Freie Universität Berlin

Fachbereich Erziehungswissenschaft und Psychologie

der Freien Universität Berlin

Adaptive e-Learning: Towards a Metacomprehension Approach

Dissertation

zur Erlangung des akademischen Grades

Doktor der Philosophie (Dr. phil.)

vorgelegt von

Julia Vössing (M. Sc.)

Berlin, 2016

Erstgutachter:	Prof. Dr. Kathrin Heinitz
Zweitgutachter:	Prof. Dr. Christian Stamov-Roßnagel
Disputation:	13. Juli 2016

Acknowledgments

For reasons of data protection, the acknowledgements are not included in this version.

Eidesstaatliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und ohne unzulässige Hilfe verfasst habe. Für die im Rahmen der Promotion durchgeführten Studien und die Erstellung der Fachartikel war ich hauptverantwortlich. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Frankfurt, den 28. Juli 2016

Julia Vössing

List of Publications

The dissertation is based on the following three articles. The co-author(s) can confirm that I have been the main responsible person for the realization of all of these projects.

Vössing, J., & Stamov-Roßnagel, C. (2016). Boosting metacomprehension accuracy in computer-supported learning: The role of judgment task and judgment scope. *Computers in Human Behavior*, *54*, 73-82.

Vössing, J., Stamov-Roßnagel, C., & Heinitz, K. (submitted). The Impact of Text Difficulty on Metacomprehension Accuracy and Knowledge Test Performance in Text Learning.

Vössing, J., Stamov-Roßnagel, C., & Heinitz, K. (2016). Images in computer-supported learning: Increasing their benefits for metacomprehension through judgments of learning. *Computers in Human Behavior, 58*, 221-230.

Table of Contents

List of Tables

List of Figures

Summary

This dissertation presents work contributing to a further improvement of e-learning using a metacomprehension approach. Research interest is increasing in using metacomprehension judgments (e.g., *judgments of learning*, JOL) in computer-assisted learning. The usefulness of such judgments as tools for research or for learner support hinges on their accuracy, which is notoriously low. Although JOL generally predict learning performance their accuracy seems to be influenced both by features of the learning material (e.g. amount and size of text) and by personal factors (e.g. self-efficacy). This dissertation therefore aims to sustainably boost metacomprehension accuracy as a vital requirement for further improving self-regulated computer-supported learning.

The rationale of Study I was that of a manipulation to increase *experience-based* judgments and decrease *theory-based* judgments. In terms of an experimental manipulation, the question therefore was whether judgments of difficulty (JOD) administered before the JOL proper influence JOL accuracy. In the simple judgments group, participants rated the likelihood of answering correctly a knowledge question on e-learning chapters (JOL). In the combined judgments group, participants rated text difficulty (JOD) before making a JOL. In the first experiment no accuracy differences emerged, but additional analyses showed that the judgment manipulation had induced cognitive processing differences. Therefore, in the second experiment of Study I judgment scope was manipulated. Both judgment accuracy and knowledge test scores were higher in the combined judgments group with term-specific judgments.

Study II sought to add knowledge about text difficulty effects on metacomprehension accuracy. Sequences of chapters with the same text difficulty and

sequences of chapters with randomly varying text difficulty may entail different anchors and thus prime higher or lower JOL accuracies, respectively. The general tenet was that orders would trigger different extents of experience-based processing and thus influence metacomprehension accuracy to different degrees. To investigate the impact of text difficulty on metacomprehension accuracy and learning performance four groups of learners who perform e-learning differing in blocking type (blocked vs. mixed) and difficulty sequence (rising vs. declining) of text material were compared. As hypothesized, accuracy was higher for blocked difficulty orders. Late-section judgment magnitude decreased more strongly in the blocked groups. At the same time, late-section judgment accuracy was higher in the blocked group.

Finally, Study III extended the previous findings by analysing the impact of the so called multimedia heuristic on judgment accuracy. For that purpose, a conceptual images group was compared to a decorative images group and a text-only group. Depending on the condition either conceptual images summarizing the chapters' content or decorative images without any relevant information were added to the e-learning. The control group worked on a text-only version of the same e-learning unit. As postulated, combined JOL benefitted accuracy and knowledge test scores; both were highest in the conceptual images group. In Experiment 2 of Study III, to further increase accuracy, judgment scope was changed from global to specific. Again, accuracy and test scores were highest in the conceptual images group. Contrary to expectations, however, JOL accuracy did not benefit from term-specific judgments.

In sum, the work presented in the this dissertation demonstrates ways to sustainably boost metacomprehension accuracy and thereby establishes the requirements needed to use the metacomprehension approach for personalizing e-learning and increasing its learning efficiency as well as effectiveness.

Deutsche Zusammenfassung

Die vorliegende Dissertation zeigt Arbeiten, die mit Hilfe eines metakognitiven Ansatzes zu einer weiteren Verbesserung von E-Learning beitragen. Das wissenschaftliche Interesse zur Nutzung von metakognitiven Urteilen (z.B. Lernurteil, *judgment of learning*, JOL) bei computergestütztem Lernen steigt stetig an. Der Wert dieser Urteile als Instrument für die Forschung oder die Unterstützung von Lernenden hängt jedoch von ihrer Genauigkeit ab, welche bisher notorisch niedrig ist. Obwohl Lernurteile insgesamt die Lernleistung vorhersagen können, scheint ihre Genauigkeit sowohl durch das Lernmaterial selbst (z.B. Textmenge und Schriftgröße) als auch durch Persönlichkeitsfaktoren der Lernenden (z.B. Selbstwirksamkeit) beeinflusst zu werden. Das Ziel dieser Dissertation ist es deshalb die Genauigkeit der metakognitiven Urteile nachhaltig zu erhöhen, um so die zentrale Voraussetzung zur weiteren Optimierung von E-Learning als selbstreguliertes Lernen am Computer zu schaffen.

Ziel von Studie I war es durch eine Manipulation *erfahrungsbasierte* Lernurteile zu erhöhen und gleichzeitig *theoriebasierte* Lernurteile zu verringern. Im Sinne einer experimentellen Manipulation lautete die Forschungsfrage, ob Schwierigkeitsurteile (*judgments of difficulty*, JOD), welche vor den eigentlichen Lernurteilen (JOL) abgegeben werden, die Genauigkeit der Lernurteile (JOL) auf geeignete Weise verbessern. In der „ein-Urteil“-Gruppe, bewerteten die Teilnehmer die Wahrscheinlichkeit eine Wissensfrage zum jeweiligen E-Learning-Kapitel korrekt zu beantworten (JOL). In der „zwei-Urteil“-Gruppe, bewerteten die Teilnehmer jeweils zunächst die Textschwierigkeit (JOD) bevor sie das eigentliche Lernurteil (JOL) abgaben. Im ersten Experiment fanden sich keine Genauigkeits-Unterschiede, aber vertiefende Analysen zeigten, dass die Manipulation zu unterschiedlicher kognitiver

Verarbeitung geführt hatte. Deshalb wurde im zweiten Experiment in Studie 1 der Urteilsumfang manipuliert. Sowohl die Urteilsgenauigkeit als auch die erzielten Punkte im Wissenstest waren nun in der „zwei-Urteil"-Gruppe mit spezifischem Urteilsumfang höher.

Studie II zielte darauf ab das Verständnis von Textschwierigkeitseffekten auf die metakognitive Genauigkeit zu erhöhen. Sequenzen mit gleicher oder zufälliger Textschwierigkeit könnten als Anker fungieren und so höhere oder niedrigere Urteilsgenauigkeiten induzieren. Um die Wirkung von Textschwierigkeit auf metakognitive Genauigkeit und Lernleistung zu untersuchen, wurden vier Gruppen von Lernern verglichen, die jeweils ein E-Learning mit unterschiedlichem Aufbau (in Blöcken vs. gemischt) und Schwierigkeits-Sequenzen (ansteigend vs. absteigend) bearbeiteten. Wie vermutet, war die Genauigkeit höher mit Schwierigkeits-Blöcken. Die Lernurteile für spätere E-Learning-Kapitel sanken in Gruppen mit Schwierigkeits-Blöcken stärker ab. Gleichzeitig stieg die Genauigkeit für spätere E-Learning-Kapitel in den Gruppen mit Schwierigkeits-Blöcken an.

Schließlich erweiterte Studie III die vorherigen Ergebnisse durch die Untersuchung von möglichen Auswirkungen der so genannten Multimedia-Heuristik auf die Urteilsgenauigkeit. Zu diesem Zweck wurde eine "konzeptionelle Bilder"-Gruppe mit einer "dekorative Bilder"-Gruppe und einer „Text"-Gruppe verglichen. Je nach Bedingung enthielt das E-Learning entweder konzeptionelle Bilder, die die Kapitelinhalte zusammenfassten oder dekorative Bilder ohne relevante Informationen. Die Kontrollgruppe erhielt eine reine Textversion des gleichen E-Learning. Wie vermutet wirken sich kombinierte Urteile positiv auf die Genauigkeit und Lernleistung aus, beide waren am höchsten mit den konzeptionellen Bildern.

In Experiment 2 von Studie III, wurde der Urteilsumfang von global zu spezifisch verändert, um die Genauigkeit weiter zu erhöhen. Erneut waren die Genauigkeit und die Lernleistung mit den konzeptionellen Bildern am höchsten. Anders als erwartet wirkten sich die spezifischen Urteile jedoch nicht positiv auf die Urteilsgenauigkeit aus.

Zusammengefasst zeigen die vorgestellten Arbeiten dieser Dissertation Möglichkeiten auf, die Genauigkeit metakognitiver Urteile nachhaltig zu erhöhen und so die benötigte Voraussetzung für die Verwendung des metakognitiven Ansatzes zur Personalisierung von E-Learning sowie der Erhöhung von dessen Lerneffizienz und Effektivität zu schaffen.

Chapter 1

1 General Introduction

1.1 Adaptive e-learning: Towards a metacomprehension approach

Lifelong learning is an important challenge of the 21st century. Regarding the growing virtual environment and so called "digital natives" e-learning seem to be a good choice to realize the challenge of lifelong learning. But to actually meet this challenge e-learning needs to be cheaper, more efficient and more effective.

Therefore, a promising approach to further improve e-learning is to use learners' metacomprehension judgments, i.e. their self-assessments of ongoing learning processes. These are usually captured by asking participants to rate the likelihood they will correctly answer questions on the content at hand (*judgments of learning*, JOL). In principle, that *metacomprehension approach* might help to personalise e-learning and thereby increase learning efficiency as well as effectiveness. Before detailing the research aims of the dissertation project, a short overview is given over the current state of research concerning e-learning adaption, metacomprehension and JOL accuracy.

Research interest in adaptive e-learning systems has increased considerably recently. Adaptive systems adjust learning contents, pedagogical models, or user interactions to individual learning needs and preferences (Stoyanov & Kirchner, 2004). Adaptive systems are thought to help save learning time whilst increasing retention (Conlan, Hockemeyer, Wade, & Albert, 2002), reduce cognitive load (Corbalan, Kester, & Van Merriënboer, 2006), increase learners' orientation (Chen, Fan, & Macredie, 2006) and promote learning motivation (Kareal & Klema, 2006).

Adaption may be achieved on two fundamentally different routes (Fiorella, Vogel-Walcutt, & Fiore, 2012). In *direct adaption,* system parameters such as learner control, e.g., fixed sequence vs. free navigation, or display complexity may be adjusted to learner variables such as prior knowledge or learning styles. Prompts, on the other hand, might be seen as *indirect adaption*. Rather than manipulating system parameters, prompts seek to influence learners' information processing, whilst leaving the "system core" unaltered. Prompts make a system adaptive to the extent that they seek to support different phases of the learning, e.g., forethought, planning or goal-setting, at certain points of system use. Also, different types of prompts like metacognitive, cognitive or semantic prompts may be used depending on learner type (Fiorella et al., 2012).

Independent of the direct-indirect dimension, adaption may be *off-line* or *on-line*. In *off-line* adaption, learner variables are assessed prior to learning and system parameters or prompts are adjusted according to this assessment. Typical variables are expertise (Kalyuga, 2006), learning styles, epistemic beliefs, attribution styles, age, and sex (Harteis, Gruber, & Hertramph, 2010; Kabassi & Virvou, 2006). In *on-line* adaption, system use or learning process variables are measured during learning. Variables include navigation behaviour (Puntambekar & Stylianou, 2005), self-reported cognitive load (Van Gog & Paas, 2008), or knowledge acquisition (Kalyuga, 2006), but also learning strategies and metacognitive skills (Schwonke, Hauser, Nückles, & Renkl, 2006). On-line and off-line adaption may be combined, e.g., by providing different adaptive prompt types (on-line component) for different learner types (off-line component).

