RAINER BROMME, STEPHANIE PIESCHL & ELMAR STAHL

EPISTEMOLOGICAL BELIEFS AND STUDENTS' ADAPTIVE PERCEPTION OF TASK COMPLEXITY

INTRODUCTION

Epistemological Beliefs

Epistemological beliefs are usually defined as beliefs about knowledge and knowing. One of the most widely used framework within educational psychology (Buehl & Alexander, 2001; Hofer & Pintrich, 1997), among others widely used (Niessen, Vermunt, Abma, Widdershoven, & van der Vleuten, 2004), comprises four identifiable and more or less interrelated dimensions of beliefs. According to Hofer and Pintrich (1997) the first two dimensions represent the "nature of knowledge": (a) the certainty of knowledge is focused on the perceived stability and the strength of supporting evidence, and (b) the structure of knowledge describes the relative connectedness of knowledge. The other two dimensions describe the nature of "knowing": (a) the justification of knowledge explains how individuals proceed to evaluate and warrant knowledge claims, and (b) the source of knowledge describes where knowledge resides, internally and/or externally. In the remainder of this paper we will primarily refer to this framework, although, from the beginning alternative frameworks have been proposed (for an overview with an emphasis on the assumed dimensions, see Buehl, 2008). More recently, Greene, Azevedo, and Torney-Purta (2008), for example, proposed an alternative framework in which "justification", either personal or by authorities, was proposed as the core epistemic dimension, whereas beliefs about the simplicity and certainty of knowledge are coined "ontological" because they refer to learners' assumptions about the structure of the categorical representation of the world. In a similar vein, Bromme, Kienhues, and Stahl (2008) have argued that epistemological judgments are based on topic- and domain-related ontological assumptions. Both propositions point to the interplay between epistemological beliefs and topic- and domain-related knowledge, because ontological assumptions are very abstract knowledge about a domain. In the present study the interplay between epistemological beliefs and domain-specific knowledge is investigated.

An important assumption in epistemological theories is that learners' epistemological beliefs develop (or should develop) from the so-called "naïve" to "sophisticated" epistemologies (Hofer & Pintrich, 1997). The term "naïve" is used, for example, to indicate a person's belief that the knowledge to be learned consists

S. Krolak-Schwerdt et al., (Eds.), Teachers' Professional Development, 123–151.

of a stock of certain facts related to each other additively and whose veracity is guaranteed by an authority. Such facts, once found, mirror the world unambiguously. Through (formal) education people become aware that knowledge is, for example, more complex and relativistic thus resulting in a focus on the evaluation of different viewpoints (King & Kitchener, 2002). Persons with sophisticated perspective believe, for example, that the veracity of knowledge claims depends on context and is continuously established within social interactions, and that knowledge is rather a complex network of facts, theories and conjectures than a pure addition of true facts. They accept uncertainty and changeability of truth and the notion that knowledge is construed rather than given; however, this does not mean that it would be reasonable to conceive each knowledge claim in each context that way, for example, to doubt that the earth is (nearly) round. On the contrary, sophistication entails adaptability to contextual demands (Bromme, Kienhues, & Porsch, 2008; Elby & Hammer 2001). In the present study students' adaptivity to differences between learning tasks when planning their use of a complex learning environment was investigated. Thereby, the assumed relationship between epistemological beliefs, domain-specific knowledge, and adaptive planning behavior was scrutinized.

An increasing number of empirical studies show that more sophisticated epistemological beliefs are related to more adequate learning strategies and better learning outcomes. For example, students' epistemological beliefs have been found to influence their processing of information (Schommer, 1990), their academic performance (Schommer, 1993), their conceptual change (Mason & Boscolo, 2004), their quality of argumentation (Weinstock & Cronin, 2003), and their engagement in learning (Hofer & Pintrich, 1997). Although there are fewer studies concerning computer-based learning environments their results are encouraging as well. For example, Jacobson and Spiro (1995) found that learners with sophisticated epistemological beliefs were better able to learn and apply their knowledge after using a hypertext system than students with naïve epistemological beliefs. Additionally, epistemological beliefs were a good predictor of learning outcomes during hypertext learning (Bendixen & Hartley, 2003; Windschitl & Andre, 1998). There is also evidence that epistemological beliefs affect students' information retrieval from the Internet (Bråten, Strømsø, & Samuelstuen, 2005; Hofer, 2004), especially in more open-ended tasks (Tu, Shih, & Tsai, 2008), and the understanding of multiple documents (Strømsø, Bråten, &Samuelstuen, 2008). With regard to *self-regulated learning*, epistemological beliefs have also been related to the use of more self-reported (Cano, 2005; Dahl, Bals, & Turi, 2005; Neber & Schommer-Aikins, 2002) as well as concurrently measured (Kardash & Howell, 2000) metacognitive strategies, better metacognitive comprehension monitoring (Schommer, 1990), and more metacognitively controlled help-seeking in a hypertext (Bartholomé, Stahl, Pieschl, & Bromme, 2006).

How are epistemological beliefs related to metacognition? Most recent views on epistemological beliefs and learning conceive such beliefs as being involved in metacognitive processes of monitoring (Kuhn, 2000). For example, Kitchener

(1983) proposed three levels of cognitive processing: (a) *cognition* (all cognitive operations such as reading, memorizing, perceiving, computing), (b) *metacognition* (all cognitions that have cognitive operations as their subjects, for example comprehension monitoring), and (c) *epistemic cognition* (cognitions about the limits of knowing, the certainty of knowledge, or the criteria for knowledge). Only the epistemic cognitions are assumed to be involved in monitoring the epistemic nature of problems and the evaluation of solutions. To give another example, Hofer (2004) details how epistemological belief dimensions can be matched to components of metacognition. Beliefs about the nature of knowledge (certainty of knowledge and simplicity of knowledge) are assumed to share similarities with declarative metacognitive knowledge (Schraw & Moshman, 1995). Beliefs about the nature of knowing on the other hand (source and justification of knowledge) can be matched to the procedural component of metacognition, for example metacognitive monitoring (Schraw & Moshman, 1995). These models, concerned with structural aspects of epistemological beliefs (where they are located in the cognitive system), are promising but more functional theories about the impact of epistemological beliefs are rare (i.e., how do they exactly exert their influence?).

The COPES Model

An encouraging theoretical framework that helps to specify such a functional relationship is given by the COPES model of self-regulated learning (Greene & Azevedo, 2007; Muis, 2007; Winne & Hadwin, 1998): Epistemological beliefs are conceptualized as important internal conditions for learning, which impact learners' internal standards for metacognitive monitoring and control and, thereby, influence the whole learning process. More specifically, self-regulated learning according to the COPES model occurs in four weakly sequenced and recursive stages: (a) task definition, (b) goal setting and planning, (c) enactment, and (d) adaptation. In the task definition stage, a student generates her or his own perception about what the studying task is and what constraints and resources are in place. An important product of this stage is the student's perception of the given goal of the task. Based on this perception the student generates idiosyncratic goal(s) and constructs a plan for addressing that study task in the second stage. In the enactment stage the previously created plan of study tactics is carried out. The adaptation stage pertains to fine-tuning of strategies within the actual learning task as well as to long-term adaptations.

All four stages are embedded in the same general cognitive architecture. In the centre of this architecture are processes of metacognitive monitoring and control that students are assumed to use to self-regulate their learning process according to perceived task demands. If and how these processes occur depends on five constituents whose acronym gave the model its name, namely conditions (C), operations (O), products (P), evaluations (E) and standards (S). Conditions (C) pertain to external task conditions (e.g., task complexity) as well as to internal conditions (e.g., epistemological beliefs) and are assumed to directly influence learners' internal standards and their operations. Operations (O) include all cognitive processes that learners utilize to solve a learning task and which create internal or external products (P) (e.g., written answers). Students' goals are represented as multivariate profiles of standards (S) (e.g., targeted level of understanding). As a result of metacognitive monitoring, evaluations (E) are generated based on a comparison of students' products and standards. When a learner notices discrepancies she or he is able to perform metacognitive control by executing fix-up operations. To summarize, the COPES model describes how students might adapt their self-regulated learning process to important external conditions such as task complexity. Furthermore, it specifies the impact of learner-related internal conditions such as epistemological beliefs.

The Present Study

Based on the COPES model but with a special focus on the impact of epistemological beliefs on adaptivity to task complexity, the present study explores if epistemological beliefs affect processes of *metacognitive calibration.*

Traditionally, calibration refers to the accuracy of a person's subjective metacognitive judgments (e.g., their judgments of learning (JOL) regarding the confidence in recall) regarding their own objective performance (e.g., in a recall task of paired associates such as "ocean-tree"; example taken from Nelson & Dunlosky, 1991). Multiple measures of accuracy have been suggested in the literature on traditional calibration: The most frequently used method is *relative calibration* (Nelson, 1996) which denotes the degree of association between judgments and performance (e.g., the Goodman-Kruskal's Gamma correlation used by Nelson & Dunlosky, 1991). Additionally, indices of *absolute calibration* are often computed that indicate the exact degree of over- or underconfidence of judgments in relation to performance (e.g., bias score, see Schraw, 1995). Furthermore, measures of *discrimination* denote the ability to discriminate between the occurrence and the nonoccurrence of an event, for example predict correct versus incorrect performance (see Weingardt, Leonesio, & Loftus, 1994).

