents the topic, we suggest three "views." First, how thoroughly does the product—Mia's essay—sample information available in sources she identified as relevant to her learning project? Second, is information sampled in her essay configured like information in the sources? Third, how well does the profile of emphasis in Mia's essay align to that in sources she studied? If learners can more accurately judge these matters, results of studies such as Cerdán and Vidal-Abarca's (2008) indicate they develop a more extensive and deeper understanding of the topic.

nStudy Data

As Mia studies source documents, she selects text representing key concepts and terms that play critical roles in her topic and creates term notes. Terms Mia creates are added to those seeded by her instructor's termnet.

Conceptual Description of Learning Analytics

Termnets can be displayed for a single bookmarked source, across the corpus of sources Mia studied, and in Mia's essay. Indexes can be computed to describe quantitative properties of terms and collections of terms in a termnet. For example, the number of terms associated with a focal term (the centrality index) can suggest the conceptual importance of a term in a different manner than counting the frequency of a term's appearance as an indicator of emphasis given that concept in a source or in Mia's essay. A betweenness centrality index gauges the degree to which a term plays the role of a key mediator between clusters (neighborhoods) of other concepts.

With the termnet as a backdrop, basic questions can be framed about particular nodes and neighborhoods of nodes. These can guide Mia to more accurately estimate whether and what she has learned about the topic she researched, setting a stage for restudying material she hasn't learned well. Creating a termnet of terms appearing in the sources Mia cites in her essay and contrasting it to another termnet of terms Mia has used in her essay sets a stage for learning analytics describing content selection and topical coverage in the essay relative to that in the sources.

Learning Analytics Report and Learning Science Principles

After studying several source documents and creating a glossary of terms appearing in them, Mia drafts her essay. She then asks nStudy for an analysis of her essay that compares it to sources she cited in the essay. Although Mia did not directly request it, nStudy will first engage her in an "interactive" analytic that helps Mia determine how well she has learned about the topic writ large.

nStudy shows Mia a termnet developed from sources cited in her essay. Nodes are labeled by terms they represent. Terms with high centrality and betweenness centrality are highlighted. Alongside this display is this instruction: "To test your understanding, please click a highlighted term in the termnet." When Mia does so, nStudy computes whether the term's betweenness centrality exceeds a threshold.

If the betweenness centrality threshold is not exceeded, a window pops up with a four-part web form (an nStudy note template). In the first part, labeled "Term," nStudy displays the term. Part two is a slider labeled "How well I know this" with a scale labeled from 0 to 100. Part three is a text box labeled "Description." Inside the text box is gray replacement text that will disappear as Mia begins typing her reply; the replacement text is "Enter what you remember about this term." Last, a label "Click if you need to review" is printed next to a button "Review." Clicking that button opens the source in which Mia first defined the term, scrolls to the text she selected when the term was created, and opens the term's note to display what Mia entered at the earlier time when she was studying that source.

If the betweenness centrality threshold is exceeded, a window pops up with the same four parts as just described plus an additional text box located between the Description and Review parts of the preceding template. The new text box is labeled "Explain." Replacement text in the box is: "What important role does this term play in the termnet?"

This interactive analytic has two purposes. First, by attempting to recall the descriptions/definitions of concepts/terms she used in her essay, Mia is engaging in retrieval practice. This form of cognition promotes learning by engaging the testing effect (Roediger & Butler, 2011). Second, because Mia estimates her knowledge by positioning the slider, she engages in a judgment of learning. As Mia engages with each key term in the termnet, nStudy computes probabilities she reviews a term's source material as a function of her judgment of learning. After Mia has completed several of these exercises for important terms, nStudy provides a further analytic describing, as would be predicted based on the research literature, her overconfidence. More accurate judgments of learning are associated with better choices about how to study in the first place and what needs to be restudied after first engagement (Soderstrom, Yue, & Bjork, 2016). In addition, explaining the conceptual role of these terms with high betweenness centrality engages the self-explanation effect (Bisra, Liu, Salimi, Nesbit, & Winne, 2017).

The display of the termnet also includes a button: "Compare to sources." When Mia clicks it, another termnet is displayed alongside the termnet reflecting terms in her essay. This new termnet reflects the conceptual structure of terms in sources Mia cited. Terms appearing in sources that are not used in Mia's essay are color coded (e.g., blue). Terms appearing in both the termnet developed from Mia's essay and the termnet of sources she cited are heat-map shaded in proportion to the centrality index of terms in cited sources. Mia can visually compare these two termnets to examine not only the extent to which her termnet "covers" information in sources, but also how emphasis in her essay compares to that in the corpus of sources she cited in her essay. We hypothesize these comparisons can serve Mia well in considering whether and how to revise her essay. Also, various indexes can be computed to quantify how well Mia's essay structurally "correlates" with sources.

Challenge 3: Benefiting from and Contributing to Peer Review

Beyond researching her topic, Mia has two roles in relation to her learning project—she is an author and, because her instructor structured the course to involve students as peer reviewers, Mia and her peers are reviewers of several students' essays. Engaging students as peer reviewers implements findings that an author's writing can benefit from peers' feedback (Crossman & Kite, 2012; see also Cho & Cho, 2011). However, coordinating, organizing, and extracting meaningful information from multiple peers' reviews can be difficult (Cerdán & Vidal-Abarca, 2008).

nStudy Data

Drawing on source materials uploaded by her instructor, Mia will research and then draft her essay using nStudy's essay tool. In nStudy, each essay is formatted in html and assigned a unique URL; i.e., it is a web page. To participate in the peer review process, authors distribute that URL to reviewers through nStudy's hub. In Mia's course, her instructor defined teams, so Mia's essay is automatically distributed to specific reviewers her instructor selected for her team.

In nStudy's hub, reviewers access their peers' essays by clicking the URLs distributed to them. Once an essay's web page has been opened, they make comments by selecting text to create a quote. Then, they use nStudy tags the instructor provided to operationalize features of a rubric that are relevant to the general theme of the course plus other tags empirical research in writing has demonstrated commonly need attention (Ferris, 2002, p. 70–71; Hacker, Keener, & Kircher, 2009; Haswell, 2000). Examples of the latter set of tags might include: vague expression, poor word choice, and needs transition.

Reviewers can associate a tag and a note to selections of text they quote in Mia's draft. To complete a note, the reviewer selects among note templates the instructor developed. Each note template available to reviewers was designed by the instructor to provide labeled text fields, checkboxes items, and radio buttons items that relate to schemas relevant to the course topic or writing skills (e.g., genre features). Specific feedback like this has proven beneficial to improve writing (e.g., Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Hattie & Timperley, 2007; Kluger & DiNisi, 1996). In addition, students perceive specific feedback is more helpful and they are more responsive to specific feedback than to peers' global comments like "good" or "needs work" (Lipnevich & Smith, 2009; Ng, Tay, & Cho, 2015).

While peers are reviewing Mia's essay and before Mia begins her reviews of essays drafted by her team members, she is re-examining her own draft that she distributed to them. In this task, Mia has access to the same nStudy tags and note templates she will use as a peer reviewer. Her task is to tag and annotate her draft to generate predictions about what her peers will identify as "needing work."

When a peer has completed a review, Mia is notified via nStudy's hub. Each reviewer's annotations of Mia's draft is accessible as a new web page with a new URL. The sidebar for a page contains all the quotes and notes reviewers created to guide Mia's attention when she revises her draft.

As Mia edits her draft, nStudy records changes she makes. Because nStudy logged text the reviewers selected to quote spots they recommended Mia attend to, the software can identify which parts in the draft Mia has and has not revised that a peer marked.

Conceptual Description of Learning Analytics

In a sidebar on the left of Mia's reviewed essay are headers labeled: tags, quotes, notes. "Inside" each header are the artifacts peer reviewers created when they reviewed Mia's draft. At the right edge of each header is a small box reporting the number of artifacts within the header. Nubs in the gutter area adjacent to the scrollbar on the right show the relative positions in the essay of all the reviewers' artifacts.

Mia can filter items in the sidebar by entering text in the search box at the top of the sidebar. Numbers of artifacts in each header change corresponding to the focus of a filter. Mia can filter artifacts to focus on those created by a specific reviewer, with a particular tag (e.g., add transition), created using a particular note template (e.g., debate issue), or with specific text a reviewer quoted in Mia's draft or entered in a text box in a note template. nStudy hides the other artifacts not within Mia's focus.

To view the text a peer reviewer quoted in context, Mia clicks an artifact. nStudy scrolls to the quote associated with that artifact and opens a popup window to display the tag or note a reviewer used to comment about that selected text.

Learning Analytics Reports and Learning Science Principles

Analytic Report 1: Sharpening Metacognitive Monitoring. After Mia opens the web page containing her peers' reviews and a second window that displays her predictions about what might need revising, Mia can assess the accuracy of her predictions. Comparing her tags and notes predicting what reviewers would identify to reviewers' actual annotations generates feedback about standards for metacognitively monitoring one's writing (25 learning principles, 2015; Butler & Winne, 1995). Mia uses a note template to explain to herself why: (a) she picked targets her reviewers did not, (b) she missed issues her reviewers identified, and (c) her predicted targets for revision match reviewers'. Self-explaining by posing and answering such "Why?" questions has demonstrated positive benefits for learning in other areas (Bisra et al., 2017; Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013) even when feedback is not given for self-explanations (Schworm & Renkl, 2006). With repetition focusing on her drafts as well as reviewing peers' drafts (Cho & Cho, 2011), we predict this process of self-explaining reviewers' annotations relative to her forecasts will help Mia sharpen standards for accurately metacognitively monitoring her drafts and when serving as a reviewer of peers' draft essays. As well, this activity may counter overconfidence about qualities of one's own writing (Goldfinch & Hughes, 2007).

Analytic Report 2: Managing Cognitive Load and Focusing Attention. A problem novice writers face when revising an essay is high cognitive load (Kozma, 1991; Piolat, Roussey, Olive, & Amada, 2004). This challenge is exacerbated when multiple peer reviewers provide diverse recommendations for revisions dispersed throughout a draft. Several features can help Mia focus on particular issues and manage cognitive load.

First, Mia can filter reviewers' artifacts that mark particular issues to consider for revision, say, filtering tags to focus only on parts of her draft her reviewers tagged as "vague." Removing whatever other issues were identified and recommendations provided by reviewers reduces cognitive load. Also, rather than having to switch among a variety of issues that might unsystematically confront Mia if she worked through the draft from word one to the end, she can focus on one issue at a time. By refreshing the count indicator for quotes within the Tags header to report only quotes the reviewers tagged as "vague," Mia can survey which issues are most pressing (show the greatest count), and focus on those that most need attention.

Second, when Mia filters the sidebar, nStudy also filters nubs in the gutter area. Only nubs marking quotes that match the filter Mia used—in this instance, the tag “vague”—are shown. The spacing of nubs reveals whether issues are dispersed across the essay or clustered in a particular area. Again, Mia can choose an approach to revision that manages cognitive load.

Third, clicking through each instance of a single problem—vague expression—affords opportunity to monitor whether there is a common fault underlying vague expression—e.g., writing “there are several …” versus “there are four …” for items in an enumerable list—or whether local coherence is the issue when quotes tagged “vague” are viewed in context. Multiple examples of one concept help learners identify key attributes of that concept without overly taxing cognition (see Lee & Anderson, 2013).

Analytic Report 3: Re-evaluating Sentence Phrasing and Paragraph Construction. As described previously, nStudy provides means for learners to create terms, artifacts that describe conceptually significant concepts in the domain of a learning project. Once a term has been created, nStudy tracks its occurrences in sources learners view and review, and in drafts they produce. Because writing requires greater considerations about presentation beyond domain-specific concepts, and writing and revising elevate cognitive load. Mia likely struggles when considering how best to convey meaning.