1.2 Potential benefits of metacomprehension prompting in e-learning

In light of an increasing reliance on and investments in e-learning, adaptive e-learning systems should appeal to companies to the extent that they increase efficiency by reducing study time and that they foster sustainable learning motivation, which is important given the need for lifelong learning. Most of the so far mentioned adaption methodologies are technical solutions. Another approach is to focus on the learners themselves and base e-learning adaption on their metacomprehension. However, little empirical research has addressed adaptive systems outside the laboratory. One potential limitation is the generalisability of adaptive systems, which analyse learners' behaviours to compute learning support requirements (Aleven, Roll, McLaren, & Koedinger, 2010). Such approaches require content modifications, which may be rather costly in terms of content author time. Also, no one learner model has yet been developed that captures all – or at least the most frequent – learner types that occur in practice. Lab studies usually cross relatively clear-cut conditions such as prior knowledge (high vs. low) with prompting (optional vs. compulsory) (Schnotz & Heiß, 2009), but a host of other factors might influence the adaption-performance relationship.

To overcome the shortcomings of content-modified systems and previous lab studies, *metacomprehension prompting* appears, beside technical adaption methodologies, to be an attractive option to increase efficiency of e-learning sustainably. The basic assumption stemming from the monitoring-affects-control hypothesis is that learners continuously monitor their learning (Nelson & Leonesio, 1988). On a meta-level they make so called judgments of learning (JOL), which are metacomprehensive estimations of their actual learning success, and adjust their learning behaviours based on these monitoring outcomes, i.e. their JOL. JOL measure people's metacomprehension accuracy and could in principle be used to trigger system responses. For instance, if a learner rates the probability of her answer correct above 90%, she may proceed to the next section. If she rates the probability below, say, 40%, she might be prompted, for instance, to immediately re-study the section or mark it for later re-study. In that way JOL magnitude is used as an indicator whether a learner needs support at a given point in her learning episode. Based on the JOL navigation is adapted.

Metacomprehension accuracy research shows that JOL are significantly correlated with objective learning outcomes, albeit correlations range from .27 to .80 (Maki, 1998). From an applied point of view, the appeal of metacomprehension prompts lies in the fact that they are largely independent of learning content. They can therefore be implemented in a rather schematic fashion like prompting on discrete system events such as proceeding to the next chapter, making the development of adaptive systems a relatively fast and inexpensive process. Also, from a solid body of metacomprehension research (Dunlosky & Lipko, 2007) a variety of prompting techniques might be derived to maximize prompting benefits.

The goal of adaption through JOL is to reconcile in the best possible way the conflicting requirements of personalised learner support and cost-effective implementation. No modification of the content will be required, nor will multiple versions of the same content need to be developed. The prompts support and modify learners' *approach to the content,* e.g., intermediate self-testing, but leave the content proper unchanged. Therefore, the prompts are *portable* and may easily be migrated to virtually any content.

First evidence suggests that the JOL approach may be applied to hypermedia settings. Mengelkamp and Bannert (2009) found in their study significant correlations between retrospective JOL and objective learning performance in a knowledge test. Serra and Dunlosky (2010) had their participants make segment-wise JOL as they read a text passage. Whilst these authors found systematic variations of JOL magnitude, a substantial proportion of those variations could be attributed to bias. In the superiority of multimedia formats epistemic beliefs were coupled with higher JOL even when text-diagram combinations were arranged so as not to support learning. In other words, multimedia beliefs had overrated on-line monitoring and influenced JOL magnitude.

Serra and Dunlosky (2010) therefore called for developing techniques to align judgments with learning and avoid misapplication of heuristics.

Ikeda, Kitagami, Takahashi, Hattori, and Ito's (2013) findings point in a similar direction. In their study, participants read a text about neuroimaging findings on which they later took a comprehension test. In the multimedia groups, the text came with either brain images or bar graphs. Metacomprehension judgments were higher in the multimedia groups than in the plain text condition, but test performance was not. Like Serra and Dunlosky's (2010) study, these findings indicate that heuristics inflated the judgments of learning.

1.3 The accuracy problem

Metacomprehension accuracy is commonly relatively low (Maki, 1998) and constrained by theory-based processing, such as the multimedia heuristic (cf. Serra & Dunlosky, 2010). Theory-based judgments of learning are based on general beliefs or domain knowledge, whereas experience-based JOL use specific cues from the information processing during learning or attempting to answer a question. In other words, directing participants' attention to their information processing during a given text passage might help to generate the more valid experience-based JOL, rather than invalid theory-based ones (Ozuru, Kurby, & McNamara, 2012). Therefore, research is needed how to boost experience-based processing in order to increase metacomprehension accuracy notably. Numerous studies have tried to find ways of increasing metacomprehension accuracy by analysing influencing factors. Such factors could be cue-utilization (Koriat, 1997), judgment scope (Dunlosky, Rawson, & Middleton, 2005) or text difficulty (Linderholm, Wang, Therriault, Zhao, & Jakiel, 2012). Son and Metcalfe (2005) stated that JOL are based on a two-stage process comprising a quick pre-retrieval stage at which stimulus familiarity is analysed, and a second stage in which judgments of learning are based on retrieval fluency or on stimulus characteristics that become apparent after attempted retrieval.

More recent research showed that the multimedia heuristic has an impact on metacomprehension accuracy (e.g. Ackerman, Leiser, & Shpigelman, 2013; Butcher, 2006; Paik & Schraw, 2013; Serra & Dunlosky, 2010), too.

To actually use JOL for adapting e-learning to personal needs and thereby increase efficiency these judgments first of all have to be completely understood. Thus, two conditions must be met for metacomprehension prompts to be a valid adaption methodology. First, metacomprehension judgments must not be epiphenomenal, but actually control learning behaviour. If metacomprehension were epiphenomenal, they were merely feelings without impact on learning behaviour. Research suggests, however, that people preferentially study items that they believe to have not learned well. Son and Metcalfe (2000) analysed the allocation of study time and found that people allocated more time to difficult items. Furthermore, Thiede, Anderson and Therriault (2003) showed that when more accurate judgments of text comprehension were induced, people restudied more strategically and performed better on a final test. Such results invite the interpretation that people were studying strategically, based on metacomprehension.

Second, for metacomprehension prompts to be a valid adaption methodology, judgments must be accurate. Whilst research seems to rule out the possibility that judgments are merely epiphenomenal, much less is known about how to reliably boost judgment accuracy. The aforementioned findings also underline the need for metacognitive judgments to be accurate and suggest that accuracy may not be the default. Indeed, systematic overconfidence and underconfidence effects have been reported (see Dunlosky & Rawson, 2012) that seem to depend on characteristics of learning materials, learner variables, or both. Simple extrapolation shows that inaccurate judgments can hardly be used to form adaptive prompting. Imagine a learner makes a JOL during a study, rating the probability of an answer correct at 100%. If this judgment was accurate, the system would let the learner proceed to the next section. If the judgment was overconfident, having the learner perform further learning on that section might be the better option. She would in this case be given a second-stage prompt inducing additional learning. In order to boost accuracy it is therefore also important to improve learners monitoring of their learning (Todorov, Kornell, Larsson Sundqvist, & Jönsson, 2013).

Furthermore the judgment task itself might influence metacomprehension accuracy. Ozuru, Kurby and McNamara (2012) distinguished between judgments of difficulty and predictions of performance each aligning with different types of processing cues. Being forced to make two types of metacomprehension judgments during reading let readers to attend to different information sources. Depending on the information source being used different factors influenced metacognitive accuracy. Zhao and Linderholm (2008) also found further constraints to metacomprehension accuracy. According to their findings the anchoring effect and the poor diagnostic validity of experimental cues reduce the accuracy of serial JOL due to test uncertainty.

To sum up, while the first condition for metacomprehension prompts to be a valid adaption methodology is met because metacomprehension is not epiphenomenal, much less is known about the second condition. How to boost accuracy of metacomprehension prompts to an appropriate level needs to be further investigated in order to adapt e-learning based on metacomprehension judgment to personal needs and thereby raise its efficiency.

1.4 Research aims and design of the dissertation project

Although JOL predict learning performance, correlations might be as low as .27 (Maki, 1998). This JOL magnitude seems to be influenced both by features of the learning material (e.g., amount and size of text) and by personal factors (e.g., self-efficacy, e-learning beliefs).

Based on the above considerations, the main research question of the upcoming studies therefore is: *What prompting format best boosts judgment accuracy?* The prompt format can be varied in two ways. First, a variation of the judgment criterion leads to prompt format variation. Transfer appropriate processing has been shown to have an influence on accuracy (Bransford, Franks, Morris, & Stein, 1979). Typically, learners are *globally* asked how well they think they will be able to answer knowledge questions about a text (cf. Appendix). Alternatively, one might ask learners how well they think they will be able to answer knowledge questions about *specific* concepts from a text or how confident they will be to generate keywords for a text (cf. Appendix). It seems essential, the more similar judgment and test criterion (e.g. global vs. specific or question vs. keyword) are, the greater the accuracy.

Second, for finding the prompting format that best boots judgment accuracy, the prompt format can be varied by making diagnostic information available prior to judgments in an effort to increase accuracy. Recent research (Toth, Daniels, & Solinger, 2011) shows that judgments may be biased by illusions of knowing stemming from feelings of familiarity. Having learners separate diagnostic from non-diagnostic information might boost accuracy. For instance, asking participants how confident they are to answer knowledge questions about a text correctly might lead them to use non-diagnostic information for their confidence rating and as consequences such ratings would not be accurate. In contrast, if participants first had to estimate the text difficulty of a given text and afterwards make a JOL they would attend to different information sources and thereby use more diagnostic and valid information for their JOL (cf. Ozuru et al., 2012). Keeping in mind that even a difficult text could be learnt and not every easy text is understood completely might lead to more accurate JOL.

A further guiding question is: *What impact does text difficulty have on judgment accuracy?* On the basis of Linderholm and colleagues (2012) who found effects of text order and text difficulty it seems necessary to further investigate the impact of text difficulty on JOL and their accuracy. More precisely, are there malleable beliefs concerning text difficulty influenced e.g. by serial JOL that may be used to gauge bias?

Additionally, the following research question is addressed: *What heuristics bias judgment accuracy* (see also Dinsmore & Parkinson, 2013)? Inspired by Serra and Dunlosky's (2010) finding on specific beliefs about the benefits of multimedia formats, a more general research question concerns the influence of heuristics biasing JOL. More specifically, are there domain-independent, and potentially malleable, beliefs that may be used to gauge bias? To foreshadow the multimedia heuristic seems to be promising a candidate.

Finally, the following (minor) research question needs to be answered: *What effects do a priori factors (e.g. self-efficacy, e-learning beliefs) have on judgment accuracy?* To investigate whether e-learning adaptions are possible by using a metacomprehension approach it is important to further analyse prompts as indirect and on-line adaption possibilities. Parameters of direct adaption such as learner control are kept constant and learner variables (e.g. learning-related self-efficacy) which could be used as off-line adaption are controlled for. Main focus is to increase judgment accuracy by exploring a priori indicators and finding an appropriate prompting format.

Answers to all these aforementioned questions will be found in the three conducted studies of this dissertation. All three of them involve an e-learning on team building (cf. Appendix). Each study has different manipulations and a certain main question, both will be described in detail below. Main objective of Study I is to find an appropriate prompting format that boosts judgment accuracy. In Study II the influence of varying text difficulties on judgment accuracy is analysed (cf. Linderholm et al., 2012). Finally, the main focus of Study III is to investigate the impact of the so called "multimedia heuristic" (cf. Serra & Dunlosky, 2010) on judgment accuracy.

In all conducted studies a ten-chapter e-learning on team building is used as learning material. Learners are given metacomprehension prompts at discrete events, such as opening a new chapter and then rate their learning of the just studied chapter. Experimental groups are compared on a variety of learning outcomes tapping judgment accuracy, learning performance (scores on a concluding knowledge test; cf. Appendix), learning efficiency (study time use) and learning-related attitudes (e.g. self-efficacy).

1.4.1 Study I (cf. Chapter 2)

Main objective of Study I is to find an appropriate prompting format that boosts judgment accuracy. The rationale of Study I is that of a manipulation to increase *experience-based* judgments and decrease *theory-based* judgments. With a view towards boosting metacomprehension accuracy in computer-supported learning, the usefulness of combined judgments of learning is assessed. In terms of an experimental manipulation, the research question of Study I therefore is whether judgments of difficulty (JOD) administered before the JOL proper increase JOL accuracy (cf. Appendix).

From these considerations follows a relatively straightforward design that compares in two experiments two groups of learners, respectively. In Experiment 1 of Study I an experimental combined judgments group (JOD and JOL) is compared to a control group of learners who make simple JOL before proceeding to the next of each of ten chapters of a learning unit on team building models. The second experiment of Study I is a replication of Experiment 1, with one crucial difference. Instead of making global judgments, participants make term-specific judgments.

1.4.2 Study II (cf. Chapter 3)

Main objective of Study II is adding knowledge about text difficulty effects on metacomprehension accuracy. Hence, the influence of varying text difficulties as a constraint on judgment accuracy is analysed (cf. Linderholm et al., 2012). Study II extends the first experiment by assessing the role of priming effects from text difficulty. If ease of processing is a criterion, text difficulty might obviously influence ratings and consequently resource allocation (Kornell, Rhodes, Castel, & Tauber, 2011). Thus, sequences of chapters with the same text difficulty and sequences of chapters with randomly varying text difficulties may entail different anchors and thus prime higher vs. lower JOL accuracies, respectively. On the basis of Linderholm and colleagues (2012) who found effects of text order and text difficulty follows a design that compares four groups of learners who perform e-learning differing in blocking type (blocked vs. mixed) and difficulty sequence (rising vs. declining) of text material (cf. Table 1, p.62). Extending the findings of Study I all groups have to deliver combined and specific judgments to assess the text difficulty as well as their learning in Study II (cf. Appendix).