For the present study, we transferred the methodology of traditional calibration research outlined above to a new context (for more detail see Pieschl, 2009): We define metacognitive calibration as the alignment between learners' subjective *task definitions, goals and plans* (measured by the COPES questionnaire, see below) and objective task demands, more specifically *task complexity* (operationalized by Bloom's revised taxonomy; Anderson et al., 2001, see below). Therefore, metacognitive calibration in this sense denotes the degree of adaptivity to task complexity. Note that our definition of calibration is conceptually different from the traditional one, but that the same methods are applied: We assume that students *discriminate* between tasks of different complexity by indicating different task definitions, goals and plans. And we assume that students systematically *calibrate* their task definitions, goals and plans to task complexity, for example by planning more elaborate learning strategies for more complex tasks.

Assuming that epistemological beliefs affect metacognitive calibration as outlined above implies not only main effects of epistemological beliefs but also potential interactions between epistemological beliefs and task complexity. To illustrate the potential *main effects* of epistemological beliefs imagine a learner with a naïve belief that knowledge is simple and stable. As epistemological beliefs are assumed to directly influence the learner's internal standards, the learner might set quite superficial goals (e.g., "The goal is achieved if I have memorized the facts"; "I will complete this task in a short time") compared to a more sophisticated learner who believes that knowledge is complex and relative (e.g., "I have to deeply understand the subject-matter in order to apply it"; "I will need much time to complete the task"). To give another example, epistemological beliefs are also assumed to directly influence the learner's operations; thus, a more naïve learner might also plan rather superficial learning tactics and strategies for task completion (e.g., memorizing) compared to a more sophisticated learner who might plan strategies of deeper elaboration (e.g., critically evaluating).

To illustrate potential *interactions* with task complexity, consider that such differences might become more pronounced for more complex tasks. Specifically, if learners are confronted, for example, with the complex task of writing a pro- and contra- argumentation about a controversial topic, this task might be interpreted in multiple ways. A student who believes that knowledge is uncertain (sophisticated belief) would probably plan to verify each argument by searching for additional information, whereas a more naïve student would probably take each argument at face value. For a very simple task like a factual question on the other hand, these potential differences should be smaller, that is, students with naïve beliefs are assumed to approach such task superficially because they are assumed to have a general bias to underestimate task complexity and thus might approach all tasks too superficially. On the other hand, students with sophisticated beliefs should be able to accurately diagnose task demands and thus also plan adequately superficial strategies. Therefore, it was hypothesized that students with more sophisticated epistemological beliefs should show a better fit between external task demands such as task complexity and their self-regulated learning process.

Within a large project (for an overview: Bromme, Pieschl, & Stahl, 2010; Pieschl, Stahl, & Bromme, 2013) this general assumption about epistemological beliefs as important predictor of metacognitive calibration was tested with regard to each stage of self-regulated learning as defined in the COPES model. The present study – as well as an exploratory study already conducted (Stahl, Pieschl, & Bromme, 2006) – focused on the first two preparatory stages of self-regulated learning, that is, on COPES' *task definition* and *goal setting and planning*. All studies within this large project have common elements, that is, students work with (or plan to work with) a hypermedia information system on "genetic fingerprinting". This topic was chosen because it was judged inherently interesting by students, there was sufficient variance in students' epistemological beliefs towards this topic, and because this domain contains certain facts (e.g., "DNA contains four bases: adenine, cytosine, guanine, and thymine.") as well as controversial issues (e.g., "Should we compile comprehensive data bases of DNA profiles?"). Additionally, in all studies within this large project students have to complete (or plan to complete) learning tasks with different levels of complexity according to Bloom's revised taxonomy (Anderson et al., 2001). In this theoretical framework, task complexity is not defined quantitatively (e.g., by the number of necessary operations), but rather by the quality of the required cognitive processes. For example, factual questions, such as "What is the capital of Germany?" (correct answer: Berlin), are always considered simple because they just require recall of information from long-term memory. This even holds if task difficulty (i.e., the proportion of incorrect solutions in an empirical sample) is high (e.g., "What is the capital of Mongolia?" correct answer: Ulaanbaatar) or if a lot of similar questions need to be answered (e.g., questions about the capitals of all countries in the world). On the other hand, questions that require more complex cognitive elaboration processes, such as evaluating evidence (e.g., regarding the suitability of DNA analysis methods) with regard to some standards (e.g., error-proneness regarding lab results or statistical analyses), are always considered more complex – independently of task difficulty.

Research Questions – Hypotheses

In the present study, three research questions were addressed based on these theoretical considerations. The first questions pertain to adaptation to task complexity: (a) Do students *discriminate* between tasks of different complexity? We predicted that students would discriminate significantly between *Task Levels* of different complexity in their task definitions, goals and plans and that this would be evident in their answers in the COPES-questionnaire (Hypothesis 1). In short (for more information see method section), scales of the COPES questionnaire indicate task definitions, goals and plans either for deep or for superficial processing. (b) Do students *calibrate* their task definitions, goals and plans to task complexity? We predicted that students would calibrate their answers on the COPES-questionnaire significantly to task complexity and this would be evident by a systematic relationship between students' answers in the COPES-questionnaire and *Task Levels* of different complexity. More specifically, we predicted that students would judge all variables indicating deep processing more important for more complex tasks and all variables indicating superficial processing less important for more complex tasks (Hypothesis 2).

Further question pertain to the impact of personal characteristics: (c) Are these *discrimination* and *calibration* processes related to students' learner characteristics? First, we predicted effects of *epistemological beliefs*. More specifically, we predicted that students with more sophisticated epistemological beliefs would judge all variables indicating deep processing more important across all tasks and would judge all variables indicating superficial processing less important across all tasks (Hypothesis 3; main effects). Additionally, we predicted that these differences between "sophisticated" and "naïve epistemological beliefs would be more pronounced in more complex *Task Levels* (Hypothesis 4; interaction effects). Second, we predicted similar effects of prior domain-specific knowledge. Prior domain-specific knowledge showed a crucial impact on learning processes in most other studies (Lind & Sandmann, 2003; McDonald & Stevenson, 1998). The COPES model (Winne & Hadwin, 1998) predicts a similar functional relationship for prior domain-specific knowledge as for epistemological beliefs. In this study, we systematically compared two Prior Knowledge Groups: Biology students with high prior biology knowledge and humanities students with almost no prior biology knowledge. More specifically, we predicted that students with higher prior domain-specific knowledge would judge all variables indicating deep processing more important across all tasks and would judge all variables indicating superficial processing less important across all tasks (Hypothesis 5; main effects). Additionally, we predicted that the differences between high and low prior domain-specific knowledge would be more pronounced in more complex *Task Levels* (Hypothesis 6; interaction effects).

METHOD

Procedure

The present study was conducted in two sessions. During the first online session, students filled in questionnaires about their domain-general (*EBI*; Jacobson & Jehng, 1999) and domain-dependent (*CAEB*; Stahl & Bromme, 2007) epistemological beliefs, which took them about 15 minutes. Sixty-five biology students and 64 humanities students completed these online-questionnaires. The second face-to-face session was held in groups with a minimum of 3 students and a maximum group size of 12 and lasted approximately one hour. Not all students continued; 52 biology students (80% of the original sample) and 50 humanities students (78% of the original sample) participated in this second session where they had to fill in paper-pencil-questionnaires. First, students had to answer a *Short Knowledge Test* about molecular genetics and the *Self-Rated Prior Biology Knowledge* item. Then, all students read a factual introduction[1] to molecular biology which adequately contextualized students to the topic of "genetic fingerprinting". In the main part of this session, students evaluated six learning tasks of different complexity according to Bloom's revised taxonomy (Anderson et al., 2001: *remember, understand, apply, analyze, evaluate*, and *create*) with the *COPES-questionnaire*. Tasks were presented in random order.

Participants

Students who participated in both sessions constitute the final sample. All students were selectively recruited to ensure two levels of biology knowledge. Biology students were recruited during regular courses in biology; humanities students were recruited by a posting at the psychological institute. All students received 10 Euros reimbursement. Although the advanced students of biology were no "real" experts in the specific topic of "genetic fingerprinting" (Chi, 2006), they can be considered discipline experts (Rouet, Favart, Britt, & Perfetti, 1997) because they know the tools of their discipline, for example how to interpret an electrophoretogram. Students of humanities on the other hand can be considered novices (Chi, 2006).