For conceptually complex but discipline general concepts, such as “model” or “conjecture,” an analytic modeled on WordVis (http://wordvis.com/; see Fig. 11.2.) and merged with nStudy’s termnet will support Mia’s re-evaluation of meaning and phrasing beyond a domain’s technical terminology. The result will visually represent the conceptual elements of an individual sentence or a paragraph and offer comparisons—synonyms from the WordVis component and associated terms from the termnet component—to invite Mia to monitor whether her text represents her ideas as clearly and thoroughly as intended (cf. Trumpower & Sarwar, 2010).

Analytic Report 4: Judging Content Representation and Thesis Presentation. Modules on the server are available to apply natural language-processing techniques that create a synopsis of text. These routines can be applied to three individual corpora: sources Mia cited in her essay, quotes and notes Mia created while mining cited sources for information, and Mia’s draft essay. The synopsis is a set of sentences extracted from a corpus that best represents the meaning conveyed within it. The “size” of a synopsis can be manipulated by setting values of a sampling parameter expressed as an absolute threshold (number of extractions) or a relative threshold (percent of extractions relative to the number of units in the corpus). Comparing these synopses may help students working on learning projects that arise because the process of developing expertise in the learning project’s topic can overload cognitive resources, thus interfering with writing strategies (Beauvais, Olive, & Passerault, 2011).

Comparing the synopsis of sources cited in her draft to notes plus quotes invites Mia to consider how well her methods for extracting information capture the information available in sources. Mia might metacognitively consider whether her strategy for mining information is effective.

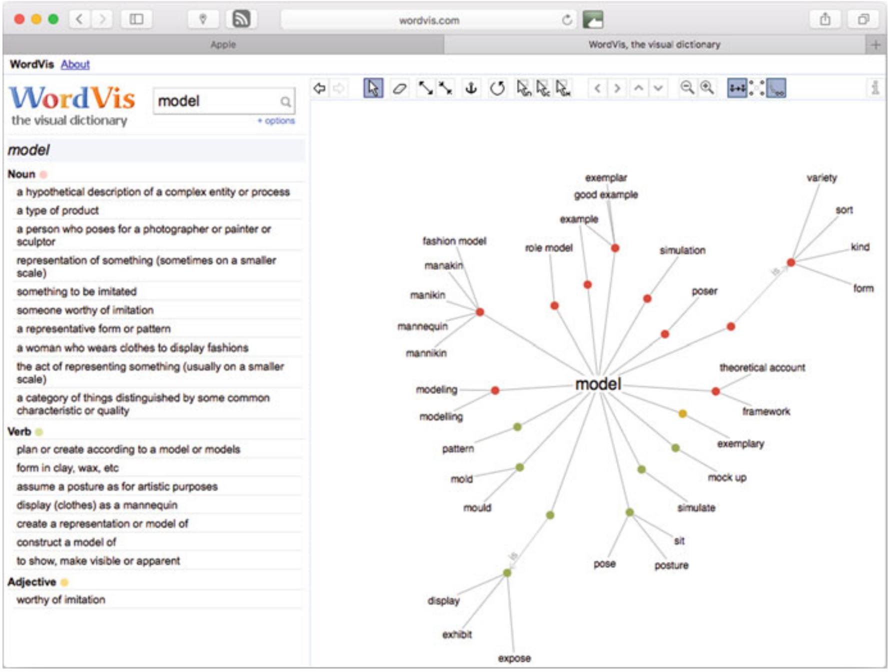

Fig. 11.2 WordVis analytic from http://wordvis.com/describing the concept "model"

Comparing the synopsis of the draft essay to the synopses of sources cited and her notes plus quotes to her draft essay affords Mia opportunity to address two metacognitive issues: First, how well does the essay reflect, or does it deliberately deviate from, information available in cited sources? In other words, is the essay biased and, if so, is the bias as intended? Second, by comparing the synopsis of her notes and quotes to that of her draft essay, Mia can metacognitively review whether criteria she uses for selecting annotations for use in her information are productive vis à vis her goals for the essay.

Conclusion

Modern computing technologies offer significant supports for learners to work on and learn from learning projects. While the Internet grants access to a hugely expanded and more diverse scope of sources for human informavores, access does not assure learners learn about the topics the research in learning projects. Nor can access to sources by itself support learners to develop knowledge and skills for solving information problems.

The information ecology we sketch includes state-of-the-art software like nStudy in a key role. nStudy mediates learners' interactions with information, recording fine-grained trace data that fuse information learners process and with an account of how information is processed and with time marks identifying when these proto-learning events occur. Trace data with this 3-tuple form supply raw material for learning analytics that can assist self-regulated learners in their quests for better marks on learning projects. Moreover, because trace data encapsulate information about skills used to address information problems, these data afford learning analytics that can be designed to advance generalizable skills for engaging metacognitively with information problems (Winne, 2013). In this newer digital learning environment, learners come to play a central role in designing effective learning environments forged as skills for solving IPS. They are supported in this task by learning analytics grounded in trace data.

We illustrated sectors of a common learning project where trace data and learning analytics developed with trace data offer strong potential to help learners cultivate both topical knowledge and skills for information problem solving. While we cited empirical support for elements within our sketches, it is important to note that direct empirical testing of our sketches is forthcoming. As the ecology of learning mediated by state-of-the-art software evolves, we predict significant benefits will arise by leveraging big data that is quite easily gathered when learners use systems like nStudy (Winne, 2017). As more and more data are amassed, and more frequent and completely authentic studies of those data are carried out, we are confident learning analytics can be designed to achieve the dual goals of boosting what learners learn in concert with advancing their skills for learning.

References

25 learning principles to guide pedagogy and the design of learning environments. (n.d.) Retrieved December 5, 2016, from https://legacy.wlu.ca/documents/60931/25-learning-principles-to-guide-pedagogy_(1).pdf

Alvermann, D. E. (1981). The compensatory effect of graphic organizers on descriptive text. *Journal of Educational Research*, *75*(1), 44–48.

Anderman, E. M., & Wolters, C. A. (2006). Goals, values, and affect: Influences on student motivation. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, *61*(2), 213–238.

Bascones, J., & Novak, J. D. (1985). Alternative instructional systems and the development of problem solving skills in physics. *European Journal of Science Education*, *7*(*3*), 253–261.

Baumann, J. F., & Bergeron, B. S. (1993). Story map instruction using children's literature: Effects on first graders' comprehension of central narrative elements. *Journal of Reading Behavior*, *25*(4), 407–437.

Beauvais, C., Olive, T., & Passerault, J.-M. (2011). Why are some texts good and others not? Relationship between text quality and management of the writing processes. *Journal of Educational Psychology*, *103*(2), 415–428.

Bisra, K., Liu, Q., Salimi, F., Nesbit, J. C., & Winne, P. H. (2017). *Inducing self-explanation: A meta-analysis.* Manuscript submitted for publication.

Boekaerts, P. P., Pintrich, P. R., & Zeidner, M. (2000). *Handbook of self-regulation*. San Diego, CA: Academic.

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education*, *53*(4), 1207–1217. https://doi.org/10.1016/j.compedu.2009.06.004

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*(1), 61–100.

Buehler, R., Griffin, D. W., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, *67*, 366–381.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281.

Cerdán, R., & Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology*, *100*(1), 209–222.

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(*1*), 73–105.

Cho, Y. H., & Cho, K. (2011). Peer reviewers learn from giving comments. *Instructional Science*, *39*, 629–643.

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *102*(3), 687–700.

Cronin, H., Sinatra, R. C., & Barkley, W. (1992). Combining writing with text organization in content instruction. *National Association of Secondary School Principals (NASSP) Bulletin*, *76*, 34–45.

Crossman, J. M., & Kite, S. L. (2012). Facilitating improved writing among students through directed peer review. *Active Learning in Higher Education*, *13*(3), 219–229.

Dehmer, M., & Emmert-Streib, F. (Eds.). (2009). *Analysis of complex networks: From biology to linguistics*. Weinheim, Germany: Wiley-VCH.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58.

Eisenberg, M. B. (2008). Information literacy: Essential skills for the information age. *DESIDOC Journal of Library & Information Technology*, *28*(2), 39.

Ferris, D. R. (2002). *Treatment of error in second language writing classes*. Ann Arbor, MI: University of Michigan Press.

Foster, N. L., Was, C. A., Dunlosky, J., & Isaacson, R. M. (2016). Even after thirteen class exams, students are still overconfident: The role of memory for past exam performance in student predictions. *Metacognition Learning*, *12*(1), 1–19. https://doi.org/10.1007/s11409-016-9158-6

Goldfinch, J., & Hughes, M. (2007). Skills, learning styles and success of first-year undergraduates. *Active Learning in Higher Education*, *8*(3), 259–273.

Gurlitt, J., & Renkl, A. (2008). Are high-coherent concept maps better for prior knowledge activation? Differential effects of concept mapping tasks on high school vs. University students. *Journal of Computer Assisted Learning*, *24*(*5*), 407–419. https://doi.org/10.1111/j.1365-2729.2008.00277.x

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, NY: Routledge.

Haswell, R. H. (2000). Documenting improvement in college writing: A longitudinal approach. *Written Communication*, *17*, 307–352.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(*1*), 81–112.

Hopkins, D. J., & King, G. (2010). A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, *54*(*1*), 229–247.
Ives, B., & Hoy, C. (2003). Graphic organizers applied to higher-level secondary mathematics. *Learning Disabilities Research & Practice*, *18*(1), 36–51.
Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS ONE*, *9*(6), e98679. https://doi.org/10.1371/journal.pone.0098679
Johnson, D., & Steele, V. (1996). So many words, so little time: Helping college ESL learners acquire vocabulary-building strategies. *Journal of Adolescent & Adult Literacy*, *39*(5), 348–357.
Kahneman, D., & Tversky, A. (1979). Intuitive prediction: Biases and corrective procedures. *TIMS Studies in. Management Science*, *12*, 313–327.
Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284.
Kozma, R. B. (1991). Learning with media. *Review of Educational Research*, *61*(2), 179–211.
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.
Lee, H. S., & Anderson, J. R. (2013). Student learning: What has instruction got to do with it? *Annual Review of Psychology*, *64*, 445–469.
Lipnevich, A. A., & Smith, J. K. (2009). Effects of differential feedback on students' examination performance. *Journal of Experimental Psychology: Applied*, *15*(4), 319.
Lipson, M. (1995). The effect of semantic mapping instruction on prose comprehension of below-level college readers. *Literacy Research and Instruction*, *34*(4), 367–378.
Locke, E. A., & Latham, G. P. (1990). Work motivation and satisfaction: Light at the end of the tunnel. *Psychological Science*, *1*(4), 240–246.
Manning, C. D., Raghavan, P., & Schütze, H. (2008a). Scoring, term weighting and the vector space model. *Introduction to Information Retrieval*, *100*, 2–4.
Manning, C. D., Raghavan, P., & Schütze, H. (2008b). Language models for information retrieval. *Introduction to Information Retrieval*, 218–233.
Marzouk, Z., Rakovic, M., Liaqat, A., Vytasek, J., Samadi, D., Stewart-Alonso, J., ... Nesbit, J. C. (2016). What if learning analytics were based on learning science? Australasian Journal of Educational Technology, 32, 1–18.
Mayer, R. E. (2001). *Multimedia learning*. NY: Cambridge University Press.
Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, *76*(3), 413–448. https://doi.org/10.3102/00346543076003413
Ng, Z. Y., Tay, W. Y., & Cho, Y. H. (2015). Usefulness of peer comments for English language writing through web-based peer assessment. In G. Chen, V. Kumar, H. R. Kinshuk, & S. Kong (Eds.), *Emerging issues in smart learning. Lecture notes in educational technology*. Berlin, Germany: Springer.
Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching*, *27*(10), 937–949.
Novak, J. D. (1991). Clarify with concept maps: A tool for students and teachers alike. *The Science Teacher*, *58*, 45–49.
Novak, J. D. (1998). *Learning, creating, and using knowledge*. Mahwah, NJ: Erlbaum.
Novak, J. D., & Wandersee, J. (1991) (Eds.) Special issue on concept mapping. *Journal of Research in Science Teaching, 28,* 10.
Ojima, M. (2006). Concept mapping as pre-task planning: A case study of three Japanese ESL writers. *System*, *34*(4), 566–585.
Patterson, E. W. (2001). Structuring the composition process in scientific writing. *International Journal of Scientific Education*, *23*, 1–16.