1.4.3 Study III (cf. Chapter 4)

The main focus of Study III is to investigate the impact of the so called “multimedia heuristic” (cf. Serra & Dunlosky, 2010) on judgment accuracy. Recent research showed that the multimedia heuristic has an impact on metacomprehension accuracy (e.g. Ackerman et al., 2013; Butcher, 2006; Paik & Schraw, 2013; Serra & Dunlosky, 2010). Furthermore, images are widely used in computer-based learning although they might bias learners’ judgments on how well they have mastered the material, which might reduce the effectiveness of metacognitive learning control. As that bias seems to result from primarily theory-based processing, combined JOL are used in Study III to induce more experience-based processes and thereby benefit metacomprehension accuracy. The research question of Study III therefore is whether judgments of difficulty (JOD) administered before the JOL proper increase JOL accuracy and reduce the multimedia bias by promoting experience-based processing.

From these considerations follows a design that compares an experimental conceptual images group to an experimental decorative images group and a text-only group. Depending on the condition either conceptual images summarizing the chapters’ content or decorative images without any relevant information were added to the e-learning. The control group worked on a text-only version of the same e-learning unit.

The *Appendix* includes figures of the e-learning, JOL, JOD, knowledge test question as well as different chapter versions to get a better impression of the conducted studies.

1.5 References

Ackerman, R., Leiser, D., & Shpigelman, M. (2013). Is comprehension of problem solutions resistant to misleading heuristic cues? Acta Psychologica, 143(1), 105-112.

Aleven, V., Roll, I., McLaren, B. M., & Koedinger, K. R. (2010). Automated, unobtrusive, action-by-action assessment of self-regulation during learning with an intelligent tutoring system. Educational Psychologist, 45(4), 224-233.

Bransford, J. D., Franks, J. J., Morris, C. D., & Stein, B. S. (1979). Some general constraints on learning and memory research. In L. S. Cermak, & F. I. M. Craik (Eds.), Levels of processing and human memory (pp. 331-354). Hillsdale; NJ: Erlbaum.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. Journal of Educational Psychology, 98(1), 182-197.

Chen, S. Y., Fan, J., & Macredie, R. D. (2006). Navigation in hypermedia learning systems: Experts vs. novices. Computers in Human Behavior, 22(2), 251-266.

Conlan, O., Hockemeyer, C., Wade, V., & Albert, D. (2002). Metadata driven approaches to facilitate adaptivity in personalized eLearning systems. The Journal of Information and Systems in Education, 1, 38-44.

Corbalan, G., Kester, L., & Van Merriënboer, J. J. G. (2006). Towards a personalized task selection model with shared instructional control. Instructional Science, 34(5), 399-422. doi:10.1007/s11251-005-5774-2

Dinsmore, D. L., & Parkinson, M. M. (2013). What are confidence judgments made of? students' explanations for their confidence ratings and what that means for calibration. Learning and Instruction, 24, 4-14. doi:10.1016/j.learninstruc.2012.06.001

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. Current Directions in Psychological Science, 16(4), 228-232.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students’ learning and retention. Learning and Instruction, 22(4), 271-280. doi:10.1016/j.learninstruc.2011.08.003

Dunlosky, J., Rawson, K. A., & Middleton, E. L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. Journal of Memory and Language, 52(4), 551-565.

Fiorella, L., Vogel-Walcutt, J., & Fiore, S. (2012). Differential impact of two types of metacognitive prompting provided during simulation-based training. Computers in Human Behavior, 28(2), 696-702. doi:10.1016/j.chb.2011.11.017

Harteis, C., Gruber, H., & Hertramph, H. (2010). How epistemic beliefs influence e-learning in daily work-life. Educational Technology & Society, 13(3), 201-211.

Ikeda, K., Kitagami, S., Takahashi, T., Hattori, Y., & Ito, Y. (2013). Neuroscientific information bias in metacomprehension: The effect of brain images on metacomprehension judgment of neuroscience research. Psychonomic Bulletin & Review, 20(6), 1357-1363.

Kabassi, K., & Virvou, M. (2006). Multi-attribute utility theory and adaptive techniques for intelligent web-based educational software. Instructional Science, 34(2), 131-158.

Kalyuga, S. (2006). Assessment of learners' organised knowledge structures in adaptive learning environments. Applied Cognitive Psychology, 20(3), 333-342.

Kareal, F., & Klema, J. (2006). Adaptivity in e-learning. A.Méndez-Vilas, A.Solano, J.Mesa and JA Mesa: Current Developments in Technology-Assisted Education, 1, 260-264.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. Journal of Experimental Psychology: General, 126(4), 349-370.

Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. Psychological Science, 22(6), 787-794. doi:10.1177/0956797611407929

Linderholm, T., Wang, X., Therriault, D., Zhao, Q., & Jakiel, L. (2012). The accuracy of metacomprehension judgements: The biasing effect of text order. Electronic Journal of Research in Educational Psychology, 10(1), 111-128.

Maki, R. H. (1998). Text prediction over text materials. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), Metacognition in educational theory and practice (pp. 117-144). Hillsdale, NJ: Erlbaum.

Mengelkamp, C., & Bannert, M. (2009). Judgements about knowledge: Searching for factors that influence their validity. Electronic Journal of Research in Educational Psychology, 7(1), 163-190.

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study-time and the "labor-in-vain effect". Journal of Experimental Psychology: Learning, Memory, and Cognition, Vol. 14(4), 676-686. doi:10.1037/0278-7393.14.4.676

Ozuru, Y., Kurby, C. A., & McNamara, D. S. (2012). The effect of metacomprehension judgment task on comprehension monitoring and metacognitive accuracy. Metacognition and Learning, 7(2), 113-131. doi:10.1007/s11409-012-9087-y

Paik, E. S., & Schraw, G. (2013). Learning with animation and illusions of understanding. Journal of Educational Psychology, 105(2), 278-289.

Puntambekar, S., & Stylianou, A. (2005). Designing navigation support in hypertext systems based on navigation patterns. Instructional Science, 33(5-6), 451-481.

Schnotz, W., & Heiß, A. (2009). Semantic scaffolds in hypermedia learning environments. Computers in Human Behavior, 25(2), 371-380. doi:10.1016/j.chb.2008.12.016

Schwonke, R., Hauser, S., Nückles, M., & Renkl, A. (2006). Enhancing computer-supported writing of learning protocols by adaptive prompts. Computers in Human Behavior, 22(1), 77-92.

Serra, M. J., & Dunlosky, J. (2010). Metacomprehension judgements reflect the belief that diagrams improve learning from text. Memory, 18(7), 698-711.

Son, L. K., & Metcalfe, J. (2005). Judgments of learning: Evidence for a two-stage process. Memory & Cognition, 33(6), 1116-1129.

Stoyanov, S., & Kirchner, P. (2004). Expert concept mapping method for defining the characteristics of adaptive e-learning: ALFANET project case. Educational Technology Research and Development, 52(2), 41-54.

Thiede, K. W., Anderson, M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. Journal of Educational Psychology, 95(1), 66.

Todorov, I., Kornell, N., Larsson Sundqvist, M., & Jönsson, F. U. (2013). Phrasing questions in terms of current (not future) knowledge increases preferences for cue-only judgments of learning. Archives of Scientific Psychology, 1(1), 7-13.

Toth, J. P., Daniels, K. A., & Solinger, L. A. (2011). What you know can hurt you: Effects of age and prior knowledge on the accuracy of judgments of learning. Psychology and Aging, 26(4), 919-931. doi:10.1037/a0023379

Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. Educational Psychologist, 43(1), 16-26.

Zhao, Q., & Linderholm, T. (2008). Adult metacomprehension: Judgment processes and accuracy constraints. Educational Psychology Review, 20(2), 191-206.

2 Boosting metacomprehension accuracy in computer-supported learning: The role of judgment task and judgment scope[1]

[1] This chapter was published as:
Vössing, J., & Stamov-Roßnagel, C. (2016). Boosting metacomprehension accuracy in computer-supported learning: The role of judgment task and judgment scope. *Computers in Human Behavior*, *54*, 73-82.
Accessible online: http://dx.doi.org/10.1016/j.chb.2015.07.066

Chapter 3

3 The Impact of Text Difficulty on Metacomprehension Accuracy and Knowledge Test Performance in Text Learning[2]

[2] This chapter is submitted at the Journal of Computer-Assisted Learning (pre-refereeing version).

3.1 Abstract

Metacomprehension as reflected in judgements of one's learning are crucial for self-regulated study, yet their accuracy is often low. We investigated text difficulty as a constraint on metacomprehension accuracy in text learning. 235 participants studied a ten-section expository text and afterwards took a knowledge test. They made judgements of learning after each section. Sections were of high, medium, or low difficulty; we manipulated between participants the order of difficulty levels across sections. In blocked orders, texts in each block (sections 1-4; sections 5-6; sections 7-10) were of the same difficulty level. In mixed orders, difficulty varied throughout the learning unit either from easy to difficult or from difficult to easy. Our general tenet was that orders would trigger different extents of experience-based processing and thus influence metacomprehension accuracy to different degrees. As hypothesised, accuracy was higher for blocked difficulty orders. Late-section judgement magnitude decreased more strongly the blocked groups. At the same time, late-section judgement accuracy was higher in the blocked group. We discuss implications and limitations of the influence of fluctuations in text difficulty on JOL accuracy together with some avenues for further research.

3.2 Introduction

During self-regulated study, learners are assumed to continuously monitor their learning and base on that monitoring decisions that control learning (e.g., which part of the material to study next; Nelson & Narens, 1990; Nelson & Leonesio, 1988). To study such processes of monitoring and control, researchers may elicit *judgements of learning* (JOL) by having learners gauge the likelihood they will correctly answer a test question on material they previously studied. The accuracy of JOL is then measured by computing the correlation between JOL and performance on the respective test questions, reflecting *metacomprehension,* i.e. one's ability to judge one's learning and/or comprehension of text materials (Dunlosky & Lipko, 2007). As a vital ingredient of self-regulated study, metacomprehension is functional in the control of study and underlies important study decisions, such as time allocation (Metcalfe & Finn, 2008; Metcalfe, 2009; Winne & Hadwin, 1998).

Reflecting the importance of metacomprehension, the first real-world applications are becoming available that use a JOL-based metacomprehension approach to support learners. Hong, Hwang, Tai, and Chen (2014), for instance, had users of a vocabulary learning Smartphone application indicate if they knew the meaning of words presented for study before letting users enter that meaning. This allowed for computing learners' levels of learning confidence, which were then used to select the words to re-study. Even a commercial version of that approach is on the market (e.g., Membean, 2015). However, whilst numerous studies have shown that JOL predict actual learning performance, judgement accuracy tends to be low (see Dunlosky & Lipko, 2007; Lin & Zabrucky, 1998; Maki, 1998; Thiede, Griffin, Wiley, & Redford, 2009), which may limit the effectiveness of such metacomprehension-based learning tools. Moreover, the constraints on metacomprehension accuracy are not yet fully understood (see

Alexander, 2013). Therefore, both from a theoretical and a practical perspective, more research into metacomprehension accuracy is clearly needed. A profound understanding of the drivers of metacomprehension accuracy is crucial for the development of approaches to improve metacomprehension and thus support learning control.

To contribute to such an understanding, we experimentally studied the influence of fluctuations in text difficulty on JOL accuracy. Earlier research demonstrated systematic relationships between text difficulty and JOL accuracy (Linderholm, Wang, Therriault, Zhao, & Jakiel, 2012; Moore, Lin-Agler, & Zabrucky, 2005; Weaver & Bryant, 1995). Across a series of texts, variations in difficulty levels between texts may even impair readers' metacomprehension. Linderholm et al. (2012) found a biasing effect of text order such that metacomprehension accuracy was lower when a set of texts that differed in difficulty were presented in an easy-to-difficulty order, than when difficult texts were presented first. This seems to conflict with the rationale of many expository texts to start easy in order to facilitate learning. More research is therefore needed to understand how difficulty order might influence metacomprehension accuracy.

In the present experiment, participants studied a ten-section expository text and took a knowledge test on that text. They made judgements of learning after each section. Sections were of either high, medium, or low difficulty; we manipulated between participants the order of difficulty levels across sections. Our general tenet was that different orders would provide different meta-informational cues and thus differentially influence metacognitive accuracy. In the following section, we outline a framework for explaining text difficulty effects before reporting an experiment on systematic variations of difficulty order. In the final section, we discuss implications and limitations of our approach together with some avenues for further research.