The 52 biology students' (35 female) mean age was 22.10 years (SD = 2.19) and they studied on average in the 3.00rd semester (SD = 0.28) biology or related majors. The 50 humanities students' (43 female) mean age was 23.61 years (SD = 4.47) and they studied on average in the 4.24th semester (SD = 2.24) psychology or other humanities. Biology students significantly outperformed humanities students on a *Short Knowledge Test* (see below; t (100) = 20.63, p < .001, Cohen's d = 4.08; biology students: M = 7.23, SD = 1.06; humanities students: M = 2.16, SD = 1.40; 8 points maximum). Furthermore, they also possessed higher *Self-Rated Prior Biology Knowledge* (see below; t (99) = 5.60, p < .001, Cohen's d = 1.11; biology students: M = 2.79, SD = 0.73; humanities students: M = 1.90, SD = 0.87; on a scale from 1 (very low) to 5 (very high)). Thus, these two quasi-experimental groups of students (*Prior Knowledge Groups:* biology vs. humanities students) were used as predictor variable in all subsequent analyses to explore the effects of prior domain knowledge.

Measures

Short Knowledge Test Background knowledge in molecular biology was tested with eight multiple-choice questions (Cronbach's α = .89) that were developed with the help of a domain expert. Sample item: "What does the abbreviation PCR stand for?" Multiple-choice options: "(1) Protein Coupling Reaction, (2) Phosphate Chain Reaction, (3) *Polymerase Chain Reaction*, (4) Polysaccharide Chain Reaction, (5) Phosphate Coupling Reaction, or (6) I don't know." Each question had one correct answer (in the example in italics).

Self-Rated Prior Biology Knowledge Students were also asked to self-assess their own knowledge in genetics with the item: "I estimate my prior domain-specific knowledge in genetics to be" Answers could be given on a Likert-type scale ranging from 1 (very low) to 5 (very high).

Epistemological beliefs For the measurement of epistemological beliefs the distinction between explicit-denotative and associative-connotative aspects of epistemological beliefs was used. The distinction has been proposed by Stahl and Bromme (2007) because of the often reported problems of measuring epistemological beliefs in a reliable way (Niessen et al., 2004; Strømsø et al., 2008). The two aspects of epistemological beliefs are not necessarily in accordance with each other and they have to be measured separately.

The *explicit-denotative* aspects of epistemological beliefs were measured with an adapted version of the Epistemological Beliefs Instrument (EBI; Jacobson & Jehng, 1999) that comprises items such as "If scientists try hard enough, they can find the answer to almost every question" that had to be rated on a Likert-type scale ranging from 1 (totally disagree) to 7 (totally agree). The original instrument consists of 61 items. However, for this study only items that refer to epistemology in a strict sense were selected, more specifically from the scales *certainty of knowledge* (9 items), *omniscient authority* (5 items), and *simple view of learning* (3 items). Furthermore, we added 4 items from Wood and Kardash's (2002) questionnaire and 2 items from our own lab. The exploratory factor analysis of these 23 items of the adapted EBI applied to the sample of the present study yielded one factor explaining 35.91 % of variance; this scale was labelled *EBI-definitude.* This scale measures whether students assume that absolute answers are attainable or whether knowledge is indefinite (9 items, Cronbach's $\alpha = .76$; sample items are "For most scientific research questions there is only one right answer.", "Most words have one clearly defined meaning.").

To capture the *associative-connotative* aspects of epistemological beliefs about knowledge in the domain of genetics a semantic differential, namely the Connotative Aspects of Epistemological Beliefs (CAEB; Stahl & Bromme, 2007), was used. This instrument consists of 24 pairs of antonymous adjective as items; on each item the degree of association could be rated on a 7-point scale. Sample item: "Knowledge in genetics is: simple (1) – complex (7)". The exploratory factor analysis of the 24 items of the CAEB yielded two factors explaining 50.39 % of the variance, namely CAEB-texture and CAEB-variability. The factor *CAEB-texture* encompasses beliefs about the structure and accuracy of knowledge (9 items loaded on this factor). Sample items are "Knowledge in genetics is: from 1 (precise / sorted / exact / etc.) to 7 (imprecise / unsorted / vague / etc.)"; Cronbach's $\alpha = .82$. The factor *CAEB-variability* encompasses beliefs about the stability and dynamics of knowledge (5 items loaded on this factor). Sample items are "Knowledge in genetics is: from 1 (irrefutable / flexible / completed / etc.) to 7 (refutable / inflexible" / uncompleted / etc.)"; Cronbach's $\alpha = .67$.

These three factors, namely *EBI-definitude, CAEB-texture*, and *CAEB-variability*, were used as predictor variables in all relevant subsequent analyses to explore the effects of epistemological beliefs.

Tasks Six tasks of different complexity according to Bloom's revised taxonomy (Anderson et al., 2001) were presented. This taxonomy distinguishes between six task classes affording cognitive processes of different complexity (in order of ascending complexity): (a) remember, (b) understand, (c) apply, (d) analyze, (e) evaluate, and (f) create. For the present study, one task for each Bloom category was constructed and selected in a cyclic process. First, two experts in biology searched through relevant textbooks for adequate tasks and constructed additional tasks for all categories.

Second, the resulting pool of about 100 tasks was independently categorized by five raters into the six Bloom categories; these raters were blind to the experts' categorization. For 39 tasks all raters immediately agreed, for further 25 tasks four of the raters agreed; the remaining tasks were either rephrased and re-categorized (15 tasks) or deleted from the pool. Third, based on content considerations six tasks per Bloom category were selected for an exploratory study (Stahl et al., 2006). Fourth, for the present study only the most prototypical task for each Bloom category was chosen based on participants' categorizations in the Stahl et al. (2006) study.

As simplest *remember* task a multiple-choice question about how to split DNA was selected; the answer only required recall of facts. As *understand* task a multiple-choice question about which errors in STR (Short Tandem Repeats) profiling could cause an erroneous match was used; to answer this question an understanding of the whole process was necessary. The *apply* task required constructing a father's DNA profile from the profiles of his wife and his biological daughters in a table; knowledge about the heredity of DNA had to be applied to this concrete problem. The *analyze* task required to detail the STR analysis process step-by-step and outline potential problems; it required participants to have a detailed mental model of the whole process. The *evaluate* task asked to evaluate the impact of DNA degradation on different methods of DNA analysis in an open answer format; it required knowledge about this topic as well as critical thinking. The most complex *create* task required describing the consequences of a law change that would allow the analysis of coding DNA regions in an open answer; this task required original and creative thinking.

All six tasks were presented in random order to each participant. Participants did not solve these tasks but had to evaluate each task with the COPES questionnaire. In this study students' adaptation to these *Task Levels* of different complexity was explored (remember, understand, apply, analyze, evaluate, and create).

The COPES Questionnaire The COPES questionnaire (Stahl et al., 2006) measures students' judgments regarding their preparatory stages of self-regulated learning, namely task definition, goal setting and planning (Winne & Hadwin, 1998) and consists of 46 items. The whole questionnaire was administered in this study for each task, but only 18 items were further analysed, namely those items where participants of an exploratory study (Stahl et al., 2006) demonstrated significant discrimination and calibration. These items cover most facets of self-regulated learning (i.e., conditions, operations, evaluations, and standards).

Two of these items required short open answers; students had to estimate the number of concepts (*estimated concepts*) and the time needed for task completion (*estimated time*). One item (*Bloom classification*) had a forced-choice format with six alternative answers that represent the *Task Levels* of different complexity according to Bloom's revised taxonomy.

The remaining fifteen items were rated on 7-point Likert-type scales, mostly ranging from very unimportant (1) to very important (7); these fifteen items were subjected to an exploratory factor analysis on the present sample that explained 62% of variance and yielded three meaningful factors: *Deep Processing* (8 items, Cronbach's α = .89; sample item: "Imagine you would have to actually solve the present task. In your opinion, how unimportant or important is it to employ the learning strategy of 'elaborating deeply'?"), dealing with *Multiple Information Sources* (5 items, Cronbach's α = .82; sample item: "… In your opinion, how unimportant or important is it to concentrate on information about 'multiple perspectives'?"), and *Superficial Processing* (2 items, Cronbach's α = .63; sample item: "… In your opinion, how unimportant or important is it to employ the learning strategy of 'memorizing'?"). These results indicated that the items were not grouped together according to the facets of the COPES model but rather according to three different approaches to learning.

These three COPES factors, namely *Deep Processing, Multiple Information Sources,* and *Superficial Processing*, as well as the three single items, namely *estimated concepts, estimated time*, and *Bloom classification*, were used as dependent variables in all subsequent analyses, each repeatedly measured six times for the six *Task Level*s representing the Bloom categories. No theoretical assumptions were made about the importance of these factors and items for the tasks of different complexity; rather the students' opinions were important.

RESULTS

Descriptives and Interrelations Regarding Learner Characteristics

The two *Prior Knowledge Groups* (biology students vs. humanities students) did not differ in their domain-related epistemological beliefs measured by the CAEB. On average students believed that knowledge in genetics is quite structured (*CAEB-texture*: M = 3.33, SD = .80; on a scale from 1 = structured – 7 = unstructured) but tentative (*CAEB-variability*: M = 3.04, SD = .85; on a scale from 1 = variable – 7 = static). However, with regard to the definitude of knowledge in general (*EBI-definitude*) the *Prior Knowledge Groups* differed significantly (F (1,100) = 10.55, p < .01, d = .64): Humanities students believed much less (M = 2.59, SD = .71) in the definitude of knowledge in general than did biology students (M = 3.09, SD = .84; on a scale from 1 = knowledge is indefinite – 7 = absolute answers are attainable).