Patzak, A., & Winne, P. H. (2016). *Using research on decision making to account for why and how students self-handicap*. Unpublished manuscript.

Pieronek, F. (1994). Using maps to teach note taking and outlining for report writing. *The Social Studies*, *85*, 165–169.

Piolat, A., Roussey, J.-Y., Olive, T., & Amada, M. (2004). Processing time and cognitive effort in revision: Effects of error type and of working memory capacity. In L. Allai, L. Chanquoy, & P. Largy (Eds.), *Revision cognitive and instructional processes* (pp. 21–38). Boston, MA: Kluwer Academic Publishers.

Reynolds, S. B., & Hart, J. (1990). Cognitive mapping and word processing: Aids to story revision. *Journal of Experimental Education*, *58*, 273–279.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*, 20–27.

Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, *2*(1), 7–12.

Roth, W. M. (1994). Student views of collaborative concept mapping: An emancipatory research project. *Science Education*, *78*(1), 1–34.

Saddler, B., Moran, S., Graham, S., & Harris, K. R. (2004). Preventing writing difficulties: The effects of planning strategy instruction on the writing performance of struggling writers. *Exceptionality*, *12*(1), 3–17.

Schultz, M. (1991). Mapping and cognitive development in the teaching of foreign language writing. *The French Review*, *64*, 978–988.

Schunk, D. H. (1995). Self-efficacy and education and instruction. In J. E. Maddux (Ed.), *Self-efficacy, adaptation, and adjustment: Theory, research, and application*. New York, NY: Plenum Press.

Schunk, D. H. (2001). Self-regulation through goal setting. In *ERIC/CASS Digest.* Washington, DC. Retrieved from www.eric.ed.gov. ED462671.

Schwarz, N. (2002). Emotion, cognition, and decision making. *Cognition and Emotion*, *4*, 433–440.

Schworm, S., & Renkl, A. (2006). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers and Education*, *46*(4), 426–445.

Trumpower, D. L., & Sarwar, G. S. (2010). Effectiveness of structural feedback provided by Pathfinder networks. *Journal of Educational Computing Research*, *43*(1), 7–24.

Washington, V. M. (1988). Report writing: A practical application of semantic mapping. *Teacher Educator*, *24*, 24–30.

Whittaker, R., Llinares, A., & McCabe, A. (2011). Written discourse development in CLIL at secondary school. *Language Teaching Research*, *15*(3), 343–362.

Winn, W. (1991). Learning from maps and diagrams. *Educational Psychology Review*, *3*(3), 211–247.

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 153–189). Mahwah, NJ: Erlbaum.

Winne, P. H. (2013). Teaching and researching open-minded inquiry in the 21st century. In J. View, D. Laitsch, & P. Earley (Eds.), *Why public schools? Voices from the United States and Canada* (pp. 166–170). Charlotte, NC: Information Age Publishing.

Winne, P. H. (2017a). Cognition and metacognition in self-regulated learning. In D. Schunk & J. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed.). New York, NY: Routledge.

Winne, P. H. (2017b). Learning analytics for self-regulated learning. In G. Siemens & C. Lang (Eds.), *Handbook of learning analytics*. Beaumont, AB: Society for Learning Analytics Research.

Winne, P. H. (2017c). Leveraging big data to help each learner upgrade learning and accelerate learning science. *Teachers College Record*, *119*(3), 1–24.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Lawrence Erlbaum Associates.

Wyatt, D., Pressley, M., El-Dinary, P. B., Stein, S., Evans, P., & Brown, R. (1993). Comprehension strategies, worth and credibility monitoring, and evaluations: Cold and hot cognition when experts read professional articles that are important to them. *Learning and Individual Differences*, *5*(1), 49–72.

Zimmerman, B. J., Bonner, S., & Kovach, R. (1996). *Developing self-regulated learners: Beyond achievement to self-efficacy*. Washington, DC: American Psychological Association.

Chapter 12
Successes and Failures in Building Learning Environments to Promote Deep Learning: The Value of Conversational Agents

Arthur C. Graesser, Anne M. Lippert, and Andrew J. Hampton

A contrast is often made between shallow and deep knowledge, although the boundaries between the two are fuzzy rather than crisp (Chi, 2009; Graesser, 2015; Hattie & Donoghue, 2016). Shallow knowledge of a subject matter includes identification of key terms, their features, and simple definitions. However, deep knowledge is needed for causal reasoning, solving difficult problems, integrating components in complex systems, justifying claims with logical arguments, identifying inaccurate information, resolving contradictions, designing useful new artefacts, and precisely quantifying ideas. It is much more difficult to acquire deep knowledge, so people tend to settle for shallow knowledge on most topics unless they are forced to dive into deeper waters by virtue of schooling, employment, obstacles to important goals, and other external pressures. However, deep knowledge and reasoning skills will fortify citizens who desire higher paying jobs in the twenty-first century (Autor, Levy, & Murnane, 2003; Carnevale & Smith, 2013; Levy & Murnane, 2006).

This chapter describes some attempts to promote deep learning through conversational pedagogical agents ("agents" for short). Learning environments with agents have been developed to serve as substitutes for humans who range in expertise from novices to experts. Agents can guide the interaction with the learner, instruct the learner what to do, and interact with other agents to model ideal behavior, strategies, reflections, and social interactions (Graesser, Li, & Forsyth, 2014; Johnson & Lester, 2016). Some agents generate speech, gestures, body movements, and facial expressions in ways similar to people. These agent-based systems are afforded with characteristics that serve both motivational and social-interactive functions in addition to the cognitive functions that are so important to learning in informational

A.C. Graesser (✉) • A.M. Lippert • A.J. Hampton
Department of Psychology and Institute for Intelligent Systems, University of Memphis, 202 Psychology Building, Memphis, TN, 38152-3230, USA
e-mail: graesser@memphis.edu

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1_12

environments. These systems have sometimes facilitated deep learning more than conventional learning environments, as we will discuss in this chapter.

Agent-based learning environments have proliferated over the last decade. *AutoTutor* and its descendants (Graesser, 2016; Nye, Graesser, & Hu, 2014) have successfully helped college students learn various skills and subject matter by holding a conversation in natural language. These conversation-based systems have covered topics such as computer literacy (Graesser et al., 2004), physics (*DeepTutor*, Rus, D'Mello, Hu, & Graesser, 2013; *AutoTutor*, VanLehn et al., 2007), biology (*GuruTutor*, Olney et al., 2012), and scientific reasoning (*Operation ARIES/ARA*, Halpern et al., 2012; Kopp, Britt, Millis, & Graesser, 2012; Millis, Forsyth, Wallace, Graesser, & Timmins, 2017). Other examples of systems with agents that have successfully improved student learning are *MetaTutor* (Azevedo, Moos, Johnson, & Chauncey, 2010); *Betty's Brain* (Biswas, Jeong, Kinnebrew, Sulcer, & Roscoe, 2010), *iDRIVE* (Craig, Gholson, Brittingham, Williams, & Shubeck, 2012; Gholson et al., 2009), *iSTART* (Jackson & McNamara, 2013; McNamara, O'Reilly, Best, & Ozuru, 2006), *Crystal Island* (Rowe, Shores, Mott, & Lester, 2011), *My Science Tutor* (Ward et al., 2013), and *Tactical Language and Culture System* (Johnson & Valente, 2009).

This chapter begins by presenting evidence that deep knowledge is not routinely acquired by conventional learning environments, such as reading from text, listening to lectures, and exploring hypermedia environments. These traditional learning environments are excellent ways of exposing learners to information and helping them acquire a broad spectrum of shallow knowledge. However, they do not have the affordances for promoting the acquisition of deep knowledge, such as active generation of information, feedback on performance or creations, and communication with others. One way to embody these affordances is through agents—the approach we cover in this chapter. Other methods may promote deep knowledge acquisition, such as interactive simulation environments, design of artefacts, and dynamic, adaptive testing during learning. However, it is the agent approach that we have pursued for the last 20 years in the interdisciplinary Institute for Intelligent Systems at the University of Memphis. This chapter identifies some of our successes and failures in designing agent-based environments to promote deep learning.

Some Notable Failures in Promoting Acquisition of Deep Knowledge

Much can be learned from failure. This section identifies a number of approaches to acquiring deep knowledge that are not very effective. We offer these with an appropriate modicum of caution because some people have excellent strategies of self-regulated learning that fortify them for deep learning from a broad spectrum of learning environments. However, the vast majority of children and adults do not have well developed self-regulated learning strategies (Azevedo et al., 2010;

Graesser, McNamara, & VanLehn, 2005). Consequently, the development of learning environments to support self-regulated learning is a critical area of research (see Chap. 9 by Scheiter et al., and Chap. 10 by Azevedo et al., of this volume).

Reading Text

Instructors typically assign texts for students to read in courses with the hope that the students will vigorously study the texts at a deep level, engage in reasoning, and generate important inferences. Inference generation is quite respectable for stories and other forms of narrative texts (Graesser, Singer, & Trabasso, 1994), but is extremely limited for technical texts on unfamiliar topics (Lorch, 2015). Students typically cannot accurately calibrate how well they have comprehended technical text. Comprehension calibration research has revealed that even college students' judgments on how well they understand expository text has only a 0.2–0.3 correlation with objective comprehension scores on the text (Dunlosky & Lipko, 2007; Maki, 1998). Extremely shallow readers view a text as comprehended if they can recognize the words in the text—a metacognitive standard confined to the vocabulary level. Readers with a sentence-level standard of comprehension are satisfied that they understand a text if they can interpret the meaning of individual sentences, without consideration of how the sentences are related at the discourse level. Deeper readers adopt a coherence-based standard of comprehension so they notice when there are coherence gaps and inconsistences between ideas (Baker, 1985; Van den Broek, Bohn-Gettler, Kendeou, Carlson, & White, 2011). These students generate inferences that attempt to fill in coherence gaps and resolve inconsistencies (Zwaan & Radvansky, 1998). The deepest readers adopt a critical stance that is sensitive to the text genre (e.g., fiction, persuasion, informational), the expertise of the author, inaccurate information, and clashes in claims between multiple texts (Bräten, Britt, Strømsø, & Rouet, 2011; Kimmerle et al., Chap. 4, this volume; Rapp & Braasch, 2014; Rouet, 2006). This level of depth is extremely important in the twenty-first century because progressively fewer documents are edited for accuracy of facts and claims (Rapp & Braasch, 2014).

We have conducted studies on the subject matters of computer literacy and physics that assess how well college students perform on tests that require reasoning after they read texts versus interact with conversational agents (Graesser et al., 2004; VanLehn et al., 2007). For example, in Newtonian physics, a widely-used test is the *Force Concept Inventory* (Hestenes, Wells, & Swackhamer, 1992). This test assesses how well a student can track the causes and consequences of an event in a physical system. Consider a problem in which a packet is dropped from an airplane directly over a target at a particular point in time; the student is asked whether the packet will land on the target, in front of the target, behind the target, or it is impossible to know. These questions require reasoning that integrates Newtonian laws of physics rather than simply retrieving factual information. VanLehn et al. (2007) compared performance on tests like the Force Concept Inventory on college students randomly assigned to one of four conditions: Intelligent tutoring systems with conversational agents (Agents), human tutors in computer mediated conversions (Tutors), reading texts for a time comparable to the first two conditions (Reading), and doing nothing (Nothing). The results showed the following pattern of performance on the Force Concept Inventory and other measures of deep learning: Agents = Tutors > Reading = Nothing. This pattern also occurred in a three-condition study without human tutors on the topic of computer literacy: Agents > Reading = Nothing (Graesser et al., 2004).