3.2.1 Text Difficulty as a Meta-Informational Cue

Demonstrating the general influence of text difficulty on metacomprehension, Weaver and Bryant (1995) found higher correlations between predicted and actual performance for passages of medium difficulty (.69) than for either easy or difficult passages (.29 and .30, respectively). As text difficulty was manipulated between participants, however, these findings do not reveal how text difficulty might affect a series of JOL made on longer texts. Moore, Lin-Agler, and Zabrucky (2005) focused on that within-person level by having their participants read 12 expository text passages, four of which were of low, medium, and high difficulty, respectively. After each passage, participants rated their confidence of answering correctly a knowledge question on that passage. One main finding was that confidence was highest for easy passages; no difference emerged between medium and difficult passages. At the same time, JOL on a given passage were more influenced by JOL on prior passages than by an assessment of the given passage. For example, JOL on passage 10 accounted for 43% and 38%, respectively, of the JOL variances at passages 11 and 12.

Linderholm, Wang, Therriault, Zhao, and Jakiel (2012) found what they called a *biasing effect* of text order. Participants studied seven expository texts with varying levels of difficulty. Half the participants were given the texts in an order from easy to difficult, whereas the second group received the difficult-to-easy order. Both groups predicted after each text the number of answers correct on questions about that text. Judgement accuracy was significantly lower in the easy-to-difficult group. Taken together, these studies show that text difficulty affects metacomprehension. Most importantly, Linderholm et al.'s (2012) data suggest that across a series of texts, the order in which difficulty levels vary (easy to difficult vs. difficult to easy) might bias judgements of learning. In real-world learning settings where difficulty is likely to fluctuate across a set of texts, this would mean that the way text sections or individual texts are arranged might adversely, albeit unintendedly, impact on metacomprehension.

To understand the basis of such bias, it is helpful to draw on Koriat's (2007; Koriat, 1997; see also Koriat, Nussinson, Bless, & Shaked, 2008) cue utilisation framework. Its essential assumption is that metacomprehension judgements on a given text are not based on direct access to what is stored about that text in memory. Instead, learners infer how well they will do from a number of plausible cues. Such cues may be experience-based, arising from the on-line monitoring of the material. Beyond the understanding of the text content proper, experience-based cues may be *meta-informational* (Leyens, Yzerbyt, & Schadron, 1994; Yzerbyt, Schadron, Leyens, & Rocher, 1994). Rather than content information, meta-informational cues employ information on stimulus structure or cognitive processing information, such as, for instance, the time needed to retrieve relevant information from memory (Kelley & Lindsay, 1993; Nelson & Narens, 1990), or the perceived completeness of recall (Brewer, Sampaio, & Barlow, 2005), In Linderholm et al.'s (2012) study, text difficulty may have served as such a meta-informational cue. The authors argued that the easy-to-difficult order had triggered shallower processing relative to the difficult-to-easy order, which lead to less accurate judgements (Linderholm et al., 2012). Consistent with this interpretation, the biasing effect was removed in a second experiment by an instruction that sought to increase learners' cognitive engagement with the task. Participants were told they would have to repeat the experiment if they failed to reach a certain performance level (Linderholm et al., 2012).

Learners may also use theory-based cues and judge a specific text on the basis of a general theory or belief. To illustrate, Serra and Dunlosky (2010) found that the "multimedia heuristic", i.e. the belief that texts accompanied by pictures lead to better learning than text-only presentations, went with higher judgements of learning even when only non-conceptual pictures were shown that contained no text-relevant information. Theory-based cues may also include self-referential beliefs, (e.g., about one's self-efficacy, ability, or expertise; Ehrlinger & Dunning, 2003; Glenberg & Epstein, 1987). The latter seem to have played a role in Moore et al.'s (2005) study. To account for the high correlation between participants' individual judgements of learning, those authors argued that learners' assessments of their performance were primarily a reflection of their self-perceptions of ability based on past performance in a domain.

An important question arises when viewing Linderholm et al.'s (2012) and Moore et al.'s (2005) studies in synopsis. From the cue utilisation perspective, it appears that the learners in those studies relied on different cues for their judgements. Whilst texts were ordered along their difficulty in Linderholm et al.'s (2012) study, they were presented in a randomised fashion in Moore et al. (2005). Therefore, participants in that latter study were unable to use text difficulty order as an experience-based meta-informational cue, which may have increased reliance on the theory-based cue of self-ability perceptions. However, in neither study were data on self-perceptions collected, so this interpretation cannot be empirically evaluated. Therefore, we gathered participants' ratings of their learning-related self-efficacy.

3.2.2 Overview of the Study

The present experiment sought to add to the knowledge about text difficulty effects on metacomprehension accuracy by modifying previous research in two respects. First, we compared blocked and mixed orders of text difficulty levels. In blocked orders, each block of four (sections 1-4; sections 7-10) and two sections (sections 5-6), respectively, of the study text were of the same difficulty level. In mixed orders, subsequent sections were always of different levels of difficulty (see Table 1). If blocked orders induce more experience-based processing and provide a more diagnostic meta-informational cue than mixed orders, then we would expect

Hypothesis 1: JOL accuracy will be higher with blocked orders than with mixed orders.

By the same logic, if blocked-order learners engage more experience-based processing, we would expect between-group differences in the adjustment of JOL magnitudes and accuracies from initial to final text sections. Serra and Dunlosky (2010) found such adjustment effects when they compared participants' judgements before, during, and after studying the six sections of an expository text. Participants' pre-study judgements were significantly higher in magnitude than both the averaged judgements after reading a section and the post-study judgements, which may reflect "participants' realisation that the task will be more difficult than they had expected" (p. 704). By analogy, assuming that blocked-order JOL and mixed-order JOL differ in their degrees of experienced-based processing and thus adjustment, we posit

Hypothesis 2: The difference in magnitude between early JOL (sections 1-3) and late JOL (sections 8-10) will be greater in the blocked than in the mixed groups.

Conversely, we expect

Hypothesis 3: The difference in accuracy between early JOL (sections 1-3) and late JOL (sections 8-10) will be greater in the mixed groups than in the blocked groups.

Furthermore, if judgements of learning do not merely reflect people's perception of their learning, but actually control learning, we would expect judgements in blocked groups to be coupled with different learning outcomes than in mixed-order groups. If blocked orders are associated with higher metacomprehension accuracy, then that accuracy might boost learning such that

Hypothesis 4: Blocked-order judgements are associated with higher knowledge test scores than mixed-order judgements.

As a third modification, to gauge the effect of theory-based processing on the grounds of one's ability self-perceptions, we assessed learning-related self-efficacy, defined as learners' perceptions of their competence and confidence to attain high performance levels in learning tasks (see Pintrich & De Groot, 1990). To the extent that mixed orders invoke more theory-based processing and thus greater reliance on ability self-perceptions, we expected that

Hypothesis 5: The correlation between learning-related self-efficacy and judgement magnitude will be higher for mixed-order JOL than for blocked-order JOL.

3.3 Pilot study

We ran a pilot study to generate for each of the ten sections of our study text an easy, medium, and difficulty version, respectively.

3.3.1 Participants

Fifty-six psychology students (none of which participated in the main studies) from the University of Münster participated for course credit. All participants were native German speakers.

3.3.2 Materials and procedure

Similar to the main experiment, participants studied a text on team building. The text comprised ten sections describing the team models by Tuckman and Jensen (1977), Gersick (1988), and Belbin (1993). We created an easy, medium, and difficult version of each section by manipulating the average length of words and sentences, which have been found to be the best indicators of reading difficulty (McLaughlin, 1969; Wasow, 1997). Easy versions had bullet points, short words and sentences and avoided the passive voice, subordinate clauses, connective words and Anglicisms. Technical terms were replaced by colloquial language words. For the medium and difficult versions, we successively increased word and sentence lengths, as well as sentence complexity. Easy sections had Flesh Index values (Flesch, 1948) greater than 60, medium-difficulty sections had Flesh values between 50 and 60, whilst difficult sections' values were below 50.

Participants were randomly assigned to one of five groups. All groups read sections of all difficulty levels, but differed in the order in which difficulty levels were arranged. For instance, in the EMD group (n = 16), each section triplet followed the easy-medium-difficult order. We used another four groups in addition to the EMD group: MDE (n = 8), DME (n = 14), EDM (n = 5), and MED (n = 13). At the end of each section, participants rated the difficulty of that section on a scale from 1 (= very easy) to 5 (= very difficult).

3.3.3 Results

In accordance with the a priori computed Flesh values, the average difficulty ratings were 1.78 (SD = 0.63), 2.26 (SD = 0.69), and 3.26 (SD = 1.05), for easy, medium, and difficult sections, respectively. The average difficulty ratings differed significantly ($F(4, 56) = 2.88$, $p = .03$, $d = .58$), yielding the learning materials for the main study.

3.4 Method

3.4.1 Participants

We recruited on-line through e-mail lists of student representatives 249 students from ten German universities. All participants were native German speakers. Their majors included Social Sciences (48.7%), Natural Sciences (40.8%) and Teacher Education (10.4%). We excluded the data from 14 participants due to excessively short or long study times. Of the remaining 235 participants, 157 (66.8%) were female, the mean age was 23.68 years (*SD* = 3.63). Participants completed the study on-line and were entered into a raffle of vouchers in the value of € 10 of an international e-commerce company.

3.4.2 Materials and procedure

Participants were randomly assigned to one of four experimental groups. In two blocked groups, texts in each block of four (sections 1-4; 7-10) and two sections (sections 5-6), respectively, were of the same difficulty level. Amongst the blocked groups, blocks were presented in easy-medium-difficult order in the bEMD group (n = 64); the bDME group (n = 58) started with difficult blocks. In two mixed-order groups, difficulty varied between each section in the sets of sections 1-4, 5-6, and 7-10, respectively. In the mEMD group (n = 56), sections were presented in the easy-medium-difficult order, whereas participants in the mDME group (n = 57) saw the reverse order (difficult-medium-easy). Table 1 shows the study design including all groups and difficulty levels. There were no significant differences between the groups in age (F(3, 231) = 0.49, p = .69, d = .35), sex (χ^2(3, N = 235) = 0.66, p = .88), or major (χ^2(110, N = 235) = 102.14, p = .63).

Participants studied the team models text described in the pilot study section. Its ten sections were of comparable length ($M_{No\ of\ words}$ = 156.73, *SD* = 30.27) but differed in text difficulty as assessed in the pilot study. Participants studied the text in one session of approx. 30 min.; they could take as much time as they wanted but were not allowed to go back to previous pages. Participants were informed that they would take a knowledge test on the text in the second part of the study. At the end of each section, participants made judgements of learning. First, participants rated on a seven-point Likert scale from 1 = very easy to 7 = very difficult the difficulty of the section (judgement of difficulty; JOD) and second, indicated on a seven-point Likert scale from 1 = very unsure to 7 = very sure how confident they were to be able to answer correctly a question on a specific concept from the section they had just read (judgement of learning; JOL). For instance, participants rated the likelihood of answering correctly "a question on team types in hospitals".

Upon finishing the text, participants took a knowledge test. It comprised 10 multiple-choice questions, each with one correct answer and with four answer options per question, and one question per section. Finally, we collected from all participants together with their demographic data ratings of their learning-related self-efficacy with Pintrich and de Groot's (1990) nine-item scale (Cronbach's α = .88) from the Motivated Strategies for Learning Questionnaire. A sample item is "I am certain I can understand the ideas taught in this course". Responses were rated on a seven-point Likert-scale from 1 = I absolutely agree to 7 = I absolutely disagree. Table 2 shows the correlations between all scores for each condition separately.

3.5 Results

Tests of normal distribution revealed that the dependent variables were skewed. We therefore ran non-parametric tests in addition to parametric tests. As results were identical with both procedures, we report below the more common parametric results. Means and standard deviations for all variables and conditions are summarised in Table 3.

3.5.1 Judgement accuracy

To assess how accurately participants predicted their learning, we computed Goodman-Kruskal Gamma correlations between JOL and the answers on the respective knowledge questions. Gamma is computed by examining the direction of one variable relative to another. Thus, if one judgment of learning increases from one chapter to another and test performance also increases, this is considered a concordance (C). By contrast, if judgments increase whilst performance decreases, this is considered a discordance (D). Concordance and discordance are computed across all pairs of items. The total number of each is used to compute the correlation coefficient $\gamma = (C - D)/(C + D)$. As a measure of relative accuracy, Gamma assesses judgment accuracy independent of test performance and has been used in most JOL studies (see Mengelkamp & Bannert, 2009).

Group mean accuracy was computed (Grand Mean $M = .27$, $SD = .36$, range = -1 - +1). A 2(blocking type: blocked vs. mixed) x 2(difficulty sequence: EMD vs. DME) ANOVA across groups yielded a significant main effect of blocking type, $F(1, 234) = 8.04$, $p = .005$, a non-significant main effect of difficulty sequence, $F(1, 234) = 0.96$, $p = .33$, and a non-significant interaction $F(1, 234) = 0.96$, $p = .32$ of these factors, revealing group differences in accuracy ($F(3, 231) = 3.26$, $p = .022$). JOL accuracy was higher for blocked difficulty orders (*M* = .34, *SD* = 0.33) than for mixed difficulty orders (*M* = .29, *SD* = 0.34). Hypothesis 1 was therefore confirmed.