Furthermore, the domain-general scale of the EBI (*EBI-definitude*) was not correlated significantly to any of the domain-related scales of the CAEB. However, the two domain-related scales were significantly interrelated (r = -.52, p < .001): A strong belief in structured knowledge in genetics (low value on *CAEB-texture*) was related to a strong belief in static knowledge in genetics (high value on *CAEB-variability,* the inverse relationship is due to the construction of the CAEB scales. Both endpoints point to a more 'naïve' view).

Do Students Discriminate between Tasks of Different Complexity?

We hypothesized that students should discriminate between tasks of different complexity which should be evident in their significantly different answers in the

COPES questionnaire regarding different *Task Levels* (Hypothesis 1). To test this hypothesis, we computed a MANOVA for the three COPES factors (*Deep Processing; Multiple Information Sources*; and *Superficial Processing*) with *Task Levels* as repeated-measure factor. We computed similar ANOVAs for the three remaining single items (*estimated time; estimated concepts*; and *Bloom classification*). Thus, we expected seven main effects of the repeated-measure factor *Task Level* (six univariate main effects plus one multivariate main effect).

The repeated-measure MANOVA for the three COPES factors (see Table 1 descriptives; Table 2 results) showed a multivariate main effect for the repeated-measure factor *Task Levels* which was replicated univariately on each single COPES factor (*Deep Processing; Multiple Information Sources*; and *Superficial Processing*). Exploring this question in more detail, we additionally compared the adjacent *Task Levels* statistically (after Bonferroni correction with alpha $p < .01$).

Table 1.Means and standard deviations (in brackets) for all dependent variables (rows) with regard to all Task Levels (columns)

Dependent Variable	*RE*	*UN*	*AP*	*AN*	*EV*	*CR*
Deep Processing	2.62 (1.18)	3.17 (1.05)	4.62 (1.16)	5.23 (.93)	4.86 (.99)	4.76 (.92)
Multiple I. Sources	2.28 (1.04)	3.24 (1.16)	2.98 (1.05)	3.57 (1.28)	3.84 (1.07)	5.56 (.92)
Superficial Pro.	5.52 (1.51)	3.94 (1.70)	4.05 (1.39)	4.21 (1.28)	4.10 (1.24)	2.71 (1.18)
Estimated time	7:42 (13:07)	9:49 (11:47)	44:19 (76:59)	52:39 (51:26)	38:13 (61:23)	37:11 (64:46)
Estimated concepts	2.11 (2.17)	2.51 (2.18)	4.20 (4.11)	5.85 (4.89)	5.12 (4.15)	4.99 (11.07)
Bloom classification	1.17 (.48)	2.39 (1.33)	3.28 (.91)	3.31 (1.32)	3.75 (1.34)	5.05 (1.01)

Task Levels: RE = remember, UN = understand, AP = apply, AN = analyse, EV = evaluate, and CR = create; Dependent Variables: Multiple I. Sources = Multiple Information Sources and Superficial Pro. = Superficial Processing.

For the COPES factor *Deep Processing remember* and *understand* (F (1,101) = 22.91, $p < .001$, $\eta^2_p = .18$), *understand* and *apply* (F (1,101) = 126.99, $p < .001$, $\eta^2_p = .56$), *apply* and *analyze* (F (1,101) = 26.28, $p < .001$, $\eta^2_p = .21$), and *analyze* and *evaluate* (F (1,101) = 9.47, $p < .01$, $\eta^2_p = .09$) tasks differed significantly. This means that students successfully discriminated between *Task Levels* of different complexity except for the two most complex ones (*evaluate* and *create*). Furthermore, the descriptive values (see Table 1) show that they considered *Deep Processing* less important for the most complex tasks (*evaluate* and *create*) than for the moderately complex *analyze* task. For the COPES factor *Multiple Information Sources*, the following *Task Levels* differed significantly: *remember vs. understand* (F (1,101) = 70.64, p < .001, η^2_p = .41), *apply vs. analyze* (F (1,101) = 31.01, p < .001, η^2_p= .24), and *evaluate vs. create* (F (1,101) = 194.37, p < .001, η^2_p = .66). Thus, students successfully discriminated between four *Task Levels*: the simplest *remember* task, the little more complex *understand* and *apply* tasks, the moderately complex *analyze* and *evaluate* tasks, and the most complex *create* task. For the COPES factor *Superficial Processing*, the following *Task Levels* differed significantly: *remember vs. understand* (F (1,101) = 79.87, p < .001, η^2_p = .44) and *evaluate vs. create* (F (1,101) = 105.75, p < .001, η^2_p = .51). Thus, students successfully discriminated between three broader *Task Levels*: the simplest *remember* task, a range of moderately complex tasks (*understand, apply, analyze*, and *evaluate*), and the most complex *create* task (for descriptives see Table 1).

The repeated-measure ANOVAs also indicate significant effects of the repeated-measure factor *Task Levels* for each of the three remaining single items (*estimated time; estimated concepts*; and *Bloom classification*; see Table 2). Exploring these results in more detail, we report all significant differences between adjacent *Task Levels* (for descriptives see Table 1). For the variable *estimated time*, the following *Task Levels* differed significantly: *understand vs. apply* (F (1,101) = 22.72, p < .001, η^2_p = .18). Thus, students successfully discriminated between two *Task Levels*: simple tasks (*remember* and *understand*) and complex tasks (*apply, analyze, evaluate*, and *create*). For the variable *estimated concepts*, the following *Task Levels* differed significantly: *understand vs. apply* (F (1,98) = 22.70, p < .001, η^2_p = .19) and *apply vs. analyze* (F (1,98) = 18.21, p < .001, η^2_p = .16). Thus, students successfully discriminated between three *Task Levels*: simple tasks (*remember* and *understand*), the mid-complex *apply* task, and complex tasks (*analyze, evaluate*, and *create*). For the variable *Bloom classification*, the following *Task Levels* differed significantly: *remember vs. understand* (F (1,94) = 69.94, p < .001, η^2_p = .43), *understand vs. apply* (F (1,94) = 30.43, p < .001, η^2_p= .25), and *evaluate vs. create* (F (1,94) = 64.43, p < .001, η^2_p = .41). Thus, students successfully discriminated between four Task Levels: the simplest *remember* task, one little more complex *understand* task, a range of moderately complex tasks (*apply, analyze*, and *evaluate*), and the most complex *create* task.

Do Students Calibrate their Judgments to Task Complexity?

We hypothesized that students should calibrate their judgments to task complexity which should be evident in systematic relationships between students' answers in the COPES-questionnaire and *Task Levels* of different complexity (Hypothesis 2). To test this hypothesis, we computed intra-individual Goodman-Kruskal Gamma correlations *(G)* between the *Task Levels* and each dependent variable (in all cases: n = 6, for six *Task Levels*) to diagnose calibration. These correlations were

Table 2.Repeated-measure (M)ANOVAs regarding the effects of Task Levels

	F	*df*	*df error*	*p*	*partial* η^2
repeated-measure MANOVA across Task Levels for the three COPES factors					
Task Levels (multivariate) +	60.22	15	87	< .001	.91
Deep Processing	148.38	5	505	< .001	.60
Multiple Information Sources	169.86	5	505	< .001	.63
Superficial Processing	62.78	5	505	< .001	.38
separate repeated-measure ANOVAs across Task Levels for the three single items					
Estimated time	17.15	5	97	< .001	.47
Estimated concepts	19.66	5	94	< .001	.51
Bloom classification	258.06	5	90	< .001	.94

+Multivariate effects; all values according to Pillai's trace; univariate effects indented.

Table 3. Calibration indices indicating the relationship between the dependent variables (rows) and Task Levels of different complexity

Dependent Variable	*Index M (SD)*	*Significance*	*G*
Deep Processing	.60 (.41)	t (101) = 14.69***, d = 1.46	.54
Multiple Information Sources	.95 (.70)	t (101) = 13.70***, d = 1.36	.74
Superficial Processing	-.61 (.88)	t (101) = -7.03***, d = .69	-.55
estimated time	.65 (.55)	t (101) = 11.96***, d= 1.18	.57
estimated concepts	.51 (.72)	t (101) = 7.13***, d = .71	.47
Bloom classification	1.23 (1.02)	t (101) = 12.23,***, d = 1.21	.85

**** p < .001; M = mean; SD = standard deviation; d = Cohen's d; G = Goodman-Kruskal Gamma correlation, in this column the G values that correspond to the calibration indices ("Index") are reported (reverse Z-transformation of the mean calibration indices).*

subsequently Z-transformed into calibration indices. We determined significance by statistically testing the magnitude of these average indices against zero. We expected six calibration indices of significant size, one for each of the six dependent variables (*Deep Processing; Multiple Information Sources; Superficial Processing; estimated time; estimated concepts;* and *Bloom classification*).