The above studies have two important implications for the acquisition of deep knowledge, as reflected in a discriminating test that requires reasoning and applications of knowledge to problems. First, reading text is inadequate, akin to doing nothing. Second, learning environments must promote active learning, feedback, communication, and other features that are absent in static materials like text.

Listening to Lectures

Lectures are another frequent learning environment that evolved to disseminate information to a large number of students efficiently. This technique became popular during the industrial revolution when shallow learning was adequate for the working class to run the factories by performing well-defined, routine, manual and cognitive tasks. Lectures remain popular in most school systems and universities today, despite greater demand for deep learning, problem solving, and reasoning. Lecturing is not suited to deep learning because students have no control over the information delivery (Chi, 2009) and run the risk of mind wandering, particularly when the material is difficult (Medimorecc, Pavlik, Olney, Graesser, & Risko, 2015). Reading allows more control over the information delivery than lectures because the student can voluntarily review excerpts. Lectures are appropriate for exposure to shallow knowledge, refresher training, and subject matters that the listener knows about, but not for deep learning of unfamiliar material.

There is an interesting thought experiment that addresses attention to and memory for lectures. Students attend to lecture content only a portion of the time (let us assume 50% for example) and the forgetting functions for verbal material indicate that only 10% of the studied information is recalled 1 week later, according to the classical Ebbinghaus forgetting functions. It follows that only about 5% of the lecture material is readily retrievable 1 week after the lecture. Teachers, experts themselves, frequently assume that 100% of the material is mastered and retrievable so they build on that in their next lecture. In truth, only a small fraction of the material is readily retrieved and even less of it can be transferred in applications of knowledge to real-world problems. In conclusion, lectures are not well suited to deep learning.

Exploring Hypertext and Hypermedia

It is conceivable that learners could acquire deep knowledge by having the freedom to explore diverse knowledge sources on their own rather than being sequestered by the tight constraints of lectures and reading of texts in a linear fashion. However, this ideal vision is rarely experienced by most learners because of their limited self-regulated learning strategies, which would benefit from explicit training (Azevedo et al., 2010).

An ideal self-regulated learner who desires deep knowledge acquisition would ask relevant deep questions that require reasoning, such as why, how, so what, what if, what if not, etc., but not the shallow factual questions such as who, what, when, and where (Graesser, Ozuru, & Sullins, 2009). Unfortunately, students, classroom teachers, and tutors tend to ask primarily shallow questions rather than deep questions (Graesser & Person, 1994). Classroom teachers and most tutors do not frequently model the asking of deep questions, so there is not adequate exposure of students to these inquiry strategies. Training students to ask good questions does improve comprehension (Rosenshine, Meister, & Chapman, 1996), but this training is infrequent in most classrooms and tutoring practices.

Graesser, Langston, and Baggett (1993) investigated whether students would ask more deep questions if simply exposed to the deep questions in a hypermedia environment. The researchers developed a hypermedia environment that covered both shallow and deep knowledge of woodwind instruments. Shallow knowledge included pictures of the instruments, what they sounded like, and historical facts. Deep knowledge consisted of the physics of sound and how pitch, timbre, and other acoustic parameters are influenced by the size of the instruments, the mouthpiece (e.g., double reed, single reed, vs. no reed), and other properties of the woodwind instruments. They had hoped that students would ask more deep questions if they were exposed to deep questions through a *Point & Query (P&Q)* interface. In this interface, students encounter a variety of questions as they explore hypermedia pages. Whenever they click on a hot spot on a screen, a list of questions appears; students can then click on any question they wish to ask and an answer appears.

A series of experiments tested whether the students would ask more deep questions when they were exposed to a mix of shallow and deep questions. College students were randomly assigned to one of three conditions that manipulated their goals before they started exploring the woodwind instruments hypermedia environment: (a) A deep goal, in which students were instructed to write an essay on designing an instrument with a deep pure tone, (b) a shallow goal in which students were instructed to organize a combo for a New Year's Eve party, and (c) a vague goal in which students were instructed to explore the hypermedia because they would be tested later. Students could only learn about woodwind instruments by exploring the hypermedia environment with the P&Q facility.

Graesser et al. (1993) analyzed the frequency and percentage of questions that were shallow versus deep. A number of illuminating conclusions were supported by the data. First, the frequency of questions per hour was quite high in the P&Q environment, where the only way students could learn was by pointing to hot spots, selecting questions, and reading answers to questions. The rate of student questions was 700 times the rate of student questions in classrooms and 5–10 times the rate of student questions in human tutoring (Graesser & Person, 1994). Second, the vast majority of questions were shallow questions (and approximately the same) in the shallow goal and vague goal conditions. This is consistent with the hypothesis that students typically swim in shallow waters when they learn about topics in these hypermedia environments. Third, the percentage of deep questions was substantially higher in the deep goal condition. Consequently, it takes external events to push the students into deeper waters before students ask deep questions and comprehend the answers.

Using Agents to Enhance Deep Learning

During the last 20 years, the Institute for Intelligent Systems (IIS) at the University of Memphis has developed a number of learning environments with conversational agents that attempt to facilitate deep learning. It started with *AutoTutor* (Graesser, 2016; Nye et al., 2014), a system that helps college students learn Science, Technology, Engineering, and Mathematics (STEM) subject matters by holding a conversation in natural language. A number of systems with conversational agents evolved over the next two decades in the IIS on a broad set of topics: computer literacy, physics, biology, research ethics, scientific reasoning, comprehension strategies, electronic circuits, medical reasoning, and group problem solving. This section summarizes our efforts to improve deep learning through the conversational agents developed in the IIS.

The number of agents interacting with the learner is an important consideration in the design of agent-based systems. In a *dialogue*, a single agent interacts with a single human, as in the case of the original AutoTutor. In a *trialogue*, there is a three-party conversation among two computer agents and a human student. The two agents take on different roles, but often serve as tutors and peers of the student. There are different trialogue designs that address different pedagogical goals for different classes of students (Graesser, Forsyth & Lehman, 2017; Graesser et al., 2014). For example, students can observe two agents interacting so that the student can model the actions and thoughts of an ideal agent. Agents can argue with each other over issues and ask what the human student thinks about the argument. A tutor agent can pit a human student against a peer agent in a game scenario in order to enhance motivation. It is also possible for a single agent to communicate with a small group of students, as in the case of AutoMentor, which was designed to facilitate group learning and problem solving (Morgan, Keshtkar, Graesser, & Shaffer, 2013).

Modeling Deep Question Asking Through iDRIVE

iDRIVE (Instruction with Deep-level Reasoning questions In Vicarious Environments) has two agents that train students to learn science content by modeling deep reasoning questions in question-answer dialogues (Craig et al., 2012; Gholson et al., 2009). A student agent asks a series of deep questions about the STEM content (such as physics, biology, or computer literacy) and the teacher agent immediately answers each question. Approximately 30 deep questions per hour are asked (by the student agent) and answered (by the tutor agent) in this learning environment. Therefore, a curious student agent models high quality inquiry with good questions and the tutor agent provides good explanation-based reasons in the answers. The large number of trials is much higher than the rate of deep questions during human tutoring (Graesser & Person, 1994). Thus, iDRIVE delivers many trials of good question asking and explanation-oriented answers which are known to improve deep learning (Chi, 2009; Pashler et al., 2007; Rosenshine et al., 1996). Training students with iDRIVE significantly increased the frequency of good questions and memory performance on transfer subject matters (Craig et al., 2012; Gholson et al., 2009).

One limitation of iDRIVE is that students vicariously observe the two agents interacting so they are not actively constructing information. Active constructive of information is an important component of deep learning according to Chi's (2009) *ICAP* model (Interactive, Constructive, Active, Passive). iDRIVE has entirely passive affordances for learning, which is suitable for students at an introductory stage of a subject matter when they cannot actively construct much information. It is less effective for students at the more intermediate and advanced stages.

Intelligent Dialogues and Trialogues with AutoTutor

AutoTutor (Graesser, 2016; Nye et al., 2014) adds on the Interactive, Constructive, and Active learning components of Chi's ICAP model (Chi, 2009), which is more suited to the intermediate and advanced learners. AutoTutor presents problems to solve and difficult questions to answer that require reasoning and that cover up to seven sentence-like conceptual expressions (e.g., semantic propositions, claims, main clauses) in an ideal response. The humans and agents co-construct a solution by multiple conversational turns. It may take a dozen to 100 conversational turns back and forth to solve a problem or answer a difficult question. AutoTutor may involve either a dialogue between the human and a tutor agent, or alternatively a trialogue between the human and two agents.

It is beyond the scope of this article to describe the mechanisms of AutoTutor in detail (see Graesser, 2016; Nye et al., 2014). However, a systematic conversational mechanism called *expectation and misconception-tailored (EMT) dialogue* (or trialogue) requires explanation. A list of *expectations* (anticipated good answers, steps in a procedure) and a list of anticipated *misconceptions* (bad answers, incorrect beliefs, errors, bugs) is associated with each task. As the students articulate their answers over multiple conversational turns, the contents of their contributions are compared with the expectations and misconceptions. Students rarely articulate a complete answer in the first conversational turn but rather their answers are spread out over many turns as the tutor generates hints and other conversational moves to enable the students to express their knowledge. The students' answers within each turn are typically short (one to two speech acts on average), vague, ungrammatical, and not semantically well formed. AutoTutor compares the students' content to expectations and misconceptions through pattern matching processes with semantic evaluation mechanisms motivated by research in computational linguistics (Rus, Lintean, Graesser, & McNamara, 2012).

Listed below are tutor dialogue moves that frequently occur in the EMT dialogues in AutoTutor and in most human tutoring sessions.

Main question or problem. This is a challenging question or problem that the tutor is trying to help the student answer.

Short feedback. The feedback to a student's previous contribution is either *positive* ("yes," "correct," head nod), *negative* ("no," "almost," head shake, long pause, frown), or *neutral* ("uh huh," "okay").

Pumps. The tutor gives nondirective pumps ("Anything else?" "Tell me more.") to say more or to take some action.

Hints. The tutor gives hints to get the students to do the talking or take action, but directs the students along some conceptual path. The hints vary from generic statements or questions ("What about X?" "Why?") to speech acts that nudge the student toward a particular answer.

Prompts. The tutor asks a leading question to get the student to articulate a particular word or phrase. Sometimes students say very little so these prompts are needed to get the student to say something specific.

Prompt completions. The tutor expresses the correct completion of a prompt.

Assertions. The tutor expresses a fact or state of affairs.

Summaries. The tutor gives a recap of the answer to the main question or solution to the problem.

Mini-lectures. The tutor expresses didactic content on a particular topic.

Corrections. The tutor corrects an error or misconception of the student.

Answers. The tutor answers a question asked by the student.

Rather than simply lecturing to the student, the tutor provides scaffolding for the student to articulate the expectations through a number of dialogue moves. AutoTutor provides a cycle of *pump* → *hint* → *prompt* → *assertion* for each expectation until the expectation is covered. As the student and tutor express information over many turns, the list of expectations is eventually covered and the main task is completed. The student may articulate misconceptions during the multi-turn tutorial dialogue. When the student content has a high match to a misconception, AutoTutor acknowledges the error and provides correct information.

These dialogue moves differ in the extent to which the student versus the tutor supplies the expectation content. For example, the tutor supplies progressively more of the expectation information from pump to hint to prompt to assertion to summaries. According to the principle of active student learning, there should be a greater onus on the student, rather than the tutor, in supplying the expectation information. Indeed, tutorial dialogues with more knowledgeable students have a higher proportion of tutor pumps and hints, requiring greater student input, rather than prompts and assertions where the tutor agent provides most of the information (Jackson & Graesser, 2006).