3.5.2 Judgement magnitude decrease

We calculated the difference between the mean judgement magnitudes on sections 1-3 ("early JOL") and on sections 8-10 ("late" JOL). Testing Hypothesis 2, we found a significant main effect of blocking type, $F(1, 234) = 6.61$, $p = .011$ and a non-significant main effect of difficulty sequence, $F(1, 234) = 0.20$, $p = .652$, as well as a non-significant interaction of these factors, $F(1, 234) = 2.81$, $p = .095$, indicating that groups differed in their degree of JOL adjustment, which was significant ($F(3, 231) = 3.31$, $p = .021$) between early and late JOL. Averaging across the blocked and mixed groups, early-section magnitudes were 5.51 (*SD* = 0.95) and 5.17 (*SD* = 1.09), respectively, dropping to 5.13 (*SD* = 1.30) and 5.17 (*SD* = 1.09), respectively, on late sections. In accordance with Hypothesis 2 the decrease was larger in the blocked groups than in mixed groups where judgement magnitude remained constant. Hypothesis 2 was therefore confirmed.

3.5.3 Judgement accuracy decrease

We calculated the difference between the mean judgement accuracies on sections 1-3 and on sections 8-10. Assessing Hypothesis 3, we found no significant main effects of blocking type, $F(1, 234) = 0.32$, $p = .570$ or difficulty sequence, $F(1, 234) = 0.17$, $p = .677$, nor an interaction of these factors, $F(1, 234) = 0.19$, $p = .663$, indicating that groups did not differ in their degree of accuracy adjustment between early and late JOL ($F(3,231) = 0.23$, $p = .878$). Averaging across the blocked and mixed groups, early-section accuracies were .59 (*SD* = .44) and .47 (*SD* = .55), respectively, dropping to .17 (*SD* = .63) and -0.01 (*SD* = .65), respectively, on late sections. In sum, we had to reject Hypothesis 3.

3.5.4 Knowledge test scores

In the knowledge test, participants scored one point for each correct answer. We summed up points to a test score ranging from 0 to 10 points. Group mean test scores were computed (Grand Mean $M = 5.78$, $SD = 1.94$, range = 0-10). Testing Hypothesis 4, we found a non-significant main effect of blocking type, $F(1, 234) = 2.14$, $p = .145$ and a significant main effect of difficulty sequence, $F(1, 234) = 8.65$, $p = .004$, as well as a non-significant interaction of these factors, $F(1, 234) = 0.21$, $p = .886$, revealing group differences in performance ($F(3, 231) = 3.71$, $p = .012$). Performance scores were higher for EMD orders ($M = 6.18$, $SD = 0.17$) than DME orders ($M = 5.46$, $SD = 0.18$). Hypothesis 4 was therefore not confirmed.

3.5.5 Role of ability self-perceptions

In order to assess the impact of ability self-perception on JOL (Hypothesis 5), we compared the correlations of learning-related self-efficacy with judgement magnitude. Correlations were significant in both groups ($r_{\text{mixed}} = .265$, $p > .001$; $r_{\text{blocked}} = .322$, $p > .001$), but the difference between groups was not significant (z=0.47, p =.32). Also, groups did not differ in their level of self-efficacy, $F(3,231) = 0.32$, $p = .809$. Accordingly, we had to reject Hypothesis 5.

3.6 Discussion

Metacomprehension as reflected judgements of one's own learning is a vital ingredient of self-regulated learning, but the accuracy of such judgements is often low. Earlier studies indicated that the difficulty of study texts can be a constraint on accuracy. In one of those studies, accuracy tended to co-vary with text difficulty; the order in which text passages of varying difficulty were presented biased learners towards lower accuracy (Linderholm et al., 2012).

The present study sought to add to the knowledge about text difficulty effects. As our principal manipulation, we compared mixed orders of text difficulty – similar to those used in Moore et al. (2005) – with blocked, linear orders as they had been used in Linderholm et al. (2012). Our general tenet was that blocked orders would provide learners with a more informative meta-informational cue that would trigger more experience-based processing and thus go with higher judgement accuracy via more experience-based processing, and eventually higher learning performance as assessed in a knowledge test. Although only two of our five hypotheses were supported by the data, the findings allow for some conclusions.

As hypothesised, judgement accuracy in terms of the correlation between one's rating to answer correctly a knowledge question and actual answer correctness was significantly higher in the blocked groups. Also, consistent with predictions, JOL magnitude decreased more strongly from early to late sections with blocked orders, suggesting a stronger influence of experience-based processing in that group. Counter to hypothesis 3, the decrease in judgement accuracy did not differ between the blocked and mixed groups, but whereas accuracy dropped to an insignificant level in the mixed group ($t(112) = -0.15$, $p = .89$), marginally significant accuracy was preserved in the blocked group ($t(121) = 2.56$, $p = .06$). In sum, therefore, the findings on the blocked difficulty orders are consistent with the idea that processing was largely experience-based.

As this implies that processing under mixed orders was more theory-based, consistent with Moore et al.'s (2005) conclusion, we had expected that learning-related self-efficacy would be correlated more strongly with judgement magnitude in that group

(Hypothesis 5). Falsifying this hypothesis, however, our data indicate that learners' self-perceptions of their learning capabilities played no major role in that processing. The correlation of learning-related self-efficacy with judgement magnitude was significant in both groups, yet correlations did not differ between these groups. Also, the self-efficacy perceptions accounted for only some 10% of judgement magnitude.

Still, the findings from Hypotheses 1-3 are consistent with the assumption of more theory-based processing in the mixed groups, relative to the blocked groups. After all, although mixed group participants' learning-related self-efficacy was not lower than blocked groups participants', participants in the mixed groups had lower judgement magnitudes in the initial sections and adjusted less in later sections than blocked-group participants. Perhaps, to the extent that the fluctuating difficulty levels in the mixed group provided a non-informative basis for adjustment, the "theory" underlying the "non-adjustment" simply was to maintain judgement magnitude. Given its greater parsimony, this "non-adjustment account" would be even more attractive than the self-perception account Moore et al. (2005) proposed. Further research in this direction is of course needed.

Turning to the limitations of our study, the data on the knowledge test scores and study times clearly show that more research is needed into the factors that drive the adjustment of initial JOL. We did find group differences in learning performance, but these differences ran along the divide of easy-to-difficult (EMD) orders and difficult-to-easy (DME) orders, independent of the block-mixed manipulation. In other words, the sheer order of text difficulties from easy to difficult made a difference, regardless of whether texts appeared in mixed or blocked fashion. Our data do not allow for a straightforward explanation of that effect.

We believe that taking into account the role of text relatedness might be a fruitful direction for further research. Moore and colleagues (2005) and Linderholm et al. (2012) had used texts without any content relationship to one another. For instance, in the latter study, texts were about topics as diverse as “Literature in the classroom”, “Obesity” or “Intelligence measurement”. In real-world learning, the text passages that learners study will most likely be sections or chapters sharing an overarching topic. In other words, there will be topical relatedness between texts, as there was between the sections of our study text.

Research on the learning of word lists has shown that pair relatedness is a cue that has got a major influence on JOL (for a review, see Mueller, Tauber, & Dunlosky, 2013). For example, Hertzog, Kidder, Powell-Moman, and Dunlosky (2002) had participants study word pairs that were either related (pasture – cow) or unrelated (salt – mayor). Immediately after studying each pair, participants made a JOL. JOL were substantially higher for related than for unrelated pairs. Important in the present context, relatedness effects also emerge when participants study individual words and make a JOL after each word. For example, when Matvey, Dunlosky and Schwartz (2006) presented sets of four words from the same category (e.g., apple, pear, orange, grape), JOL were higher than for sets of unrelated words (e.g., corn, football, perch, lily). Those effects result from the ease or fluency of processing related stimuli. Mueller et al. (2013) argued that in contrast to reading two unrelated words, reading the first word in a related pair may facilitate reading the second word, and such conceptual fluency, in turn, could lead to higher JOL for related than unrelated pairs. This logic also holds for the word list effect in Matvey et al.’s (2006) study. Further research might evaluate the possibility that text relatedness influenced the difficulty order effects. After all, the processing fluency that comes with relatedness is a “fundamental heuristic in metacognition that guides, and biases, judgements about memory" (Kornell, Rhodes, Castel, & Tauber, 2011, pp. 792-793). Relatedness might have overridden the difficulty order effects, which would be consistent with the fact that in our study, no difference emerged between the bEMD and bDME orders, as might have been expected in light of Linderholm et al's (2012) findings.

Also, relatedness is positively linked with building during reading a structured text base (e.g., Kintsch, 1994) that facilitates the processing of incoming sentences. Sanford and Garrod (1994) pointed out that beyond being an objective structural feature of texts, coherence can arise from the way learners process texts. Consistent with this idea, van den Broek, Lorch, Linderholm, and Gustafson (2001) showed readers' active coherence building in generating explanatory and predictive inferences in think-aloud protocols under instructions to study for a subsequent test, but not when simply reading "for pleasure". The DME effect in the knowledge test data would be compatible with the assumption that starting with easier texts facilitated that coherence building, which was in turn positively related to test performance. At the same time, higher coherence is not necessarily linked to higher judgement accuracy. In Linderholm et al.'s (2012) study, the EMD group showed lower judgement accuracy than the DME group, presumably because their perception that texts were easy lowered the cognitive resources they devoted to monitoring their comprehension, which compromised judgement accuracy. This would explain the dissociation between judgement accuracy and learning performance as a function of group membership; the blocked groups had higher accuracy, but lower performance, whereas the EMD groups had higher performance, but lower accuracy. This reasoning is consistent with Shaddock and Carroll's (1997) findings on the effect of coherent sequences of sentences on metamemory judgements. Participants were able to recall more of the target sentences presented in an ordered sequence, although they could not predict this when making JOL beforehand. Also, using texts with varying degrees of coherence, Rawson and Dunlosky (2002) found that the magnitude of metacomprehension judgements was significantly higher for more coherent texts (for similar findings, see Dunlosky, Baker, Rawson, & Hertzog, 2006).

In sum, although not all hypotheses were confirmed, the present study shows systematic effects of text difficulty such that the order in which texts of different difficulty levels are presented affect metacomprehension accuracy. Future research could fruitfully explore the factors that drive the adjustment of learning judgements, which in turn seems to enhance accuracy. Text relatedness might play an important role in that regard.

3.7 References

Alexander, P. A. (2013). Calibration: What is it and shy it matters? An introduction to the special issue on calibrating calibration. *Learning and Instruction, 24,* 1-3.

Belbin, R. M. (1993). *Team roles at work: A strategy for human resource management*. Oxford: Butterworth-Heinemann.

Brewer, W. F., Sampaio, C., & Barlow, M. R. (2005). Confidence and accuracy in the recall of deceptive and non-deceptive sentences. *Journal of Memory and Language, 52*(4), 618-627.

Dunlosky, J., Baker, J., Rawson, K. A., & Hertzog, C. (2006). Does aging influence people's metacomprehension? Effects of processing ease on judgments of text learning. *Psychology and Aging, 21*(2), 390-400.

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science, 16*(4), 228-232.

Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *Journal of Personality and Social Psychology, 84*(1), 5-17.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221-233.

Gersick, C. J. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal, 31*(1), 9-41.

Glenberg, A. M., & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition, 15*(1), 84-93.

Hertzog, C., Kidder, D. P., Powell-Moman, A., & Dunlosky, J. (2002). Aging and monitoring associative learning: Is monitoring accuracy spared or impaired? *Psychology and Aging, 17*(2), 209-225.

Hong, J., Hwang, M., Tai, K., & Chen, Y. (2014). Using calibration to enhance students' self-confidence in English vocabulary learning relevant to their judgment of over-confidence and predicted by smartphone self-efficacy and English learning anxiety. *Computers & Education, 72*, 313-322.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32*(1), 1-24.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*(4), 294-303.

Koriat, A. (2007). Remembering: Metacognitive monitoring and control processes. In H. L. Roediger, III, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of Memory: Concepts* (pp. 243-246). Oxford: University Press.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*(4), 349-370.

Koriat, A., Nussinson, R., Bless, H., & Shaked, N. (2008). Information-based and experience-based metacognitive judgments: Evidence from subjective confidence. In J. Dunlosky, & R. A. Bjoerk (Eds.), *A Handbook of Memory and Metamemory,* (pp. 117-136). New York Hove: Psychology Press.

Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psychological Science, 22*(6), 787-794. doi:10.1177/0956797611407929

Leyens, J., Yzerbyt, V., & Schadron, G. (1994). *Stereotypes and social cognition.* Sage Publications, Inc.

Lin, L., & Zabrucky, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology, 23*(4), 345-391.

Linderholm, T., Wang, X., Therriault, D., Zhao, Q., & Jakiel, L. (2012). The accuracy of metacomprehension judgements: The biasing effect of text order. *Electronic Journal of Research in Educational Psychology, 10*(1), 111-128.

Maki, R. H. (1998). Text prediction over text materials. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117-144). Hillsdale, NJ: Erlbaum.

Matvey, G., Dunlosky, J., & Schwartz, B. (2006). The effects of categorical relatedness on judgements of learning (JOLs). *Memory, 14*(2), 253-261.

McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading, 12*(8), 639-646.

Membean (2015, November 25th), Incredibly effective vocabulary learning [commercial training]. Retrieved from: http://membean.com/

Mengelkamp, C., & Bannert, M. (2009). Judgements about knowledge: Searching for factors that influence their validity. *Electronic Journal of Research in Educational Psychology, 7*(1), 163-190.