We found significant calibration indices for all dependent variables (see Table 3). For example, the positive correlation of $G = .54$ between students' answers regarding *Deep Processing* and *Task Levels* indicates that students judged *Deep Processing* to be quite unimportant for simple tasks and of ascending importance for more complex tasks (for descriptives see Table 1). Similar positive relationships were detected for *Multiple Information Sources, estimated time, estimated concepts,* and *Bloom* *classification* (Table 3). The negative correlation of $G = -.55$ (see Table 3) between students' answers regarding *Superficial Processing and Task Levels* on the other hand indicates the following: Students judged *superficial processing* to be quite important for simple tasks and of descending importance for more complex tasks.

In addition to relative calibration (see above), *absolute calibration* was explored for students' *Bloom classifications*. This was the only instance where absolute calibration could be analysed within this study. For all other dependent variables we had no comparative standard indicating what constitutes correct answers. But for *Bloom classifications*, students' classifications could be directly compared to the correct classifications (see methods section). Students on average classified 47.17 % of the six tasks correctly ($M = 2.83$, $SD = 1.26$) which is significantly more than could be randomly expected (namely one out of six; t (97) = 14.35, $p < .001$, Cohen's $d = 1.45$). The corresponding calibration graph (see Figure 1) shows that students slightly overestimated the complexity of simpler tasks (*remember – apply*) while they underestimated the complexity of more complex tasks (*analyze - create*) compared with hypothetically perfect classifications (indicated by the "line of perfect calibration" in Figure 1).

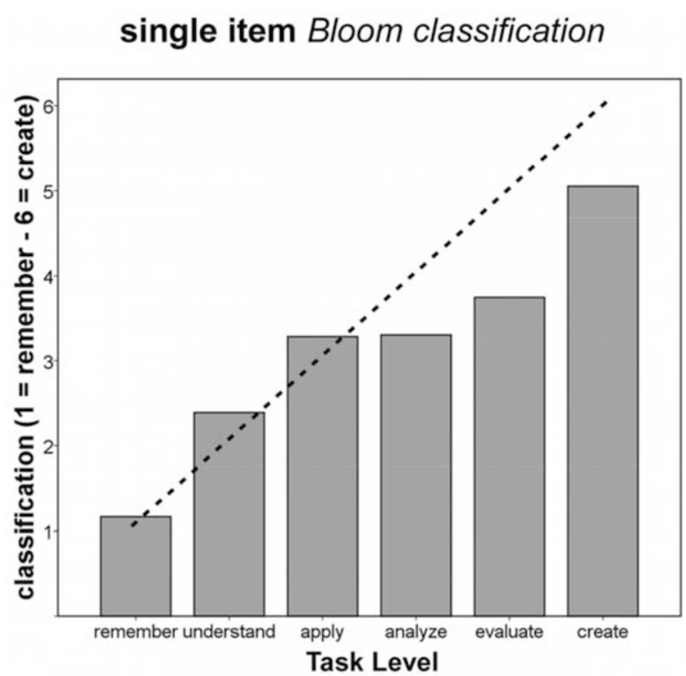


Figure 1.Calibration graph depicting students' Bloom classifications (Y-axis) as a function of Task Levels of different complexity (X-axis). The dotted line represents the hypothetical "line of perfect calibration" (perfectly correct classifications).

Are these Metacognitive Discrimination and Calibration Processes Related to Students' Learner Characteristics?

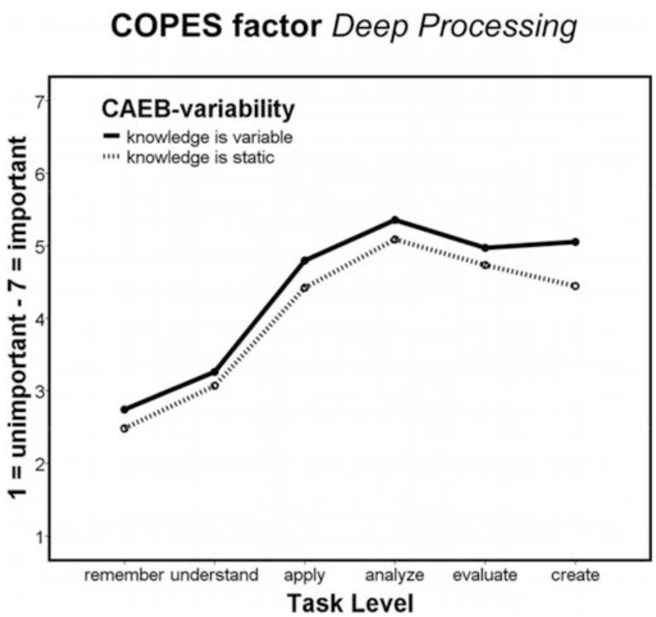
To test our hypotheses regarding learner characteristics and *discrimination*, repeated-measure analyses were computed including all learner characteristics simultaneously. More specifically, a repeated-measure MANCOVA was computed across the three COPES scales (*Deep Processing, Multiple Information Sources*, and *Superficial Processing*) with *Task Levels* as repeated-measure factor, the epistemological beliefs scales (*EBI-definitude, CAEB-variability*, and *CAEB-texture*) as covariates and the *Prior Knowledge Groups* (biology students vs. humanities students) as factor. Additionally, repeated-measure ANCOVAs were computed separately for each of the remaining single items (*estimated concepts, estimated time*, and *Bloom classification*) with the same covariates, repeated-measure and between-subject factors.

To test our hypotheses regarding learner characteristics and calibration (also see interaction effects above), correlations were computed between the calibration indices of all dependent variables (*Deep Processing, Multiple Information Sources, Superficial Processing, estimated concepts, estimated time*, and *Bloom classification*) and the epistemological beliefs scales (*EBI-definitude, CAEB-variability*, and *CAEB-texture*). Additionally, the calibration indices of all dependent variables were statistically compared between *Prior Knowledge Groups* (biology students vs. humanities students).

Note that even though in all cases all learner characteristics (epistemological beliefs scales *and* Prior Knowledge Groups) were simultaneously included in the analyses, we report the results separately: We will first report all results regarding epistemological beliefs, namely the results regarding main effects (Hypothesis 3) and the results regarding interaction effects (Hypothesis 4). Subsequently, we will report all results regarding prior domain-specific knowledge, namely the results regarding main effects (Hypothesis 5) and the results regarding interaction effects (Hypothesis 6). We will only report the significant results but we will point out the number of non-significant effects in each analysis.

Effects of Epistemological Beliefs We hypothesized that more sophisticated beliefs should be associated with judging all variables indicating deep processing more important across all tasks and with judging all variables indicating superficial processing less important (Hypothesis 3). Therefore, we expected a total of twenty-one main effects (18 univariate main effects of three epistemological belief scales regarding six dependent variables and 3 multivariate main effects of three epistemological beliefs scales).

In the MANCOVA across the three COPES factors we found a significant multivariate main effect of *CAEB-variability* (F (3,95) = 2.85, $p < .05$, $\eta^2_p = .083$) that was univariately replicated significantly on the COPES factors *Deep Processing* (F (1,97) = 5.23, $p < .05$, $\eta^2_p = .051$, Figure 2, top) and *Multiple Information Sources*



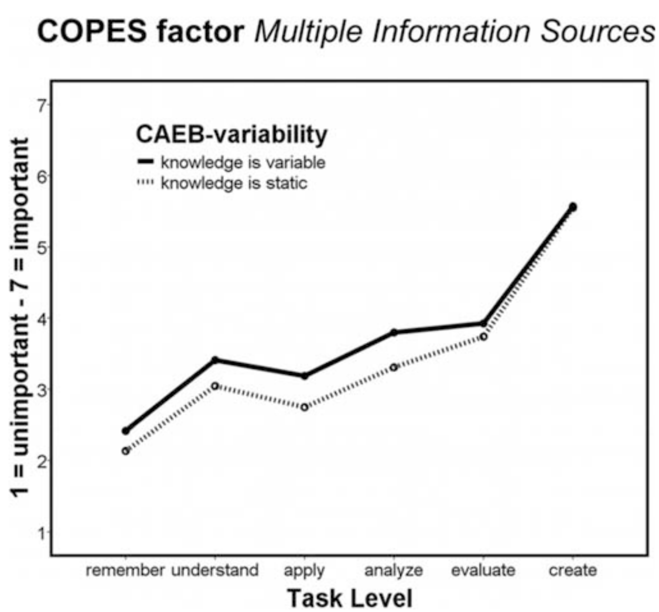


Figure 2. *Calibration graphs depicting students' judgments on the COPES factors Deep Processing (top) and dealing with Multiple Information Sources (bottom) as a function of Task Levels (X-axis) and CAEB-variability (median-split; lines).*

(F (1,97) = 7.58, $p < .01$, $\eta^2_p = .072$, Figure 2, bottom): Students, who considered knowledge in genetics variable (sophisticated view on *CAEB-variability*) also considered *Deep Processing* and *Multiple Information Sources* more important across all *Task Levels* than more naïve students. Note that these effects were visualized by median-splitting the scale *CAEB-variability* (Figure 2), but that *CAEB-variability* was included as covariate in the analyses!