AutoTutor has shown learning gains of approximately 0.80 standard deviations compared with reading a textbook for an equivalent amount of time (AutoTutor, 2016). Approximately a dozen measures of learning have been collected in these assessments, including multiple-choice questions, essay quality when students attempt to answer challenging questions, a cloze task that has students fill in missing words of texts that articulate explanatory reasoning on the subject matter, and performance on tasks that require problem solving. AutoTutor is most impressive for increasing learning gains on measures of deep rather than shallow knowledge. Moreover, AutoTutor is most suited for students at intermediate and advanced knowledge on a subject matter because it requires them to actively construct information during the interaction.

AutoTutor with Interactive Simulation

AutoTutor can scaffold student action in addition to articulating explanations in natural language. This is exemplified by AutoTutor 3-D, a version of AutoTutor with interactive simulation on Newtonian physics problems (Graesser, Chipman, Haynes, & Olney, 2005). Figure 12.1 shows the interface on one of the problems: "When a car without headrests on the seats is struck from behind, the passengers often suffer neck injuries. Why do passengers suffer neck injuries?" The main question is presented at the top of the screen. Below the question are two windows that show the car and truck (middle window) and the driver in the car (right window). These components move whenever a simulation runs.

Beneath the question on the left is the animated agent that guides the interaction with hints, suggestions, assertions, and other dialogue moves. These suggestions include having the student manipulate parameters, such as truck speed, mass of the car, and mass of the truck. The students also have several binary options: having the head rests in the car on, showing the skin on the driver, slowing down the simulation, and displaying vector arrows that depict forces. The vector arrows can sometimes be informative in the explanations. For example, some students have a billiard ball mental model that leads them to assume that the vehicle in the back pushes all parts of the vehicle in the front to move forward (including the driver's head) and that pushes the head into the windshield. However, they miss the importance of the force vectors that explain that the driver initially gets a whiplash from impact forces and then on the recoil goes through the windshield.

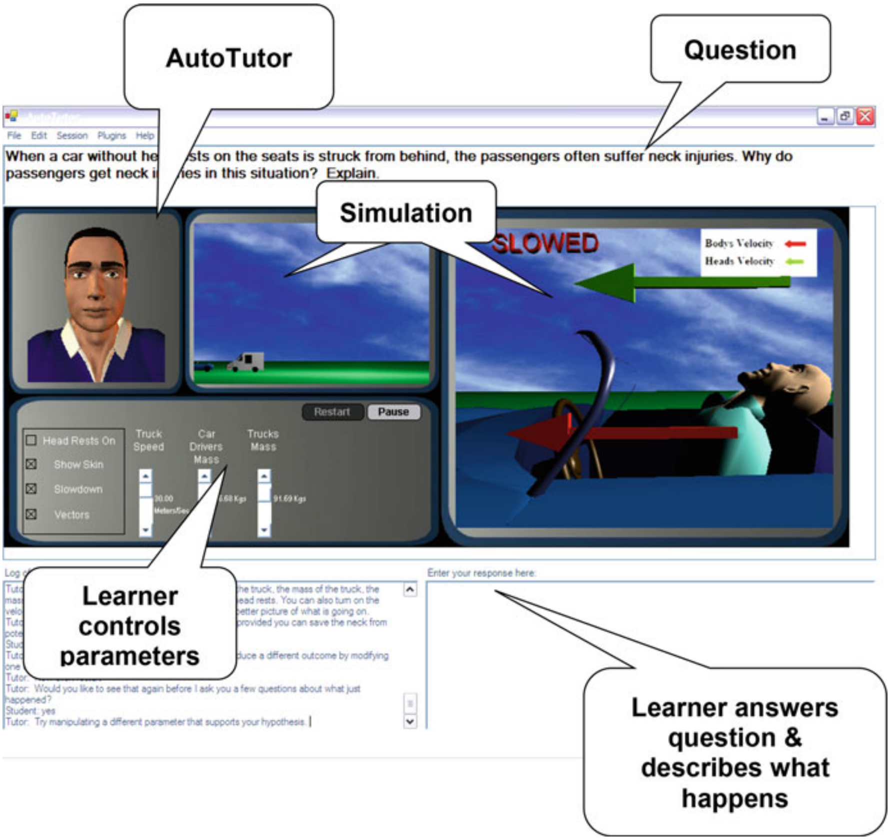

Fig. 12.1 A screenshot of AutoTutor on conceptual physics with 3D simulation (adapted from Graesser, Chipman, et al., 2005)

The student has the freedom to manipulate these parameters and options, as shown in the bottom left, before a simulation runs. The activity of manipulating these inputs and viewing the simulation can provide a referentially grounded and embodied representation of the problem and a deeper understanding of physics. The students can run as many simulations as they wish until they feel they understand the relationship between parameters and outcomes of simulations. However, interacting with and viewing the simulations is not all there is. The participants are also prompted to describe what they see and answer the main question. Therefore, deep learning of physics theoretically emerges from the combination of interactivity, perceptual simulation, feedback on the simulation, and explaining what happens. All of these are important components of Chi's ICAP model to promote deep learning.

The central question is whether the augmentation of AutoTutor with interactive simulation helps deep learning. There are reasons for optimism and pessimism. AutoTutor 3-D does have an advantage over the version of AutoTutor without

3-D, but only for those learners who use the 3-D reasonably often. The bad news is that most students ran the simulation on a problem only once or twice, if at all. Disappointingly few students applied the Vary-One-Thing-at-a-Time (VOTAT) strategy in which one component is varied while all others are held constant (Greiff, Wüstenberg, Holt, Goldhammer, & Funke, 2013). None of the students assessed the impact of two variables (e.g., mass of vehicle 1 and mass of vehicle 2) on an output measure (e.g., head velocity) that requires a minimum of four simulations: high versus low mass of vehicle 1 crossed with high versus low mass of vehicle 2. Strategies of control over variables in these interactive simulations require training and do not come naturally (Lorch et al., 2010). The AutoTutor agent can play a role in providing this training. It remains to be seen whether this strategy training would have added value in AutoTutor 3-D.

SEEK (Source, Evidence, Explanation, and Knowledge) Web Tutor

Critical thinking about science requires learners to actively evaluate the truth and relevance of information, the quality of information sources, the plausibility of causal systems, and the implications of evidence (Braasch, Rouet, Vivert, & Britt, 2012; Bräten et al., 2011; Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012; Kimmerle et al., Chap. 4, this volume; Rouet, 2006; Wiley et al., 2009). A deep understanding is needed to construct causal reasoning, integration of the components in complex systems, and logical justifications of claims. This is difficult to achieve without sufficient prior domain knowledge and reasoning skills (Kendeou & Van den Broek, 2007). However, a major start is for the student to acquire a thinking strategy with a *critical stance*—essentially being a skeptic.

SEEK Web Tutor (Graesser et al., 2007) was designed to improve college students' critical stance while they search web pages on the topic of plate tectonics. Some of the web sites provided reliable information on the topic (e.g., National Aeronautics and Space Administration, the Public Broadcasting Service, and Scientific American) whereas others were erroneous accounts of earthquakes and volcanoes that appealed to the stars, the moon, and oil drilling. The students' goal in the experiments was to search the web for the purpose of writing an essay on what caused the eruption of the Mount St. Helens volcano in 1980.

SEEK Web Tutor took a direct approach to training students on critical stance by using three facilities with different instruction methods. The first was a *Hint* button on the Google search engine page that contained suggestions on how a student can strategically search and read a set of web sites. This page approximated Google, with accessible titles and URLs for reliable and unreliable web sites. When a student clicked the Hint button, spoken messages gave reminders of the goal of the task (i.e., writing an essay on the causes of the Mount St. Helens volcano eruption in the state of Washington) and suggestions on what to do next (i.e., reading web sites with reliable information). The agent in one version of the SEEK Tutor had a talking head, but the final version had "voice only" facility because there was a worry that the visual animation of the head would distract from learning web material. This facility was presented only when the student clicked the Hint button. On reflection, this was an unfortunate design decision because students rarely ask for help when they need it (Graesser, McNamara, et al., 2005).

The second computer facility was a pop-up *Rating and Justification* that asked students to evaluate the expected reliability of the information in a site. This facility appeared after the students first viewed a particular website for 20 s. The student rated the reliability of the information and typed in a verbal justification of the rating.

The third computer facility consisted of a pop-up *Journal* with five questions about the reliability of the site that the learner just visited. These questions addressed some of the core aspects of critical stance: *Who authored this site? How trustworthy is it? What explanation do they offer for the cause of volcanic eruptions? What support do they offer for this explanation? Is this information useful to you? If so, how will you use it?* Each of the questions had a Hint button that could be pressed to receive spoken hints on answering each question. The pop-up Journal launched whenever the student left one of the web sites. It forced the student to think about each of the five core aspects of critical stance. The student gave a rating for each question and typed in a verbal justification for each rating.

The three facilities were expected to improve students' critical stance by direct didactic training and application of the principles to the exemplar web sites. We conducted experiments to test the effectiveness of this approach (Graesser et al., 2007). College students explored the web sites for approximately 1 h with the goal of writing an essay on the causes of the eruption of Mount St. Helens. Participants were randomly assigned to either the SEEK Web Tutor condition or to a "navigation" condition that had no training on critical stance. Students were assessed on critical stance by over a dozen measures, including an essay on the causes of a volcano. Unfortunately, we were quite mistaken that the intervention would be effective. An hour of intense training on critical stance via SEEK Web Tutor had little impact on college students when compared to the navigation control condition. SEEK Web Tutor did not improve learners' ability to detect reliable versus unreliable information sources during training, the amount of study time they allocated to reliable versus unreliable sites, their judgments of the truth/falsity of 30 statements about plate tectonics after the training was completed, or the articulation of core ideas about plate tectonics in the essays. Only one measure showed benefit from the SEEK Web Tutor over the navigation control: Students had more expressions in the essay with language about causal explanations (such as "cause" and "explanation") compared to the navigation control.

SEEK Web Tutor essentially failed when confined to a 1-h experiment. The question remains unanswered on what would happen if there was a much larger dosage of training that is woven into practical experiences with many topics. How much training is needed before noticeable improvements appear?

Serious Games with Operation ARIES/ARA

A different approach to acquiring deep knowledge is to immerse the student in motivating learning environments, such as games. Learners would ideally be challenged and motivated to improve on mastering complex topics that might not be acquired with traditional didactic training methods. There is the vision of a student spending hundreds of hours in a hunt for a solution to a problem that few have solved, for the optimal trade-off between two or more factors, or for a resolution to a set of incompatible constraints. Games are one place to look for a motivating learning environment. Meta-analyses have reported the impact of various game features on learning and motivation (Clark, Tanner-Smith, & Killingsworth, 2014; Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013).

We had the opportunity to apply agents to a STEM learning environment with games and conversational agents. More specifically, studies with agent trialogues held conversations with students as they answered questions and critiqued case studies of scientific research with respect to scientific methodology. The design of the case studies with trialogue critiques was part of a serious game called *Operation ARIES!*, an acronym for **A**cquiring **R**esearch **I**nvestigative and **E**valuative **S**kills (Millis et al., 2017). The system indeed helped students learn scientific reasoning and was beta tested by Pearson Education as *Operation ARA* (Halpern et al., 2012; Millis, Graesser, & Halpern, 2014). Students learned how to critically evaluate research they encounter in various media, such as the Web, TV, magazines, and newspapers. The game taught students how to critically evaluate aspects of scientific investigations (e.g., the need for control groups, adequate samples of observations, operational definitions, etc.) and how to ask appropriate questions that uncovered problems with design or interpretation.