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science, 18*(3), 159-163. doi:10.1111/j.1467-8721.2009.01628.x

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review, 15*(1), 174-179.

Moore, D., Lin-Agler, L. M., & Zabrucky, K. M. (2005). A source of metacomprehension inaccuracy. *Reading Psychology, 26*(3), 251-265.

Mueller, M. L., Tauber, S. K., & Dunlosky, J. (2013). Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review, 20*(2), 378-384.

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study-time and the "labor-in-vain effect". *Journal of Experimental Psychology: Learning, Memory, and Cognition, Vol. 14*(4), 676-686. doi:10.1037/0278-7393.14.4.676

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation, 26*, 125-141.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40.

Rawson, K. A., & Dunlosky, J. (2002). Are performance predictions for text based on ease of processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 69-80.

Sanford, A. J., & Garrod, S. C. (1994). Selective processing in text understanding. In M. A. Gernsbacger (Ed.), *Handbook of psycholinguistics* (pp. 699-719). San Diego, CA: Academic Press.

Serra, M. J., & Dunlosky, J. (2010). Metacomprehension judgements reflect the belief that diagrams improve learning from text. *Memory, 18*(7), 698-711.

Shaddock, A., & Carroll, M. (1997). Influences on metamemory judgements. *Australian Journal of Psychology, 49*(1), 21-27.

Thiede, K. W., Griffin, T. D., Wiley, J., & Redford, J. S. (2009). Metacognitive monitoring during and after reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education,* (pp. 85-106). New York and London: Routledge.

Tuckman, B. W., & Jensen, M. A. C. (1977). Stages of small-group development revisited. *Group & Organization Management, 2*(4), 419-427.

Van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition, 29*(8), 1081-1087.

Wasow, T. (1997). Remarks on grammatical weight. *Language Variation and Change, 9*(01), 81-105.

Weaver, C. A., & Bryant, D. S. (1995). Monitoring of comprehension: The role of text difficulty in metamemory for narrative and expository text. *Memory & Cognition, 23*(1), 12-22.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. *Metacognition in Educational Theory and Practice, 93*, 27-30.

Yzerbyt, V. Y., Schadron, G., Leyens, J., & Rocher, S. (1994). Social judgeability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology, 66*(1), 48-55.

3.8 Tables

Table 1

Study design and overview of the four conditions

Blocked	E	E	E	E	M	M	D	D	D	D	
	D	D	D	D	M	M	E	E	E	E	Knowledge
Mixed	E	M	D	E	M	D	E	M	D	E	Test
	D	M	E	D	M	E	D	M	E	D	

Note. Text difficulty of the ten sections was easy (E), medium (M) or difficult (D).

Table 2

Correlations between all scores (separate for each condition)

	Self-Efficacy	Judgement items (JOL)	Knowledge test items
Self-Efficacy			
bEMD	1	-.370**	-.036
mEMD	1	-.253	.063
bDME	1	-.275**	-.199
mDME	1	-.274*	-.186
Total	1	-.296**	-.099
Judgement items (JOL)			
bEMD	-.370**	1	-.019
mEMD	-.253	1	.319*
bDME	-.275*	1	.398**
mDME	-.274*	1	.113
Total	-.296**	1	.203**
Knowledge test items			
bEMD	-.036	-.019	1
mEMD	.063	.319*	1
bDME	-.199	.398**	1
mDME	-.186	.113	1
Total	-.099	.203**	1

Note. * significant at .05 (two-tailed); ** significant at .001 (two-tailed).

Table 3

Means and standard deviations for all variables and conditions

	bEMD (*n* = 64)		mEMD (*n* = 56)		bDME (*n* = 58)		mDME (*n* = 57)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
JOL	5.34	1.00	5.28	1.01	5.35	1.00	5.16	1.10
Accuracy	0.29	0.38	0.20	0.38	0.38	0.32	0.20	0.36
Knowledge test	6.38	1.68	5.98	1.77	5.62	1.92	5.30	2.11
Self-efficacy	3.46	0.91	3.50	0.95	3.37	0.99	3.53	1.01

Note. JOL 1 = very unsure to 7 = very sure; accuracy range -1 to 1; performance range 0 to 10; self-efficacy 1 = I absolutely agree to 7 = I absolutely disagree.

Chapter 4

4 Images in Computer-Supported Learning: Increasing their Benefits for Metacomprehension through Judgments of Learning[3]

[3] This chapter was published as:
Vössing, J., Stamov-Roßnagel, C., & Heinitz, K. (2016). Images in computer-supported learning: Increasing their benefits for metacomprehension through judgments of learning. *Computers in Human Behavior, 58*, 221-230.
Accessible online: http://dx.doi.org/10.1016/j.chb.2015.12.058

Chapter 5

5 General Discussion

5.1 Review of the main findings

The overall aim of this dissertation project was to investigate how to boost accuracy of metacomprehension prompts (JOL) to an appropriate level. Based on metacomprehension judgments e-learning could then be adapted to personal needs and rise in efficiency. Although JOL usually predict learning performance, correlations might be as low as .27 (Maki, 1998). This JOL magnitude seems to be influenced both by features of the learning material (e.g., amount and size of text) and by personal factors (e.g. self-efficacy, e-learning beliefs).

The rationale of Study I was that of a manipulation to increase experience-based judgments and decrease theory-based judgments. In accordance with Son and Metcalfe (2005) who claimed JOL underlying a two-stage process as well as Ozuru, Kurby and McNamara (2012) who stated different types of JOL alter the information source to which learners attend and thereby influence judgment accuracy the research question of Study I was whether JOD administered before JOL could proper increase JOL-accuracy. Initially all five hypotheses had to be rejected. Neither accuracy nor any other differences emerged between simple and combined groups, but a comparison of early-chapter and late-chapter judgment magnitudes showed that the judgment manipulation had induced cognitive processing differences. This corresponds to Zhao and Linderholms' (2008) anchoring-and-adjustment account. Study I therefore includes a second experiment, where judgment scope was manipulated. Dunlosky, Rawson and Middleton (2005) noted a mismatch between judgement scope and tested information. If the global scope had caused the combined-manipulation to fail, significantly higher judgment accuracy was expected as a function of term-specific judgments. Both judgment accuracy and knowledge test scores were as hypothesized higher in the combined judgments group in Experiment 2. Moreover, while judgment accuracy dropped to an insignificant level between early and late chapters in the simple judgments group, accuracy remained constant with combined judgments. Higher knowledge scores at short study times indicate that learners in the combined-JOL group learned both more effectively and more efficiently. In sum, therefore, Study I was a step in the right direction.

To add knowledge about text difficulty effects as a constraint on metacomprehension accuracy in text learning Study II compared four groups of learners who conducted an e-learning differing in blocking type (blocked vs. mixed) and difficulty sequence (EMD vs. DME). Based on Linderholm, Wang, Therriault, Zhao and Jakiel's (2012) finding of a biasing effect of text order as well as on Moore, Lin-Agler and Zabrucky (2005) who demonstrated prior JOL to have a huge influence on later JOL the general tenet was that orders would trigger different extents of experience-based processing and thus influence metacomprehension accuracy to different degrees. As hypothesized, accuracy was higher for blocked difficulty orders. Late-section judgment magnitude decreased more strongly in the blocked groups. At the same time, late-section judgment accuracy was higher in the blocked groups. Contrary to expectations, the data indicate that learners' self-perceptions of their learning capabilities played no major role in that processing. Performance differences only emerged between difficulty sequences with EMD outperforming DME groups instead as hypothesized between blocking type. Still, the findings from three out of five hypotheses are consistent with the assumption of more theory-based processing in the mixed groups, relative to the blocked groups.

Study III investigated the impact of the multimedia heuristic (cf. Serra & Dunlosky, 2010; Ikeda, Kitagami, Takahashi, Hattori, & Ito, 2013) on judgment accuracy. The underlying research question was whether JOD administered before JOL would proper increase JOL accuracy and reduce the multimedia bias by promoting experience-based processing. As postulated, combined JOL benefited accuracy and knowledge test scores; both were highest in the conceptual images group. The differences in judgment magnitude and accuracy between early and late chapters suggest that the combined JOL manipulation induced cognitive processing differences, as postulated. The findings on multimedia beliefs run counter the hypothesis; a marginal effect in the opposite direction than postulated was obtained. Although two out of five hypotheses had to be rejected, the findings largely support the general assertion that combined judgments lead to more experience-based processing and thus higher judgment accuracy. In a second experiment of Study III the usefulness of term-specific judgments to further increase judgment accuracy was tested. Again, accuracy and test scores were highest in the conceptual images group. Contrary to expectations, however, JOL accuracy did not benefit from term-specific judgments. What is worse, term-specific judgments may even harm accuracy. Taken together, findings of Study III are consistent with prior findings that images may trigger multimedia heuristics that lead learners to judge their learning higher than is justified by actual learning performance. But Study III also demonstrates that metacognitive judgment tasks may be designed so that they support learning. The combined judgments of learning had learners attend more closely to the specific study texts, which increased both judgment accuracy and knowledge test performance over accuracy and performance with standard simple JOL.

In sum, therefore combined judgments seem to be “better” than simple judgments, text blocks of the same difficulty level “better” than mixed texts and conceptual images “better” than decorative images or text only. As one reviewer noted, we “are definitely on to something”.

5.2 Returning to the Research Questions

5.2.1 What prompting format best boosts judgment accuracy?

The main research question of this dissertation was: *What prompting format best boosts judgment accuracy?* What have we learnt considering the conducted studies and experiments about judgment accuracy? First of all, it was possible to increase judgment accuracy throughout the studies notably. Thanks to two-step judgments and term-specific JOL judgment accuracy reached a level of 0.32 in Study I. In light of the fact that Maki (1998) found 0.27 as lowest level for judgment accuracy, 0.32 is just a slight improvement. In Study II text difficulty effects were considered beside two-step-judgments. For blocked orders of text difficulty the most successful groups reached an accuracy level of 0.38 which is a further gain. Finally, in Study III it was possible with combined judgments and conceptual images to boost judgment accuracy up to 0.42. Overall, there is still some room for improvement left. However, an accuracy level of 0.42 is a satisfying result due to the fact that it was not reached coincidentally but with sophisticated manipulations. To improve judgment accuracy notably combined or two-step judgments (JOD and JOL), judgment scope (specific), text difficulty (blocked) and images (conceptual) should be taken into account. These conclusions provide a framework or "baseline paradigm" in which further effects and impact factors like learner variables, text genres or answer format on metacomprehension accuracy could be studied (cf. 5.5).

5.2.2 What impact does text difficulty have on judgment accuracy?

Study II analysed the impact of text difficulty on judgment accuracy (cf. Linderholm et al., 2012). In general it was found that blocks of the same text difficulty level are superior to mixed text difficulty levels. As participants had to deliver a judgment of difficulty before rating their learning text difficulty is an important impact factor for judgment accuracy. That way it seems helpful if participants are able to recognize the actual text difficulty level which in turn is easier with blocks of the same difficulty level in comparison to mixed difficulty levels. Furthermore, difficulty sequences should have an effect on judgement accuracy. In real-world learning settings difficulty is likely to fluctuate across a set of learning material, i.e. starting with easier texts and becoming more difficult to the end of a session. Results of Study II demonstrate advantages of a rising text difficulty sequence instead of a declining but here more research is needed. What factors drive the adjustment of initial JOL needs further investigation, i.e. text relatedness seems to have an influence (cf. 5.5). For now, data suggests to consider text difficulty in order to increase judgment accuracy. Thereby text difficulty blocks lead to higher accuracy than mixed text difficulty levels.

5.2.3 What heuristics bias judgment accuracy?

Study III investigated whether judgments of difficulty (JOD) administered before the JOL proper increase JOL accuracy and reduce the multimedia bias by promoting experience-based processing (cf. Serra & Dunlosky, 2010; Ikeda et al., 2013). Results are consistent with the finding that images may trigger multimedia heuristics and lead learners to judge their learning higher than justified by actual learning performance. This bias may even nullify the positive impact on learning that multimedia materials can have. At the same time Study III contributes to research by showing that metacognitive judgment tasks may be designed so that they support learning. The combined judgments of learning had learners attend more closely to the specific study texts, which increased both judgment accuracy and knowledge test performance over accuracy and performance with standard simple JOL. Moreover, with mean accuracies of .42 (conceptual images), .25 (decorative images), and .28 (text-only), learners in the first experiment of Study III fared better than the respective groups in two similar studies where accuracies of .34, .11, and .29 (Jaeger & Wiley; 2014), as well as of .12, .26, and .16. (Serra & Dunlosky, 2010) had been reported. Also, consistent with these latter studies, we found the lowest accuracy with decorative images, relative to the other two groups. This strengthens the evidence for the "metacognitive harm" that non-informative images might do and suggests that multimedia content designers should not only pay attention to making illustrations informative. Content designers might also be well-advised to avoid using merely "motivational" decorative images with no substantive relationship to text content. While such images seek to make learning materials more attractive and thus engage learners more, they might actually "backfire" by impairing the metacognitive control of learning. Although clearly more research is needed, e.g.to understand why term-specific JOL did not further benefit judgment accuracy (Study III, Experiment 2), we demonstrated the multimedia heuristic as bias constraining metacomprehension accuracy. Furthermore, results show the advantages of conceptual images to increase judgment accuracy in comparison to decorative images or plain text.