Additionally, the ANCOVA for the single item *Bloom classification* indicates a significant main effect of *CAEB-variability* (F (1,90) = 4.59, $p < .05$, $\eta^2_p = .049$, without Figure): More sophisticated students who believed in variable knowledge in genetics classified tasks in more complex *Task Levels* (especially *analyze* tasks). To summarize: We found four significant main effects in the expected direction, all of the scale *CAEB-variability*; all other main effects of epistemological beliefs were not significant.

We hypothesized that the effects of epistemological beliefs would be more pronounced on more complex *Task Levels* (Hypothesis 4). Therefore, we expected a total of twenty-one interaction effects (18 univariate interactions between *Task Levels* and three epistemological belief scales regarding six dependent variables and 3 multivariate interactions between *Task Levels* and three epistemological beliefs scales). Additionally, we expected a total of eighteen significant correlations with calibration indices (for each of three epistemological beliefs scales with six dependent variables).

In the MANCOVA across the three COPES factors we found a significant univariate interaction between *CAEB-variability* and the repeated-measure factor *Task Levels* for the COPES factor *Multiple Information Sources* (F (5,485) = 2.34, $p < .05$, $\eta^2_p = .024$, Figure 2, bottom). The above-mentioned main effect of *CAEB-variability* was most pronounced for the *Task Levels remember* through *analyze*, while it disappeared for the more complex tasks *evaluate* and *create*. Furthermore, we found one significant correlation with a calibration index: More naïve beliefs in the definitude of knowledge in general (*EBI-definitude*) were significantly associated with higher calibration indices regarding *estimated concepts* ($r = .26$, $p = .009$). To summarize: We found two effects indicating interactions between epistemological beliefs and task complexity (*Task Levels*), both counterintuitive. All other interaction effects and effects on calibration were not significant.

Effects of Prior Domain-Specific Knowledge We hypothesized that more domain-specific knowledge should be associated with judging all variables indicating deep processing more important across all tasks and with judging all variables indicating superficial processing less important (Hypothesis 4). Therefore, we expected a total of seven main effects (6 univariate main effects of Prior Knowledge Groups regarding six dependent variables and 1 multivariate main effect of Prior Knowledge Groups). However, we found no significant main effects of prior domain-specific knowledge at all.

We hypothesized that the effects of prior domain-specific knowledge would be more pronounced on more complex *Task Levels* (Hypothesis 6). Therefore, we expected a total of seven interaction effects (6 univariate interactions between *Task Levels* and *Prior Knowledge Groups* regarding six dependent variables and 1 multivariate interaction between *Task Levels* and *Prior Knowledge Groups*). We found a significant multivariate interaction between the *Task Levels* and *Prior Knowledge Groups* (F (15,83) = 2.03, $p < .05$, $\eta^2_p = .268$) that was univariately only replicated on the COPES factor *Deep Processing* (F (5,485) = 2.94, $p < .05$, $\eta^2_p = .029$, Figure 3): Biology students judged *Deep Processing* to be of ascending importance from *remember* tasks through *analyze* tasks and their judgments reached a plateau for *analyze, evaluate* and *create* tasks. Humanities students did not discriminate on such a fine-grained level. They judged *Deep Processing* to be quite unimportant for *remember* and *understand* tasks and quite important for all more complex tasks. Furthermore, we found one significant difference in calibration indices (t (100) = 2.09, $p = .039$, $d = .41$): Biology students (calibration: $M = .65$, $SD = .83$) displayed significantly higher calibration indices with regard to *estimated concepts* than humanities students (calibration: $M = .36$, $SD = .56$). To summarize: We found

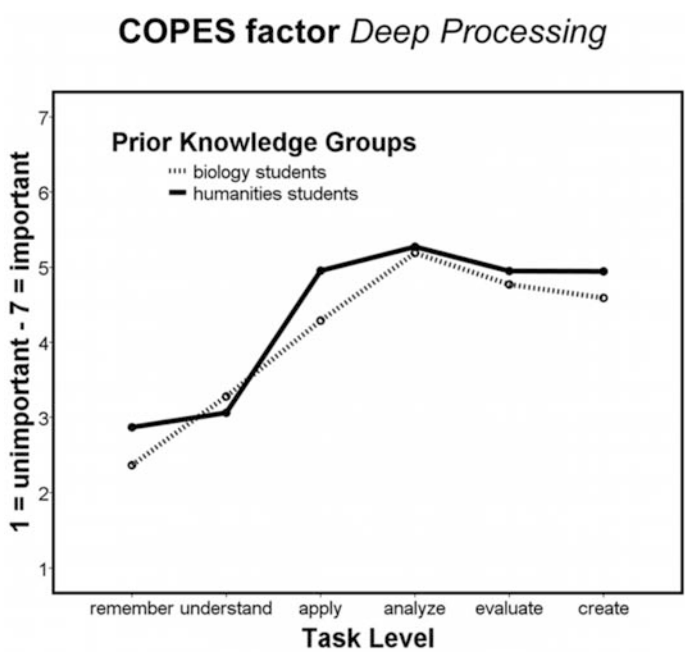


Figure 3. Calibration graph depicting students' judgments on the COPES factor Deep Processing as a function of Task Levels (X-axis) and Prior Domain Knowledge Groups (biology students vs. humanities students; lines).

two effects indicating interactions between prior domain-specific knowledge and task complexity. All other interaction effects and effects on calibration were not significant.

DISCUSSION

Discrimination and Calibration

The empirical data of the present study confirm Hypothesis 1. The repeated-measure factor *Task Levels* elicited significant main effects on *all* dependent variables (*Deep Processing, Multiple Information Sources, Superficial Processing, estimated concepts, estimated time*, and *Bloom classification*). This means that students in fact discriminate between tasks of different complexity as evident in their significantly different answers on the COPES questionnaire.

The empirical data of the present study also confirm Hypothesis 2. *Task Levels* of different complexity were significantly correlated with scores on *all* dependent variables. This means that students in fact calibrate their answers in the COPES questionnaire systematically to task complexity. More specifically, they consider all indicators of deep processing (*Deep Processing, Multiple Information Sources, estimated concepts, estimated time*, and *Bloom classification*) more important for more complex tasks and they consider indicators of superficial processing (*Superficial Processing*) less important for more complex tasks.

Therefore, the results regarding the first two research questions are consistent with the COPES-model (Winne & Hadwin, 1998) that assumes that students systematically adapt their learning process to external conditions. Furthermore, these results are mostly consistent with those of previous empirical studies about task complexity (e.g. Gall, 2006; Klayman, 1985; Rouet, 2003; Winne & Jamieson-Noel, 2003). Most of these studies focused on the *enactment* of learning strategies and indicate that learners demonstrate good self-regulation for simple tasks but less adequate self-regulation for complex tasks. The results of this study are consistent because in all cases learners processed different complex tasks differently and systematically adapted their (planned) behavior to task complexity. However, the results of this study are inconsistent with regard to the quality of students' self-regulation: While results from other studies indicate insufficient self-regulation for complex tasks the results of this study indicate that students are well aware of the special demands of complex tasks and plan to use adequate approaches. One potential explanation for this inconsistency concerns the different stages of learning: Students might be able to *plan* adequate self-regulation based on their adequate metacognitive knowledge about tasks and strategies (this study) but they might be unable to *enact* the planned approaches, for example due to cognitive overload or due to production or motivation deficits (other studies).

One directly related open issue concerns the *absolute* quality of students' calibration. Even though students in general are quite successful at discriminating and calibrating (see above) they might be still far from perfect. Overestimating the complexity of simple tasks might not be detrimental for learning, just not be the most parsimonious way to solve these simple task. Misjudging the complexity of more complex tasks on the other hand might have more detrimental effects. Not only would the answer be less adequate, but also the gained understanding would be more superficial than required. Data from this study (as well as from the corresponding exploratory study; Stahl et al., 2006) tentatively indicates that students might in fact underestimate the complexity of very complex tasks – which would be in line with the finding of less adequate self-regulation for more complex tasks in other studies (see above). For example, the calibration graph depicting students' *absolute calibration* for the item *Bloom classification* indicates that students classify complex tasks into less complex *Task Levels* than warranted (Figure 1). However, this interpretation requires further caution because of our definition of task complexity: Bloom's revised taxonomy (Anderson et al., 2001) assumes a cumulative hierarchy. But empirical results – testing Bloom's original taxonomy (Bloom et al., 1956) – show that the most complex tasks can often not be discriminated with regard to complexity or difficulty (Kreitzer & Madaus, 1994; Kunen, Cohen, & Solman, 1981). On the other hand empirical results strongly support the hierarchical order of less complex tasks, especially for understand, apply, analyze, and create (Gierl, 1997; Kreitzer & Madaus, 1994; Kunen et al., 1981). To conclude this argument: Bloom's revised taxonomy may define task complexity a bit too fine-grained because for some complex levels very similar cognitive processes might be adequate for students. However, this potential problem does not invalidate our conclusion that students in general could – or probably should – consider indicators of deep processing even more important for complex tasks than they currently do – for *analyze* through *create* tasks.