Unfortunately, there were two major obstacles to this gamification approach to enhancing deep learning, over and above AutoTutor's benefits of conversational dialogue and explanation-based reasoning. First, the available evidence suggests that games favor shallow learning over deep learning (Graesser, Hu, Nye, & Sottilare, 2016). Deep learning takes effort, is often frustrating, and is normally regarded as work rather than play (Baker, D'Mello, Rodrigo, & Graesser, 2010; D'Mello, Lehman, Pekrun, & Graesser, 2014). Indeed, the correlation between liking and deep learning tends to be negative in current learning environments without game attributes (Graesser & D'Mello, 2012). Perhaps game features could turn deep learning into play with sufficient entertaining features, learner freedom, and self-regulated activities (Lepper & Henderlong, 2000), and thereby shift the correlation from negative to positive. If not, then games may be reserved for the acquisition of shallow knowledge and skills, such as perceptual motor procedures, memorization of facts, simple skills, and rigid procedures. At this point, examples of games to facilitate deep learning are conspicuously rare, if not absent.

The second plausible obstacle to the marriage between games and deep learning is that games are pursued for pleasurable experiences and fantasy that have minimal explicit connections to the practical world. The pleasurable experience of a game may necessarily be disconnected from any practical value so the game can successfully be an escape. A serious game may be an oxymoron between extrinsic and intrinsic motivation.

Agent Disagreements and Cognitive Disequilibrium

One promising approach to prompting deep learning is to plant contradictions and information that clashes with prior knowledge so that the student experiences cognitive disequilibrium. Cognitive disequilibrium is a state that occurs when people face obstacles to goals, interruptions, contradictions, incongruities, anomalies, impasses, uncertainty, and salient contrasts. Cognitive conflicts can provoke information-seeking behavior (Buder et al., Chap. 3, this volume), which engages the student in inquiry, reasoning, and deep learning (Festinger, 1957; Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005; Otero & Graesser, 2001; Piaget, 1952).

A series of studies had trialogue agents express false information and contradictions during a process of critiquing case studies adopted from ARIES (D'Mello et al., 2014; Lehman et al., 2013). Students critiqued experiments that may or may not have a number of flaws with respect to scientific methodology. For example, one case study described a new pill that purportedly helps people lose weight, but the sample size was small and there was no control group. The trialogue had an expert tutor agent and a peer agent who conversed with the human during the process of identifying potential flaws in the experiment. That is, the tutor agent and student agent engaged in a short exchange about (a) whether there was a flaw in the study and (b) if there *was* a flaw, where it occurred. The tutor agent expressed a correct assertion and the student agent agreed with the tutor in a *True-True* control condition. In the *True-False* condition, the tutor expressed a correct assertion but the student agent disagreed by expressing an incorrect assertion. In the *False-True* condition the student agent provided the correct assertion and the tutor agent disagreed. In the *False-False* condition, the tutor agent provided an incorrect assertion and the student agent agreed.

During the course of the trialogue conversation, the system periodically asked students for their views. For example, the agents asked the human "Do you agree with Chris (student peer agent) that the control group in this study was flawed?" The human's response was coded as correct if he/she agreed with the agent who had made the correct assertion about the flaw of the study. If the human experienced uncertainty and was confused, this should be reflected in either the incorrectness of the human's answer or vacillation when asked multiple questions during the debate. This uncertainty would ideally stimulate or reflect thinking and learning.

The data indeed confirmed that the contradictions and false information had an impact on the humans' answers to these Yes-No questions immediately following a contradiction. The proportion of correct student responses showed the following order: True-True > True-False > False-True > False-False conditions. Students were not frequently confused when both agents agreed and were correct (True-True, no contradiction), but frequently became confused when there was a contradiction between the two agents (True-False or False-True). Confusion was defined as occurring if both (a) the student manifests uncertainty or incorrectness in his decisions when asked by the agents and (b) the student either reports being confused or the computer automatically detects confusion (through technologies that track discourse interaction, facial expressions, and body posture—an area of research that is outside of scope of this chapter; D'Mello & Graesser, 2010; Graesser & D'Mello, 2012).

The ARIES-inspired trialogue studies also showed that contradictions, confusion, and uncertainty caused more learning at deep levels of mastery, as reflected in a delayed test on scientific reasoning. Contradictions in the False-True condition produced higher performance on multiple choice questions that tapped deeper levels of comprehension than performance in the True-True condition. Identification of flaws on far-transfer case studies in a delayed test also showed significant benefits over the True-True condition without contradictions. The results indicated that the most uncertainty occurred when the tutor agent made false claims that the student agent disagreed with, that the contradictions stimulated thought and reasoning at deeper levels, and that scores on a delayed post-test were improved by this experience. The data suggest that there may be a causal relationship between contradictions (and the associated cognitive disequilibrium) and deep learning, with the experience of confusion playing a mediating, moderating, or causal role.

It is illuminating that the False-False condition did not engender much uncertainty and confusion. The students generally accepted what the tutor and student agents expressed when they agreed, even if the claims were false. An alternative possibility could have been that the claims of the two agents would clash with the reasoning and belief system of the human. However, that rarely occurred among the college students in this study. This result is compatible with models that predict that it takes a large amount of world knowledge before a student can detect what they do not know (Miyake & Norman, 1979) and false information (Rapp & Braasch, 2014). A strategic, skeptical, critical stance may be the only hope when the person is not equipped with sufficient subject matter knowledge.

One limitation of this contradiction approach to stimulating deep learning is that the contradictory claims need to be presented contiguously in time and possibly even directly pointed out to the learner. That is, the contradiction is likely to be missed if one agent makes a claim and then another agent makes a contradictory claim 10 min later. Contradictory claims must be co-present in working memory before they are noticed, unless the student has a high amount of subject matter knowledge. Similarly, it is difficult for many students to integrate information from multiple texts and spot contradictions between the texts (Bräten et al., 2011; Goldman et al., 2012; Rouet, 2006; Wiley et al., 2009) unless there is a high amount of subject matter knowledge.

Training Metacognition and Self-Regulated Learning with MetaTutor

MetaTutor was designed to directly train students on metacognitive strategies and self-regulated learning (Azevedo et al., 2010) in the context of hypermedia (see Chap. 10 by Azevedo et al., this volume). MetaTutor trains students on 13 strategies that are theoretically important for self-regulated learning. That is, the process of self-regulated learning involves the learners' constructing a plan, monitoring metacognitive activities, implementing learning strategies, and reflecting on their progress and achievements. The MetaTutor system has a main agent (Gavin) that coordinates the overall learning environment and three satellite agents that handle three phases of self-regulated learning: planning, monitoring, and applying learning strategies. Each of these phases can be decomposed further, under the guidance of the assigned satellite conversational agent. One agent handles metacognitive monitoring, which can be decomposed into judgments of learning, feeling of knowing, content evaluation, monitoring the adequacy of a strategy, and monitoring progress towards goals. Another agent handles learning strategies, including searching for relevant information in a goal-directed fashion, taking notes, drawing tables or diagrams, re-reading, elaborating the material, making inferences, and coordinating information sources (text and diagrams). Each of these metacognitive and self-regulated learning skills had associated measures based on the student's actions, decisions, ratings, and verbal input. The frequency and accuracy of each measured skill was collected throughout the 2-h tutoring session during direct training. This was followed by a test of transfer on a new topic in the area of biology. That is, the training encompassed mechanisms of the heart whereas the transfer subject matter was digestion.

The limitations of MetaTutor are similar to those of SEEK web tutor. Training of these skills takes hundreds if not thousands of hours in many different subject matters and contexts before a noticeable improvement can transfer to new subject matters and problems. A 2-h intervention with direct training is woefully insufficient to show appreciable gains in self-regulated learning.

Personal Assistant for Life-Long Learning (PAL3)

The Personal Assistant for Life-Long Learning (PAL3) project aims to provide training for hundreds to thousands of hours over months or years (Swartout et al., 2016). The long-term goal is to create an agent that will accompany a learner throughout his or her career. To do that, it needs to know and track the individual's background, what the learner has studied, how the learner has performed, various cognitive and noncognitive characteristics of the learner, the goals of the individual, what is needed to successfully meet these goals, and what learning resources to recommend based on the learner's progress.

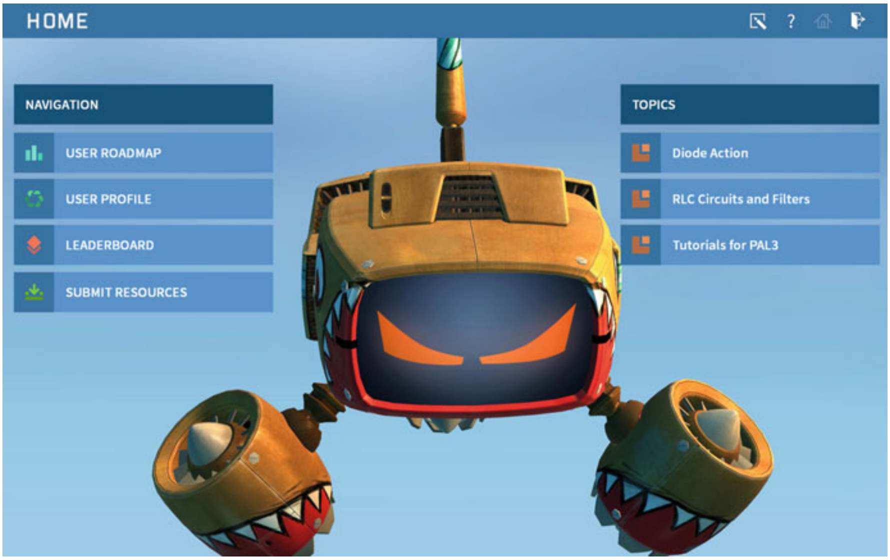

Fig. 12.2 PAL3 home screen with Pal showing emotion (adopted from Swartout et al., 2016)

A first version of PAL3 has been developed for the US Navy. PAL3 taps into a broad range of learning resources tailored to the unique profile of the individual. These learning resources include AutoTutor on the topic of electronics (called ElectronixTutor), several other intelligent tutoring systems, conventional computer-based training systems on a variety of topics, instructional videos, Wikipedia entries, and electronic texts prepared by the Navy. Thus, PAL3 is an intelligent learning guide that tracks learners' historical profiles and recommends learning resources and strategies to help them meet their goals. The major elements of the PAL3 system are: A persistent learning record, the PAL mentoring agent (see Fig. 12.2), a library of learning resources, an algorithm for recommending resources, and mechanisms to promote engagement (including tracking learners' emotions and PAL3 displaying emotions, see Fig. 12.2).

The representation of subject matter knowledge in PAL3, as well as most Auto-Tutor systems, follows the *Topic + Knowledge Component* framework proposed by researchers at the Pittsburgh Science of Learning Center (Koedinger, Corbett, & Perfetti, 2012). A subject matter (e.g., electronic circuits) is defined as a set of topics, with each topic having an associated set of *knowledge components* (KCs). For example, in the ElectronixTutor subject matter, each topic included at least three KCs to cover the *structure* of the circuit, its *behavior*, and its *function*. The system is not strictly hierarchical because one KC can be linked to multiple topics. There are hundreds of KCs in this topic by KC matrix. Mastery of each KC is assessed by the various learning resources in PAL3 that range from conventional computer-based training modules to intelligent tutoring systems.

A profile of learner scores accumulates in a *Learning Record Score* as the learning resources are completed by the individual learner. This information guides the Recommender System on what problems, questions, hints, messages, or other content items to present to the learner next. Sometimes the course curriculum guides the recommended topics, as reflected in a syllabus or day-by-day outline of content covered. At the other end of the continuum, self-regulated learning is available wherein the students select the topics they want to cover in whatever order they wish. In between, the intelligent *Recommender System* can recommend topics considering the history of their performance and psychological attributes. Some of the content in the Learning Record Store addresses subject matter knowledge (i.e., topics and KCs), but other content addresses generic characteristics that range from verbal fluency to grit (i.e., persistence on difficult tasks).