5.2.4 What effects do a priori factors (e.g. self-efficacy, e-learning beliefs) have on judgment accuracy?

Throughout the studies learner characteristics which might have an influence on judgment accuracy were analysed. In Study I and II self-efficacy was measured but contrary to expectations self-efficacy perceptions accounted for only some 10% of judgment magnitude. In detail, we obtained 11.36% in the first experiment of Study I, 10.76% in the second experiment of Study I and 8.76% in Study II. Therefore, in Study III we collected in line with Serra and Dunlosky (2010) data on participants "beliefs about multimedia learning" (BAML) instead of learning-related self-efficacy. Results were even worse with BAML accounting best for 8.23% of judgment magnitude in the combined group with conceptual images. For all other groups BAML accounted for only some 1% of judgment magnitude. Therefore according to the measures we used neither self-efficacy nor e-learning beliefs seem to have a notable influence on judgment accuracy. Further research should use different measures to investigate possible effects of self-efficacy and e-learning beliefs. Additionally, different learner characteristics should be taken into account (cf. 5.5).

5.2.5 Additional Findings

Beside answers to the research questions of this dissertation, results of the studies include additional interesting findings. First, throughout the experiments a JOL magnitude decrease emerged from early to late chapters being more pronounced with simple JOL and demonstrating a robust effect that should be considered in further research on metacomprehension accuracy. Second, in Study II and III an accuracy decrease from early to late chapters being more pronounced with combined JOL was postulated. Although hypotheses could not be confirmed, results support the underlying idea of more experience-based processing leading to less overcorrection in JOL adjustment. Third, as a side-effect in Study I longer study times for simple groups were found promising that eventually a metacomprehension approach could lead to more efficiency in e-learning.

5.3 Scientific and practical relevance

This dissertation project contributes to establishing a platform for further research into metacomprehension processes in computer-supported learning. The combined JOD-JOL ratings together with our findings concerning the impact of text difficulty and the multimedia heuristic predict learning performance significantly and consistently across an entire learning session and might therefore provide a "baseline paradigm" in which the effect of learner variables such as overconfidence or gender effects on metacomprehension accuracy could be studied. Exploring the nature of JOL and developing a deeper understanding of what factors influence them helps to use JOL in an appropriate manner i.e. as learner support. Knowing a prompting format that best boosts learning performance and leads to accurate judgments is the prerequisite for a broad application of a metacomprehension approach in adaptive e-learning. To the extent that metacomprehension bias appears to affect low-proficiency learners more than high-proficiency learners, our findings could serve as starting point for analysing how metacognitive trainings might be designed to help learners become "immune" to bias. Additionally, the findings are not solely limited to computer-supported learning. Although the approaches were applied in an online e-learning module they should work for printed text or traditional paper learning as well. Here too, the findings could serve as a "baseline paradigm" for further studies. As a side effect, the *testing effect* which is well-established for word-pair also occurred with the longer and more complex materials used in the present studies. Some of the potential future research questions are discussed in Section 5.5.

Beside an eminent contribution to research in the area of computer-supported learning the conducted studies will help to finally realise a JOL application in practice. Due to JOL constructing adaptive e-learning will be less costly in the future. At the same time these types of e-learning will be more efficient, boost learning performance and foster sustainable learning motivation which should appeal companies as well as users. In terms of practical relevance, the findings might therefore help to work towards an adaptive learner support that is easy to implement. The idea is that judgments of learning might in principle be used to "personalise" learning. For instance, learners might be asked to make JOL after a chapter or section. In case of low-magnitude JOL (e.g., subjective probability of correct answer below, say, 40%), they might be given that chapter for re-study or might be offered additional elaboration tasks (e.g., keyword generation). For high-magnitude JOL, on the other hand, the system would let learners proceed to the next chapter. Such support might work largely independent of learning content and might therefore be implemented in a standardized fashion. This adaptive learner support could therefore scaffold the learning process using less time-consuming interventions than re-reading (e.g., Dunlosky & Rawson, 2005; Rawson, Dunlosky, & Thiede, 2000) or summarizing (e.g., Anderson & Thiede, 2008; Thiede & Anderson, 2003).

While there is still research needed before implementing adaptive e-learning based on JOL successfully, the present findings also allow some concrete recommendation for practice. First, content designers are well-advised to avoid merely "motivational" decorative images with no substantive relationship to text content. While such images seek to make learning materials more attractive and thus engage learners more, they might actually "backfire" by impairing the metacognitive control of learning. Instead conceptual images should be used, i.e. charts that are informational equivalent to the text, summarizing the chapter content to support comprehension of the key message. Second, as systematic effects of text difficulty were found learning material should be presented in blocked order to evoke experience-based processing and thereby improve metacomprehension accuracy. Third, two-step judgments should be used to further foster experience-based processing in contrast to theory-based processing. Additionally, "cue utilization trainings" might help learners overcome bias by making them aware of which cues might influence their metacognitive judgments and by explicitly attending to those cues in order to “correct” their judgment for inappropriate cue influence.

5.4 Limitations

Of course, this dissertation project has got limitations that need addressing in future research. For instance, resulting from Study I, a more specific scale for learning-related self-efficacy should be applied to analyse relevant learner characteristics and overcome limitations. As a further constraint Study I points out that the effects on adjustment behaviour are not yet completely understood and should therefore be addressed in future studies. Based on the anchoring and adjustment account, we had expected that due to their greater reliance on experience-based processing, combined-JOL learners would adjust their judgments more strongly than simple-JOL participants, but we found the opposite. There might be a systematic interaction between judgment type and adjustment behaviours future studies could systematically search for in order to overcome limitations. This is also a result of Study II were we called for further research finding and understanding factors that drive adjustment of initial JOL. Beside, as another limitation, Study II revealed the need of taking into account the role of text relatedness in future studies. Chances are that text relatedness further influences the difficulty order effects. Study III revealed as constraint that the role of images is not yet understood well enough and needs future addressing. We found a dissociation between accuracy (that was unaffected by combined JOL) and test performance (that benefited from combined JOL) depending on the type of image. Additionally, results of Study III are in contrast to both Jaeger and Wiley (2014) and Serra and Dunlosky (2010). All studies are similar in their general set-up but differ in important details like number and length of study texts, topical coherence and JOL instructions suggesting manipulations to explore in future studies.

Turning to the general limitations that need addressing in future research, all studies ran on-line rather than in the laboratory, which limits experimental control. However, as system-generated time stamps were used to ensure only participants were included who worked on the module in one uninterrupted sitting, no artefacts from that on-line setting were seen that could have biased results. A further limitation concerns the generalisability of the results to "real" computer-supported learning given that only text materials and pictures were used.

5.5 Future research directions

Beside the above mentioned points, future studies should use richer materials as could be found in professional e-learning. In those e-learning, e.g. videos and interactive parts are included. Another issue for further research concerns extending the metacomprehension paradigm beyond recall questions in a multiple-choice format that was used in the present experiments. Real-world learning modules more often than not also include transfer questions, and sometimes even open-ended answer formats. Furthermore, to increase generalisability an application to different text genres and additional knowledge fields seems an interesting extension.

Another important extension should consider the learners themselves. The same e-learning is conducted by people differing e.g. in their attitudes towards e-learning, learning self-efficacy or overconfidence, respectively underconfidence. Analysing and understanding the impact of different learner characteristics on metacomprehension is a further vital requirement to improve adaptive computer-supported learning.

Further experiments also should aim to analyse learners' response to JOL. All in all a metacomprehension approach should be used to adapt e-learning to personal needs and thereby increase efficiency and fosters sustainable learning motivation. Therefore it is important to investigate the consequences that follow JOL, e.g. which type of learner decides to restudy a section and when? It seems further interesting to investigate the impact of JOL on learning motivation.

Overall the metacomprehension approach should be applied and investigate in practice. For that purpose, the findings should be transferred to real-world e-learning with different topics and applied with learners, e.g. from different companies.

5.6 Conclusion

Taken together this dissertation presents work contributing to a further improvement of e-learning using a metacomprehension approach. To boost judgment accuracy the advantage of two-step judgments in comparison to simple judgments was demonstrated. According to the reported findings term-specific judgments are superior to the more common global judgments at least as long as only text material is used. Judgment scope could therefore further improve judgment accuracy in computer-supported learning. As shown in Study II text difficulty does also have an impact on metacomprehension accuracy. Varying difficulty levels can either constraint or facilitate computer-supported learning. Here, blocked sequences of the same difficulty level instead of mixed sequences are recommended. Additionally, an e-learning starting with easier sections and a rising text difficulty could further benefit learning. Concerning the impact of images on metacomprehension accuracy more research is needed. For now, Study III demonstrated that conceptual images improve judgment accuracy as well as learning performance. In contrast, decorative images might foster multimedia heuristics and thereby even nullify the positive impact on learning that multimedia materials can have. It seems therefore highly advisable to solely use conceptual images in computer-supported learning.

Ultimately, this dissertation project presents important findings contributing to a further improvement of e-learning. The understanding of metacomprehension in self-regulated learning and its impact on accuracy are the obligatory requirements for adapting computer-supported learning to individual needs and also raises its efficiency. Becoming cheaper, more efficient and more effective will finally enable e-learning to meet the challenge of lifelong learning.

5.7 References

Anderson, M., & Thiede, K. W. (2008). Why do delayed summaries improve metacomprehension accuracy? *Acta Psychologica, 128*(1), 110-118.

Dunlosky, J., & Rawson, K. A. (2005). Why does rereading improve metacomprehension accuracy? Evaluating the levels-of-disruption hypothesis for the rereading effect. *Discourse Processes, 40*(1), 37-55.

Dunlosky, J., Rawson, K. A., & Middleton, E. L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. *Journal of Memory and Language, 52*(4), 551-565.

Ikeda, K., Kitagami, S., Takahashi, T., Hattori, Y., & Ito, Y. (2013). Neuroscientific information bias in metacomprehension: The effect of brain images on metacomprehension judgment of neuroscience research. *Psychonomic Bulletin & Review, 20*(6), 1357-1363.

Jaeger, A. J., & Wiley, J. (2014). Do illustrations help or harm metacomprehension accuracy? *Learning and Instruction, 34*, 58-73.

Linderholm, T., Wang, X., Therriault, D., Zhao, Q., & Jakiel, L. (2012). The accuracy of metacomprehension judgements: The biasing effect of text order. *Electronic Journal of Research in Educational Psychology, 10*(1), 111-128.

Maki, R. H. (1998). Text prediction over text materials. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117-144). Hillsdale, NJ: Erlbaum.

Moore, D., Lin-Agler, L. M., & Zabrucky, K. M. (2005). A source of metacomprehension inaccuracy. *Reading Psychology, 26*(3), 251-265.

Ozuru, Y., Kurby, C. A., & McNamara, D. S. (2012). The effect of metacomprehension judgment task on comprehension monitoring and metacognitive accuracy. *Metacognition and Learning, 7*(2), 113-131. doi:10.1007/s11409-012-9087-y

Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition, 28*(6), 1004-1010.

Serra, M. J., & Dunlosky, J. (2010). Metacomprehension judgements reflect the belief that diagrams improve learning from text. *Memory, 18*(7), 698-711.

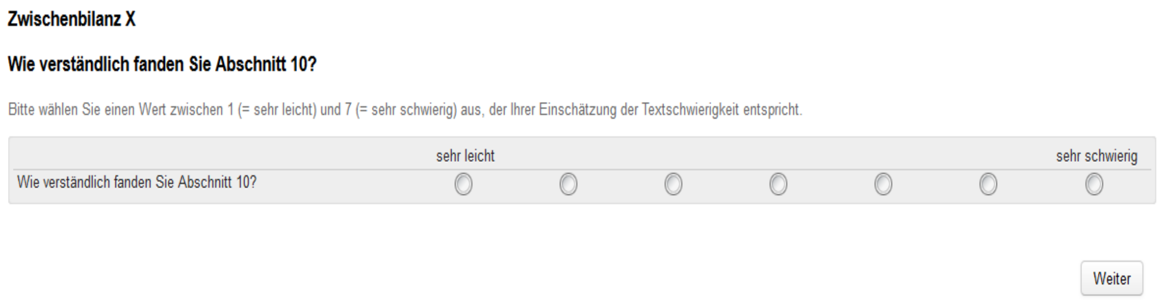
Son, L. K., & Metcalfe, J. (2005). Judgments of learning: Evidence for a two-stage process. *Memory & Cognition, 33*(6), 1116-1129.

Thiede, K. W., & Anderson, M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology, 28*(2), 129-160.

Zhao, Q., & Linderholm, T. (2008). Adult metacomprehension: Judgment processes and accuracy constraints. *Educational Psychology Review, 20*(2), 191-206.

Appendix

Figure 1. Judgment of Difficulty (JOD)



Zwischenbilanz X

Wie verständlich fanden Sie Abschnitt 10?

Bitte wählen Sie einen Wert zwischen 1 (= sehr leicht) und 7 (= sehr schwierig) aus, der Ihrer Einschätzung der Textschwierigkeit entspricht.