Another issue concerns methodology transfer of calibration measures: Recall the major conceptual differences between traditional conceptualizations of calibration (i.e., accuracy of metacognitive judgments regarding one's own performance) and our conceptualization of calibration (i.e., alignment between students' task definitions, goals and plans and the external variable task complexity). Presumably, learners possess more metacognitive awareness about their own internal cognitive processes (traditional conceptualizations) than about the fit of these processes with the external world (our conceptualizations). Therefore, if we compared these conceptually different indices we would expect higher indices of relative calibration in traditional calibration research. Thus, it is surprising that we detected indices of relative calibration (Goodman-Kruskal Gamma correlations) within this new application context that range from G = .47 to G = .85. The size of these calibration indices would even be considered substantial if compared to calibration indices from the traditional calibration paradigm (e.g. G =.38 for immediate and G = .90 for delayed confidence judgments; Nelson & Dunlosky, 1991).

Epistemological Beliefs

The empirical data of the present study partly confirm Hypothesis 3 (main effects of epistemological beliefs). More sophisticated beliefs in variable knowledge in genetics (*CAEB-variability*) were significantly associated with judging variables indicating deep processing more important (for the dependent variables *Deep Processing, Multiple Information Sources*, and *Bloom classification*). However, no significant effects were detected for other epistemological beliefs scales (*CAEB-texture* and *EBI-definitude*) or other dependent variables (*Superficial Processing, estimated time*, and *estimated concepts*). To summarize: All detected main effects of epistemological beliefs (n = 4) point in the hypothesized direction, but the majority of the hypothesized effects was not significant (n = 17).

Regarding the significant effects, most likely students who believe that knowledge is variable and dynamic automatically consider all kinds of tasks more complex per se (effect on *Bloom classification*). In order to counteract this perceived complexity and in order to adequately deal with the perceived variability of knowledge they might plan deep elaboration approaches (effects on *Deep Processing* and *Multiple Information Sources*). It could be said that these sophisticated students discriminated between tasks on a higher level. These results are in line with other empirical results indicating beneficial main effects of sophisticated beliefs (Bartholomé et al., 2006; Kardash & Scholes, 1996; Mason & Boscolo, 2004; Mason & Scirica, 2006; Schommer, 1990; Schommer, Crouse, & Rhodes, 1992; Schommer-Aikins & Hutter, 2002; Muis, 2007). For example, in other studies concentrating on the *preparatory* stages of learning, students with sophisticated beliefs perceived the affordances of ill-structured tasks more accurately (King & Kitchener, 2002) and set more adequate goals (Bråten & Strømsø, 2004; Ryan, 1984).

However, the number of non-significant main effects, especially regarding other dimensions of epistemological beliefs is surprising. There were no significant (main) effects of connotative beliefs about the structure of knowledge in genetics (*CAEB-texture*) and also of the denotative beliefs about the definitude of knowledge in general (*EBI-definitude*). Possibly beliefs about structural aspects (*CAEB-texture*) of knowledge in genetics have been conceived by our subjects as issues which apply to the field of genetics in general, while issues of variability (*CAEB-variability*) might be more topic-specific and therefore they might have been more important in order to decide how different learning tasks should be tackled differently. Note that in this study epistemological beliefs were also measured in a rather abstract way, especially regarding *EBI-definitude* which was measured for knowledge in general. This also might explain why there were weaker effects than we would have expected based on our predictions. Note, that these explanations may also be relevant for explaining the non-significant interactions between epistemological beliefs and task complexity (see below).

The empirical data of the present study do not confirm Hypothesis 4 (interaction between epistemological beliefs and task complexity). We expected that the effects of epistemological beliefs would be more pronounced regarding more complex *Task Levels*. However, we found two effects that explicitly contradicted this expectation, namely the effects of *CAEB-variability* on *Multiple Information Sources* disappeared for the most complex *Task Levels* and more "naïve" beliefs in definite knowledge (*EBI-definitude*) were associated with higher calibration indices regarding *estimated concepts*. To summarize: All detected interaction or calibration effects of epistemological beliefs (n = 2) point in directions contrary to our hypotheses; the majority of hypothesized effects was not significant (n = 20 interaction effects; n = 17 correlations with calibration indices).

These effects are inconsistent with our predictions as well as with previous research findings. In this study, students with more naïve epistemological beliefs appear to be better at adapting their task definitions, goals and plans to task complexity while sophisticated students showed less flexibility. On the other hand, we assumed theoretically that students with more sophisticated beliefs should be more flexible in their adaptations to task complexity (Hammer & Elby, 2002). Consistent with this theoretical assumption, previous empirical studies investigating the relationship between students' epistemological beliefs and their calibration, found that sophisticated beliefs in gradual learning (*quick learning*, Schommer, 1990) as well as in complex knowledge (*simple knowledge*, Schommer et al., 1992) were associated with less overestimation of comprehension. Furthermore, the corresponding exploratory study from our lab (Stahl et al., 2006) also demonstrated that sophisticated beliefs were associated with better *calibration* indices in the preparatory stages of learning.

One potential explanation for the counterintuitive effects detected in this study is related to the measurement of epistemological beliefs: The scale *EBI-definitude* reaches from views that knowledge is definite (naïve *absolutist*) to views that knowledge is indefinite (sophisticated *relativist*) but does not capture most sophisticated flexible *evaluativist* epistemologies (Kuhn, Cheney, & Weinstock, 2000). Such an evaluativist position with regard to *EBI-definitude* would mean that although knowledge in general is considered indefinite such a person would be aware that some pieces of knowledge are well-validated by scientific inquiry and thus almost absolute answers are attainable in some cases. Most likely, students with such epistemological beliefs would give judgments in the mid-range of the scale *EBI-definitude*. The frequency distribution for *EBI-definitude* reveals that judgments in this sample range from very indefinite conceptualizations (*relativist*) to moderately definite ones (*probably evaluativist*); no very definite judgments were given. Students with moderately definite views on *EBI-definitude* – probably the most sophisticated students according to this proposed explanation – possess higher *calibration* indices than students who considered knowledge very indefinite.

Regarding epistemological beliefs, we conclude that epistemological beliefs elicited fewer effects than predicted, but that our predictions of main effects were correct, at least regarding *CAEB-variability*: Sophisticated beliefs in variable knowledge in genetics were mainly associated with judging indicators of deep processing more important across all tasks. These effects are consistent with our theoretical assumption that epistemological beliefs foster learning because they entail general assumptions about the forthcoming knowledge and task structures which have to be dealt with by the learner. Of course, the reported relationships between epistemological beliefs and task definitions, goals and plans are only correlational. Therefore we conceive the results with some reserve as evidence for our theoretical proposition about epistemological beliefs as standards for the calibration in the preparatory phases of learning as proposed in the COPES model.

Prior Domain-Specific Knowledge

The empirical data of the present study does not confirm Hypothesis 5 (main effects of prior domain-specific knowledge). We expected that biology students with high prior domain-specific knowledge would judge indicators of deep processing more important and indicators of superficial processing less important across all tasks. However, we found none of the six expected main effect of *Prior Knowledge Groups*. The empirical data of the present study show an unexpected pattern regarding Hypothesis 6 (interaction between prior domain-specific knowledge and task complexity). We expected that the effects of prior knowledge would be more pronounced on more complex *Task Levels*. However, we found two effects just indicating more fine-grained and differentiated calibration of biology students. *Prior Knowledge Groups* showed an interaction with *Task Levels* on the COPES factor *Deep Processing* indicating more fine-grained discrimination of biology students. Furthermore, biology students displayed higher calibration indices with regard to *estimated concepts*. To summarize: The detected interaction or calibration effects of prior domain-specific knowledge show more fine-grained discrimination for students with higher domain-specific knowledge – which differs from the predicted pattern of interaction; however the majority of hypothesized effects was not significant.

Prior knowledge might have helped students to perceive more fine-grained nuances of differences in tasks while students without adequate domain-specific knowledge might have based their judgments on surface cues. These results are mostly consistent with those of previous empirical studies demonstrating that prior domain-specific knowledge has little quantitative impact (consistent with the small number of detected effects) but some qualitative impact (consistent with the detected effects) on *planning* processes: Experts seem to use more elaborate criteria to evaluate tasks and seem to judge task difficulty more accurately (Chi, 2006; Lodewyk & Winne, 2005). However, considering the ubiquitous impact of prior domain-specific knowledge on learning processes detected in other empirical studies, prior domain-specific knowledge had surprisingly little impact on students' preparatory stages of self-regulated learning in this study. A potential explanation concerns the domain-specificity versus domain-generality of expertise: Students' *task definitions, goals and plans* might be more dependent on domain-general approaches to learning (e.g. students' metacognitive knowledge about tasks and adequate strategies) than on prior domain-specific knowledge. However, we assume that prior domain-specific knowledge might become more relevant in subsequent stages of learning.