PAL3 is currently being tested on small samples of sailors in the Navy over short time segments so it is too early to report empirical data on learning gains, engagement, and learner impressions. The potential benefits of PAL3 are substantial. It gives *guidance* on planning one's learning experiences and tasks that fulfill *life-relevant goals* aligned with the learner's *self-concept*—all of which are essential for developing self-regulated learning. These broad skills are hopefully internalized when the learners wean themselves off of the PAL3 technologies. PAL3 facilitates *motivation* since individuals work on goals that *interest* them and are *valued* (Pekrun, 2006). The recommender system of PAL3 also always gives the learners a set of *options* to consider rather than forcing them to do one thing—an element of *choice* factors heavily in motivation. A game element adopted by PAL3 is *competition* because the leader board of PAL shows how the learner compares to outstanding peers. And finally there is *feedback* because the learner can see how well they are performing on tasks and projects at varying levels of grain size.

Additional Agent-Based Systems to Promote Deep Learning

IIS researchers are currently working on other agent-based learning environments to promote deep learning, but they have not yet been tested on learning gains, engagement, and learner impressions. These current projects cover knowledge and skills as diverse as reading comprehension strategies, the conceptual foundations of algebra, research methods, medical reasoning, and collaborative learning and problem solving. Below we describe two of these systems.

A system with AutoTutor trialogues has been developed by researchers in the Center for the Study of Adult Literacy (*CSAL*; Graesser et al., 2017) to help struggling adult readers improve their comprehension skills. Approximately 50 million adults in the USA do not read at a level where they can obtain gainful employment (NRC, 2012). CSAL is a national reading research center that focuses on adults who read between the 3.0 and 7.9 grade equivalency levels. In an international survey, it was found that literacy skills and frequency of computer use are related to the ability to use the computer to complete tasks (OECD, 2015).

Both literacy skills and computer proficiency are related to one's wage earnings. Therefore, in CSAL, we believe that it is important to increase both reading skills and practice in using computers to complete tasks.

CSAL AutoTutor primarily emphasizes deeper levels of comprehension, including vocabulary, sentence comprehension, inferences, text cohesion, and rhetorical structures in different types of discourse: narrative, persuasive, and informational texts, as well as forms and digital media. The computer system is grounded in science and practice by aligning: (a) multilevel theoretical frameworks of comprehension (Perfetti, 1999), (b) interventions to improve comprehension (Lovett, Lacerenza, De Palma, & Frijters, 2012), (c) reading curriculum standards, and (d) assessments of comprehension (Sabatini & Albro, 2013). The AutoTutor system also addresses motivational variables, such as independent reading and the selection of documents that consider adult interests, vocations, and practical value in the digital age.

The CSAL AutoTutor curriculum has 35 lessons that focus on specific comprehension strategies (Graesser et al., 2017). The AutoTutor lessons each take 10–50 min to complete and are assigned by teachers on the CSAL research team (in adult literacy centers) in a hybrid intervention with both teachers and computer facilities. The strategies are aligned with the PACES intervention that teachers have successfully implemented (face-to-face) to improve reading comprehension in middle and high school students with reading difficulties (Lovett et al., 2012). The PACES strategies include (a) **P**redicting the purpose and structure of the text with text signals, (b) **A**cquiring vocabulary using context clues and affix knowledge, (c) **C**larifying sources of confusion in text through questioning, (d) **E**valuating, explaining, and elaborating texts through inferences and questioning, and (e) identifying, using, and summarizing text **S**tructures.

The CSAL adult learners typically have substantial challenges with writing, so AutoTutor relies on point & click (or touch) interactions, multiple-choice questions, drag & drop functions, and other conventional input channels. However, the system does include writing components that require semantic evaluation of open-ended student contributions—the signature feature of AutoTutor systems. Other agent-based systems developed in the IIS encourage open-ended student contributions in natural language to promote self-explanation strategies during reading, and these have improved subject matter comprehension (such as *iSTART*, McNamara et al., 2006). However, given that struggling readers tend to write very little, the CSAL AutoTutor has many pictures, diagrams, and multimedia that grab and maintain the attention of the adult learner and that require only mouse clicks or a screen touch to proceed. CSAL AutoTutor can also read texts aloud when the learner asks for such assistance by clicking on a screen option. This is an important feature because many of the adult learners have limited decoding and word identification skills.

The web-based CSAL AutoTutor has been tested on 52 adult readers on 30 lessons in a feasibility study that included both human teachers and AutoTutor, a blended learning environment. These learners completed 71% of the lessons and answered 55% of the questions correctly. The 100-h, 4-week intervention produced increases in comprehension scores in a pretest–post-test design, with effect sizes between 0.4 and 0.7 on major standardized tests of comprehension. However, we do not yet know which component of the intervention was responsible for the increase in comprehension scores.

A second agent-based system, called *AutoMentor*, is being developed to facilitate collaborate learning and problem solving in small groups of three to four learners (Morgan et al., 2013). Much of the planning, problem solving, and decision making in the modern world is performed by teams. Many problems are so difficult that it takes a group of experts with different perspectives and talents to converge on good solutions. The success of a team can be threatened by a social loafer, a saboteur, an uncooperative unskilled member, or a counterproductive alliance; the problem can be mitigated by a strong leader who draws out different perspectives, helps negotiate conflicts, assigns roles, and promotes team communication (Salas, Cooke, & Rosen, 2008). Therefore, there has been discussion in national assessments and the development of educational curricula for including collaborative problem solving (CPS) as an important 21st century skill (Care, Scoular, & Griffin, 2016; Griffin & Care, 2015; Hesse, Care, Buder, Sassenberg, & Griffin, 2015; National Research Council, 2011).

At the international level, CPS was selected by the Organisation for Economic Co-operation and Development (OECD) as a new development for the Programme for International Student Assessment (PISA) in the 2015 international survey of student skills and knowledge (Graesser, Forsyth, & Foltz, 2016; OECD, 2013). Fifteen-year-old students from over 50 countries completed this PISA CPS 2015 assessment in addition to assessments of mathematics, science, literacy, and other proficiencies. Conversational agents were used in the PISA CPS 2015 assessment, following the definition in the assessment framework (OECD, 2013): *Collaborative problem solving competency is the capacity of an individual to effectively engage in a process whereby* ***two or more agents*** *attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills, and efforts to reach that solution.* An agent could be either a human or a computer agent that interacts with the student by this definition. The committee decided to have computer agents in the final assessment. That is, a single student interacted with one to three computer agents in each problem rather than interacting with other humans.

Students theoretically learn best from these collaboration environments when they actively contribute, receive feedback on their contributions, are exposed to multiple perspectives on the problem, and coordinate progress among team members in solving the problem. Agents can play a role in facilitating these group processes. Some example rules are listed below.

1. If the team is stuck and not producing contributions on the relevant topic, then the agent says "What's the goal here?" or "Let's get back on track."
2. If the team meanders from topic to topic without much coherence, then the agent says "I'm lost!" or "What are we doing now?"
3. If the team is saying pretty much the same thing over and over, then the agent says "So what's new?" or "Can we move on?"

4. If a particular team member is loafing, the agent says "What do you think, Harry?"
5. If a particular team member is dominating the conversation excessively, the agent says "I wonder what other people think about this?"
6. If one or more team members express unprofessional language, the agent says "Let's get serious now. I don't have all day."

We are in the process of implementing these rules in a system that involves team learning and problem solving. In the virtual internship *Land Science* (Bagley & Shaffer, 2015), students play the role of interns at Regional Design Associates, a fictional urban and regional planning firm. Their problem-solving task is to prepare a re-zoning plan for the city of Lowell, Massachusetts that addresses the requests of various stakeholder groups (business, environment, industry, or housing) that have views on socioeconomic and ecological issues, some of which are incompatible. The students read about the different viewpoints and preferences of stakeholders and prepare individual reports on how to handle competing concerns. During the course of making these decisions, students discuss options with their project teams through online chat. They also use professional tools, such as a geographic information system model of Lowell and preference surveys to model the effects of land-use changes and obtain stakeholder feedback. At the end of the internship, students write a proposal in which they present and justify their re-zoning plans. During this process, a mentor keeps the small group of three to four students moving forward, but does not encourage any particular solution to the problem-solving tasks at hand. Some of the above rules have been implemented in versions of AutoMentor that have been used in a small sample of participants. Future research is needed to assess the incremental impact of the production rules that guide agents.

Closing Comments

This chapter described some learning environments with computer agents that are designed to help students acquire deep knowledge, strategies, and learning. Some agents implement strategies of good teachers and tutors whereas other agents implement intelligent pedagogical mechanisms that are too complex to be employed by human instructors. There have been notable successes in these agent-based computer systems with respect to enhancing learning gains, motivation, and student impressions of the system, as highlighted in this chapter. However, there have also been disappointments. For example, high knowledge learners often get impatient waiting for the agents to finish speaking. It is easier to cut off the instructor in face-to-face human interactions through barging, intonation, eye gaze, gestures, and body posture. Nevertheless, low-knowledge students often appreciate the computer agents, as in the case of struggling adult readers.

At this point in the science, researchers are sorting out which subject matters, materials, discourse strategies, and learner characteristics account for successful versus unsuccessful applications of agent-based learning environments. When is it desirable to have one or two versus three or more computer agents? What subject materials and strategies do or do not benefit from the agent-based applications? What learner profile benefits most from the agents? What sort of adaptivity in agent conversation enhances learning? What is the role of motivation and emotions in successful agent environments? How do learners perceive the agents initially versus after an hour, a week, a month, or a year? The quality of computer agents has been increasing over the years and particularly during the last decade. The paradigm of developing conversational agents is destined to have a vibrant future.

Acknowledgments The research was supported by the National Science Foundation (0325428, 633918, 0834847, 0918409, 1108845, 1443068), the Institute of Education Sciences (R305B07460, R305B070349, R305A080594, R305C120001), and the Office of Naval Research (N00014-00-1-0600, N00014-12-C-0643; N00014-16-C-3027). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these funding sources.

References

Autor, D., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, *118*, 1279–1334.

Azevedo, R., Moos, D., Johnson, A., & Chauncey, A. (2010). Measuring cognitive and metacognitive regulatory processes used during hypermedia learning: Issues and challenges. *Educational Psychologist*, *45*, 210–223.

Bagley, E., & Shaffer, D. W. (2015). Learning in an urban and regional planning practicum: The view from educational ethnography. *Journal of Interactive Learning Research*, *26*(4), 369–393.

Baker, L. (1985). Differences in standards used by college students to evaluate their comprehension of expository prose. *Reading Research Quarterly*, *20*, 298–313.

Baker, R. S., D'Mello, S. K., Rodrigo, M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*, 223–241.

Biswas, G., Jeong, H., Kinnebrew, J., Sulcer, B., & Roscoe, R. (2010). Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology-Enhanced Learning*, *5*, 123–152.

Braasch, J., Rouet, J.-F., Vivert, N., & Britt, M. (2012). Readers' use of source information in text comprehension. *Memory & Cognition*, *40*, 450–465.

Bräten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist*, *46*, 48–70.

Care, E., Scoular, C., & Griffin, P. (2016). Assessment of collaborative problem solving in education environments. *Applied Measurement in Education*, *29*(4), 250–264.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73–105.

Carnevale, A. P., & Smith, N. (2013). Workplace basics: The skills employees need and employers want. *Human Resource Development International*, *16*, 491–501.

Clark, D., Tanner-Smith, E., & Killingsworth, S. (2014). *Digital games, design and learning: A systematic review and meta-Analysis*. Menlo Park, CA: SRI International.

Craig, S. D., Gholson, B., Brittingham, J. K., Williams, J. L., & Shubeck, K. (2012). Promoting vicarious learning of physics using deep questions with explanations. *Computers & Education*, *58*, 1042–1048.

D'Mello, S., & Graesser, A. C. (2010). Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-adapted Interaction*, *20*, 147–187.

D'Mello, S., Lehman, S., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170.

Dunlosky, J., & Lipko, A. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, *16*, 228–232.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Gholson, B., Witherspoon, A., Morgan, B., Brittingham, J. K., Coles, R., Graesser, A. C., ... Craig, S. D. (2009). Exploring the deep-level reasoning questions effect during vicarious learning among eighth to eleventh graders in the domains of computer literacy and Newtonian physics. *Instructional Science*, *37*, 487–493.

Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012). Comprehending and learning from internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly*, *47*, 356–381.

Graesser, A. C. (2015). Deeper learning with advances in discourse science and technology. *Policy Insights from Behavioral and Brain Sciences*, *2*, 42–50.

Graesser, A. C. (2016). Conversations with AutoTutor help students learn. *International Journal of Artificial Intelligence in Education*, *26*, 124–132.

Graesser, A. C., Cai, Z., Baer, W., Olney, A. M., Hu, X., Reed, M., & Greenberg, D. (2016). Reading comprehension lessons in AutoTutor for the Center for the Study of Adult Literacy. In S. Crossley & D. McNamara (Eds.), *Adaptive educational technologies for literacy instruction* (pp. 288–292). New York, NY: Routledge.

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions in Education*, *48*, 612–618.

Graesser, A., & D'Mello, S. K. (2012). Emotions during the learning of difficult material. In B. Ross (Ed.), *Psychology of learning and motivation* (Vol. 57, pp. 183–225). New York, NY: Elsevier.

Graesser, A. C., Forsyth, C. M., & Foltz, P. (2016). Assessing conversation quality, reasoning, and problem solving performance with computer agents. In B. Csapo, J. Funke, & A. Schleicher (Eds.), *On the nature of problem solving: A look behind PISA 2012 problem solving assessment* (pp. 275–297). Heidelberg, Germany: OECD Series.

Graesser, A. C., Forsyth, C., & Lehman, B. (2017). Two heads are better than one: Learning from agents in conversational trialogues. *Teacher College Record*, *119*(3), 1–20.

Graesser, A. C., Hu, X., Nye, B., & Sottilare, R. (2016). Intelligent tutoring systems, serious games, and the Generalized Intelligent Framework for Tutoring (GIFT). In H. F. O'Neil, E. L. Baker, & R. S. Perez (Eds.), *Using games and simulation for teaching and assessment* (pp. 58–79). Abingdon, Oxon, UK: Routledge.

Graesser, A. C., Langston, M. C., & Baggett, W. B. (1993). Exploring information about concepts by asking questions. In G. V. Nakamura, R. M. Taraban, & D. Medin (Eds.), *The psychology of learning and motivation*, *Categorization by humans and machines* (Vol. 29, pp. 411–436). Orlando, FL: Academic Press.

Graesser, A. C., Li, H., & Forsyth, C. (2014). Learning by communicating in natural language with conversational agents. *Current Directions in Psychological Science*, *23*, 374–380.

Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments, and Computers*, *36*, 180–193.

Graesser, A. C., Lu, S., Olde, B. A., Cooper-Pye, E., & Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory and Cognition*, *33*, 1235–1247.

Graesser, A. C., McNamara, D. S., & VanLehn, K. (2005). Scaffolding deep comprehension strategies through Point&Query, AutoTutor, and iSTART. *Educational Psychologist*, *40*, 225–234.

Graesser, A., Ozuru, Y., & Sullins, J. (2009). What is a good question? In M. G. McKeown & L. Kucan (Eds.), *Threads of coherence in research on the development of reading ability* (pp. 112–141). New York, NY: Guilford.

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, *31*, 104–137.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371–395.

Graesser, A. C., Wiley, J., Goldman, S. R., O'Reilly, T., Jeon, M., & McDaniel, B. (2007). SEEK Web tutor: Fostering a critical stance while exploring the causes of volcanic eruption. *Metacognition and Learning*, *2*, 89–105.

Greiff, S., Wüstenberg, S., Holt, D. V., Goldhammer, F., & Funke, J. (2013). Computer-based assessment of Complex Problem Solving: Concept, implementation, and application. *Educational Technology Research and Development.*, *61*, 407–421.

Griffin, P., & Care, E. (2015). ATC21S method. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills: Methods and approach*. Dordrecht: Springer.

Halpern, D. F., Millis, K., Graesser, A. C., Butler, H., Forsyth, C., & Cai, Z. (2012). Operation ARA: A computerized learning game that teaches critical thinking and scientific reasoning. *Thinking Skills and Creativity*, *7*, 93–100.

Hattie, J. A. C., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *Nature Partner Journal: Science of Learning*, *1*, 1–13.

Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century sills* (pp. 37–55). Heidelberg, GA: Springer.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*, 141–153.

Jackson, G. T., & Graesser, A. C. (2006). Applications of human tutorial dialog in AutoTutor: An intelligent tutoring system. *Revista Signos*, *39*, 31–48.

Jackson, G. T., & McNamara, D. S. (2013). Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology*, *105*, 1036–1049.

Johnson, W. L., & Lester, J. C. (2016). Face-to-face interaction with pedagogical agents, Twenty years later. *International Journal of Artificial Intelligence in Education*, *26*(1), 25–36.

Johnson, L. W., & Valente, A. (2009). Tactical language and culture training systems: Using artificial intelligence to teach foreign languages and cultures. *AI Magazine*, *30*, 72–83.

Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, *35*, 1567–1577.

Koedinger, K. R., Corbett, A. C., & Perfetti, C. (2012). The Knowledge-Learning-Instruction (KLI) framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, *36*(5), 757–798.

Kopp, K., Britt, A., Millis, K., & Graesser, A. (2012). Improving the efficiency of dialogue in tutoring. *Learning and Instruction*, *22*(5), 320–330.

Lehman, B., D'Mello, S. K., Strain, A., Mills, C., Gross, M., Dobbins, A., … Graesser, A. C. (2013). Inducing and tracking confusion with contradictions during complex learning. *International Journal of Artificial Intelligence in Education*, *22*, 85–105.

Lepper, M. R., & Henderlong, J. (2000). Turning "play" into "work" and "work" into "play": 25 years of research on intrinsic versus extrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 257–307). San Diego, CA: Academic Press.

Levy, F., & Murnane, R. J. (2006). Why the changing American economy calls for 21st century learning: Answers to educators' questions. *New Directions for Youth Development*, *110*, 53–62.

Lorch, R. F. (2015). What about expository text? In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp. 348–361). Cambridge, UK: Cambridge University Press.

Lorch, R. F., Jr., Lorch, E. P., Calderhead, W. J., Dunlap, E. E., Hodell, E. C., & Freer, B. D. (2010). Learning the control of variables strategy in higher and lower achieving classrooms: Contributions of explicit instruction and experimentation. *Journal of Educational Psychology*, *102*(1), 90–101.

Lovett, M. W., Lacerenza, L., De Palma, M., & Frijters, J. C. (2012). Evaluating the efficacy of remediation for struggling readers in high school. *Journal of Learning Disabilities.*, *45*, 151–169.

Maki, R. H. (1998). Test predictions over text material. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117–144). Mahwah, NJ: Erlbaum.

McNamara, D. S., O'Reilly, T., Best, R., & Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research*, *34*, 147–171.

Medimorecc, M. A., Pavlik, P., Olney, A., Graesser, A. C., & Risko, E. F. (2015). The language of instruction: Compensating for challenge in lectures. *Journal of Educational Psychology*, *107*, 971–990.

Millis, K., Forsyth, C., Wallace, P., Graesser, A. C., & Timmins, G. (2017). The impact of game-like features on learning from an intelligent tutoring system. *Technology, Knowledge, and Learning*, *22*(1), 1–22.

Millis, K., Graesser, A. C., & Halpern, D. F. (2014). Operation ARA: A serious game that combines intelligent tutoring and learning principles to teach science. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum* (pp. 169–183). Washington, DC: Society for the Teaching of Psychology Series.

Miyake, N., & Norman, D. A. (1979). To ask a question one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior*, *18*, 357–364.

Morgan, B., Keshtkar, F., Graesser, A., & Shaffer, D. W. (2013). Automating the mentor in a serious game: A discourse analysis using finite state machines. In C. Stephanidis (Ed.), *Human computer interaction international 2013* (Vol. 374, pp. 591–595). Berlin: Springer.

National Research Council. (2012). *Improving adult literacy instruction: Options for practice and research*. Washington, DC: The National Academies Press.

National Research Council. (2011). *Assessing 21st century skills*. Washington, DC: National Academies Press.

Nye, B. D., Graesser, A. C., & Hu, X. (2014). AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education*, *24*(4), 427–469.

OECD. (2013). *PISA 2015 collaborative problem solving framework*. Paris, France: OECD. Retrieved from http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf

OECD (2015). *Adults, computers and problem solving: What's the problem?* Paris, France: OECD Publishing. Retrieved August 27, 2016, from http://dx.doi.org/10.1787/9789264236844-en.

Olney, A., D'Mello, S. K., Person, N., Cade, W., Hays, P., Williams, C., ... Graesser, A. C. (2012). Guru: A computer tutor that models expert human tutors. In S. Cerri, W. Clancey, G. Papadourakis, & K. Panourgia (Eds.), *Proceedings of Intelligent Tutoring Systems (ITS) 2012* (pp. 256–261). Berlin, Germany: Springer.

Otero, J., & Graesser, A. C. (2001). PREG: Elements of a model of question asking. *Cognition & Instruction*, *19*, 143–175.

Pashler, H., Bain, P., Bottge, B., Graesser, A. C., Koedinger, K., McDaniel, M., & Metcalf, J. (2007). *Organizing instruction and study to improve student learning: A practice guide (NCER 2007–2004)*. Washington, DC: Institute of Education Sciences.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*, 315–341.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 167–208). Oxford: Oxford University Press.

Piaget, J. (1952). *The origins of intelligence*. New York, NY: International University Press.

Rapp, D. N., & Braasch, J. L. G. (Eds.). (2014). *Processing inaccurate information*. Cambridge, MA: MIT Press.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, *66*, 181–221.

Rouet, J.-F. (2006). *The skills of document use*. Mahwah, NJ: Erlbaum.

Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2011). Integrating learning, problem solving, and engagement in narrative-centered learning environments. International Journal of Artificial Intelligence in Education, 21, 115–133.

Rus, V., D'Mello, S., Hu, X., & Graesser, A.C. (2013). Recent advances in intelligent systems with conversational dialogue. AI Magazine, 34, 42–54.

Rus, V., Lintean, M., Graesser, A. C., & McNamara, D. S. (2012). Text-to-text similarity of statements. In P. McCarthy & C. Boonthum-Denecke (Eds.), *Applied natural language processing: Identification, investigation, and resolution* (pp. 110–121). Hershey, PA: IGI Global.

Sabatini, J. P., & Albro, E. (2013). *Assessing reading in the 21st century: Aligning and applying advances in the reading and measurement sciences*. Lanham, MD: R&L Education.

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries and developments. *Human Factors*, *50*(3), 540–547.

Swartout, W., Nye, B. D., Hartholt, A., Reilly, A., Graesser, A. C., VanLehn, K., ... Rosenberg, M. (2016). Designing a personal assistant for life long learning (PAL3). In Z. Markov & I. Russel (Eds.), *Proceedings of the 29th International Florida Artificial Intelligence Research Society Conference* (pp. 491–496). Palo Alto, CA: Association for the Advancement of Artificial Intelligence.

Van den Broek, P., Bohn-Gettler, C., Kendeou, P., Carlson, S., & White, M.J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M.T. McCrudden, J. Magliano, & G. Schraw (eds.), Relevance instructions and goal-focusing in text learning (pp. 123-140). Greenwich, CT: Information Age Publishing.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rose, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, *31*, 3–62.

Ward, W., Cole, R., Bolaños, D., Buchenroth-Martin, C., Svirsky, E., & Weston, T. (2013). My science tutor: A conversational multimedia virtual tutor. *Journal of Educational Psychology*, *105*, 1115–1125.

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*, *46*, 1060–1106.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*, 249–265.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185.

Index

J. Buder, F.W. Hesse (eds.), *Informational Environments*,
DOI 10.1007/978-3-319-64274-1