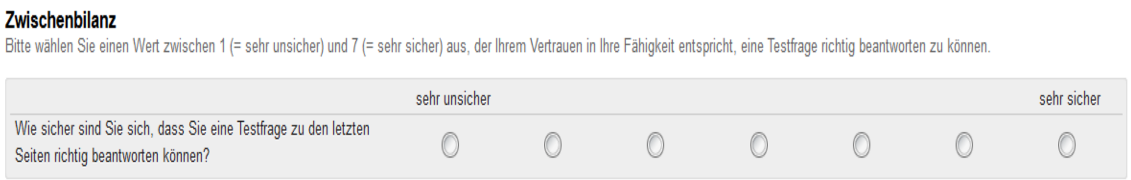
	sehr leicht						sehr schwierig
Wie verständlich fanden Sie Abschnitt 10?	○	○	○	○	○	○	○

Weiter

Figure 1. Judgment of difficulty corresponding to chapter 10. JOD were used at least in one experimental group of each study.

Figure 2. Comparison of global and specific JOL

Global JOL

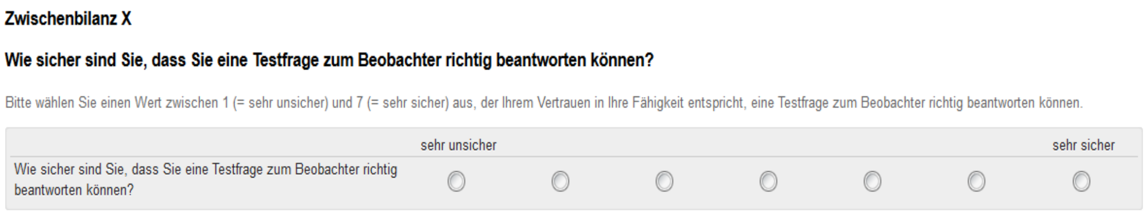


Zwischenbilanz

Bitte wählen Sie einen Wert zwischen 1 (= sehr unsicher) und 7 (= sehr sicher) aus, der Ihrem Vertrauen in Ihre Fähigkeit entspricht, eine Testfrage richtig beantworten zu können.

	sehr unsicher						sehr sicher
Wie sicher sind Sie sich, dass Sie eine Testfrage zu den letzten Seiten richtig beantworten können?	○	○	○	○	○	○	○

Specific JOL



Zwischenbilanz X

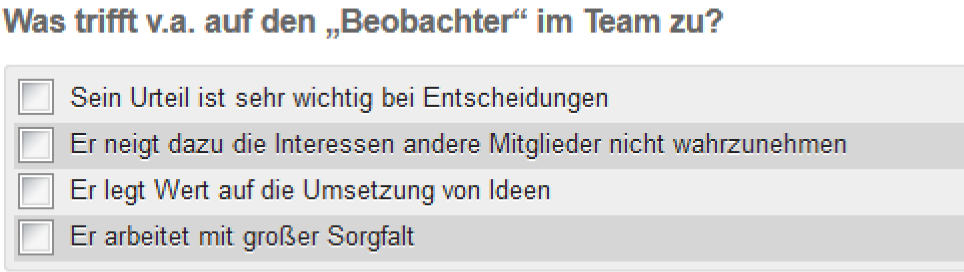
Wie sicher sind Sie, dass Sie eine Testfrage zum Beobachter richtig beantworten können?

Bitte wählen Sie einen Wert zwischen 1 (= sehr unsicher) und 7 (= sehr sicher) aus, der Ihrem Vertrauen in Ihre Fähigkeit entspricht, eine Testfrage zum Beobachter richtig beantworten können.

	sehr unsicher						sehr sicher
Wie sicher sind Sie, dass Sie eine Testfrage zum Beobachter richtig beantworten können?	○	○	○	○	○	○	○

Figure 2. Judgment of learning corresponding to chapter 10. The global JOL were used in Experiment 1 of Study I and Experiment 1 of Study III. The specific JOL were used in Experiment 2 of Study I and III as well as in Study II.

Figure 3. Knowledge test question 10



Was trifft v.a. auf den „Beobachter“ im Team zu?

- ☐ Sein Urteil ist sehr wichtig bei Entscheidungen
- ☐ Er neigt dazu die Interessen andere Mitglieder nicht wahrzunehmen
- ☐ Er legt Wert auf die Umsetzung von Ideen
- ☐ Er arbeitet mit großer Sorgfalt

Figure 3. Knowledge test question corresponding to chapter 10. The test comprised 10 multiple-choice questions, each with one correct answer and four answer options per question, and one question per chapter. For all conducted studies and experiments the same knowledge test was used.

Figure 4. E-learning extracts (Chapter 1)

Verschiedene Teamformen

Teams können unterschiedliche Merkmale haben, was verschiedene Problemen mit sich bringen kann.

Es gibt verschiedene Teams:

– Generalisierte Teams.

– Expertenteams.

– Projektteams.

Es ist wichtig zu wissen, um welche Form des Teams es sich handelt.

Generalisierte Teams

• Die Mitglieder haben ähnliche Qualifikationen und arbeiten für längere Zeit zusammen.
• Alle können dieselben Aufgabe übernehmen wie es beispielsweise bei Ärzten in der Notaufnahme der Fall ist.

Expertenteams

• Die Mitglieder sind Spezialisten in verschiedenen Bereichen.
• Jedes Mitglied hat universelles Expertenwissen wie zum Beispiel ein Facharzt im Krankenhaus.

Projektteams

• Die Arbeit der Mitglieder besteht aus einer befristeten Projektarbeit.
• Diese Aufgabe gehört nicht zum Tagesgeschäft der Teamglieder.
• Ein Beispiel ist eine neue Produkteinführung.

Figure 4. Chapter 1 (section 1-4) of the e-learning as used in Study I Experiment 1 (medium text difficulty, text-only).

Figure 5. Difficulty versions of Chapter 7 (Section 1)

Das Modell von Gersick (1988)

- Es gibt ein weiteres Modell zur Entwicklung von Teams. Dieses Modell hat Gersick (1988) erstellt.
- Gersick nannte es das „Punctuated-Equilibrium"-Modell.
- Das Modell eignet sich für Teams, die für kurze Zeit zusammenarbeiten.

Entwicklungsphasen

- Das Modell von Gersick hat nur zwei Phasen.
- Die Phasen sind dem Modell von Tuckman und Jensen zum Teil ähnlich.
- Die zwei Phasen umfassen:

– Erste Phase: Beginn der Arbeit und Übergang.

– Zweite Phase: Leistungsphase des Teams.

Easy version

Das Modell von Gersick (1988)

Gersick (1988) erstellte ein weiteres Modell zur Teamentwicklung, das „Punctuated-Equilibrium"-Modell.

Das Modell eignet sich besonders für Teams, die miteinander für befristete Zeiträume arbeiten.

Entwicklungsphasen

Im Gegensatz zu dem vorherigen Modell, hat das „Punctuated-Equilibrium"-Modell nur zwei Phasen, sie beinhalten jedoch einige derselben Prozesse wie das Modell von Tuckman und Jensen.

Die zwei Phasen umfassen Beginn der Arbeit und Übergang sowie die Leistungsphase des Teams.

Medium version

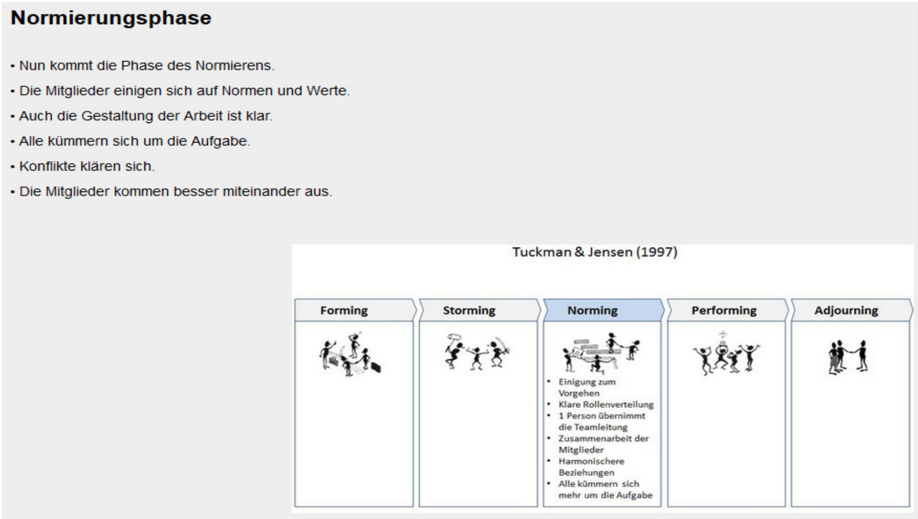
Das von Gersick 1988 erstellte Teamentwicklungsmodell wird als „Punctuated-Equilibrium"-Modell bezeichnet und eignet sich insbesondere für Teamzusammensetzungen, die miteinander für befristete und entsprechend kürzere Zeiträume zusammenarbeiten sollen.

Im Gegensatz zu dem vorherigen Teamentwicklungsmodell, differenziert das „Punctuated-Equilibrium"-Modell nur zwischen zwei Entwicklungsphasen, die allerdings einige Prozesse beinhalten, die sich mit den Entwicklungsphasen aus dem Modell von Tuckman und Jensen vergleichen lassen, so umfasst zum Beispiel die erste Phase den Beginn und den Übergang zur Arbeit während die zweite Phase als die Leistungsphase des Teams bezeichnet wird.

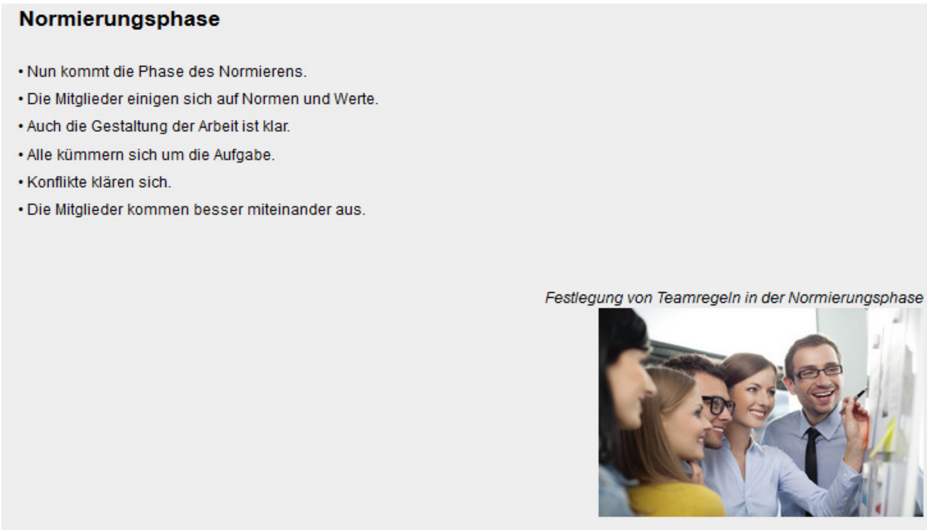
Difficult version

Figure 5. Each e-learning chapter has an easy, medium and difficult version. In Study I and III the medium text difficulty chapters were used. In Study II text difficulty was the main manipulation and consequently all text difficulty levels were used for different groups.

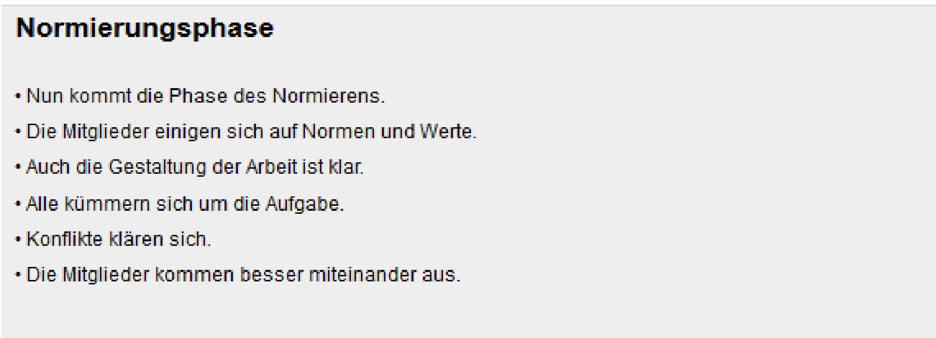
Figure 6. Multimedia versions of Chapter 4 (Section 1)



Functional images



Decorative images



Text-only

Figure 6. Manipulating the image condition in Study III, each chapter in the conceptual images group was supplemented by a chart that was informational equivalent to the text, summarizing the chapter content to support comprehension of the key message. Texts in the decorative images group came with photographs that were related to the topic of the text, but did not offer any information relevant for understanding the concept described in the text. As a control group, the text-only group received no images. In Study I and II the text-only versions were used.

Curriculum Vitae

For reasons of data protection, the curriculum vitae is not included in this version.

Curriculum Vitae

For reasons of data protection, the curriculum vitae is not included in this version.

Curriculum Vitae

For reasons of data protection, the curriculum vitae is not included in this version.