IMPLICATIONS

These results imply that students are able to successfully monitor tasks with regard to complexity and seem to know reasonably well what kind of *task definitions, goals and plans* are adequate. Of course, it cannot be taken for granted that the planning and anticipation processes which were scrutinized here do really result in appropriate learning behaviour. Thus, if students should fail to enact appropriate strategies in the subsequent stages of learning, this should not be attributed to monitoring or knowledge deficits, but rather to production or motivation deficits.

If these findings could be corroborated in further studies it would have some practical implications. If students (at least of this age group) are able to apprehend the complexity of tasks in advance, such capabilities could be used in instruction. In order to make students aware about their pre-existing knowledge and ideas (sometimes also: about their misconceptions) it might be helpful to ask them for reflections about the next tasks, similarly to the procedure with the COPES questionnaire of this study. Asking students why they judge some tasks as less complex than others and asking them what they think about the knowledge laying before them, might be a successful teaching approach just because it can build on the calibration capabilities which became evident in this study. Furthermore they could be asked about their ideas with regard to the nature of the knowledge which they have got to acquire next. While our findings with regard to the relationship between such beliefs and calibration were mixed, they nevertheless allow for the conclusion that thinking about the forthcoming learning tasks involves some epistemological belief aspects. Again such relationships could be made more aware by explicit discussing denotative as well as connotative aspects of students' ideas about the knowledge they have to acquire next. We have shown that even such general associations about the variability of knowledge as they were measured here (with the CAEB) are related to the choice of study strategies. Therefore it should be feasible to use these associations as a topic of instructional discussions.

NOTES

[1] Two versions of this introduction were administered, but because this experimental treatment elicited no significant effects, we ignored this factor subsequently. More specifically, we matched two groups of participants with regard to their prior biology knowledge and their epistemological beliefs based on the results obtained in the first online session. One matched sub-sample read a neutral version of the introduction and the other sub-sample an epistemological version that was enriched with comments about epistemological issues and which was intended to elicit more sophisticated beliefs. As a treatment check the CAEB was re-administered after this epistemological sensitization. However, we found no significant differences in epistemological beliefs after this treatment. Therefore, we ignored this attempted experimental manipulation in all subsequent analyses.

REFERENCE LIST

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et al. (2001). *A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Bartholomé, T., Stahl, E., Pieschl, S., & Bromme, R. (2006).What matters in help-seeking? A study of help effectiveness and learner-related factors. *Computers in Human Behavior, 22*, 113-129.

Bendixen, L. D., & Hartley, K. (2003). Successful learning with hypermedia: The role of epistemological beliefs and metacognitive awareness. *Journal of Educational Computing Research, 28*(1), 15-30.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: Handbook I: Cognitive domain.* New York: David McKay.

Bråten, I., & Strømsø, H. I. (2004). Epistemological beliefs and implicit theories of intelligence as predictors of achievement goals. *Contemporary Educational Psychology, 29*, 371-388.

Bråten, I., Strømsø, H. I., & Samuelstuen, M. S. (2005). The relationship between internet-specific epistemological beliefs within internet technologies. *Journal of Educational Computing Research, 33*(2), 141-171.

Bromme, R., Kienhues, D., & Porsch, T. (2009). Who knows what and who can we believe? Epistemological beliefs are beliefs about knowledge (mostly) attained from others. In L. D. Bendixen, & F. C. Feucht (Eds.), *Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice* (pp. 163-193). Cambridge: Cambridge University Press.

Bromme, R., Kienhues, D., & Stahl, E. (2008).Knowledge and epistemological beliefs: An intimate but complicate relationship. In M. S. Khine (Ed.), *Knowing, Knowledge, and Beliefs: Epistemological Studies Across Diverse Cultures* (p.423 - 444). New York: Springer.

Bromme, R., Pieschl, S., & Stahl, E. (2010). Epistemological beliefs are standards for adaptive learning: A functional theory about epistemological beliefs and metacognition. *Metacognition and Learning*, 5(1), 7-26.

Buehl, M. (2008). Assessing the multidimensionality of students' epistemic beliefs across diverse cultures. In M. S. Khine (Ed.), *Knowing, knowledge and beliefs. Epistemological studies across diverse cultures* (pp. 65-112). New York: Springer.

Buehl, M. M., & Alexander, P. A. (2001). Beliefs about academic knowledge. *Educational Psychology Review, 13*(4), 385-418.

Cano, F. (2005). Epistemological beliefs and approaches to learning: Their change through secondary school and their influence on academic performance. *British Journal of Educational Psychology*, 75, 203-221.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 21-30). New York: Cambridge University Press.

Dahl, T. I., Bals, M., & Turi, A. L. (2005). Are students' beliefs about knowledge and learning associated with their reported use of learning strategies? *British Journal of Educational Psychology, 75*, 257-273.

Elby, A., & Hammer, D. (2001). On the substance of a sophisticated epistemology. *Science Education, 85*(5), 554-567.

Gall, J. (2006). Orienting tasks and their impact on learning and attitudes in the use of hypertext. *Journal of Educational Multimedia and Hypermedia, 15*(1), 5-29.

Gierl, M. J. (1997). Comparing cognitive representations of test developers and students on a mathematics test with Bloom's taxonomy. *The Journal of Educational Research, 91*(1), 26-32.

Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research, 77*(3), 334-372

Greene, J. A., Azevedeo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist, 45*(3), 142-160.

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer, & P. R. Pintrich (Eds.), *Personal epistemology. The psychology of beliefs about knowledge and knowing* (pp. 169-190). Mahwah, NJ: Lawrence Erlbaum Associates.

Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: Thinking aloud during online-searching. *Educational Psychologist, 39*(1), 43-55.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140.

Jacobson, M. J., & Jehng, J.-C. (1999). *Epistemological beliefs instrument: Scales and items*. Retrieved June 20, 2007, from http://mjjacobson.net/publications/Epist_Beliefs_ Instrument98.PDF.

Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research, 12*(4), 301-333.

Kardash, C. M., & Howell, K. L. (2000). Effects of epistemological beliefs and topic-specific beliefs on undergraduates' cognitive and strategic processing of dual-positional text. *Journal of Educational Psychology, 92*(3), 524-535.

Kardash, C. M., & Scholes, R. J. (1996). Effects of preexisting beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology, 88*(2), 260-271.

King, P. M., & Kitchener, K. S. (2002). The reflective judgment model: Twenty years of research on epistemic cognition. In B. K. Hofer, & P. R. Pintrich (Eds.), *Personal epistemology. The psychology of beliefs about knowledge and knowing* (pp. 37-62). Mahwah, NJ: Lawrence Erlbaum Associates.

Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. *Human Development, 26*, 106-116.

Klayman, J. (1985). Children's decision strategies and their adaptation to task characteristics. *Organizational Behavior and Human Decision Processes, 35*, 179-201.

Kreitzer, A. E., & Madaus, G. F. (1994). Empirical investigations of the hierarchical structure of the taxonomy. In L. W. Anderson, & L. A. Sosniak (Eds.), *Bloom's taxonomy: A forty-year retrospective* (pp. 64 - 82). Chicago: University of Chicago Press.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Sciences, 9*(5), 178-181.

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15*(3), 309-328.

Kunen, S., Cohen, R., & Solman, R. (1981). A levels-of-processing analysis of bloom's taxonomy. *Journal of Educational Psychology, 73*(2), 202-211.

Lind, G., & Sandmann, A. (2003). Lernstrategien und Domänenwissen [learning strategies and domain knowledge]. *Zeitschrift für Psychologie, 211*(4), 171-192.

Lodewyk, K. R., & Winne, P. H. (2005). Relations among the structure of learning tasks, achievement, and changes in self-efficacy in secondary students. *Journal of Educational Psychology, 97*(1), 3-12.

Mason, L., & Boscolo, P. (2004). Role of epistemological understanding and interest in interpreting a controversy and in topic-specific belief change. *Contemporary Educational Psychology, 29*, 103-128.

Mason, L., & Scirica, F. (2006). Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learning and Instruction, 16*(5), 492-509.

McDonald, S., & Stevenson, R. J. (1998). Navigation in hyperspace: An evaluation of the effects of navigational tools and subject matter expertise on browsing and information retrieval in hypertext. *Interacting with Computers, 10*, 129-142.

Muis, K. R. (2007). The role of epistemic beliefs in self-regulated learning. *Educational Psychologist*, *42*(3), 173-190.

Neber, H., & Schommer-Aikins, M. (2002). Self-regulated science learning with highly gifted students: The role of cognitive, motivational, epistemological, and environmental variables. *High Ability Studies, 13*(1), 59-74.

Nelson, T. O. (1996). Gamma is a measure of the accuracy of predicting performance on one item relative to another item, not of the absolute performance on an individual item. Comments on Schraw (1995). *Applied Cognitive Psychology, 10*, 257-260.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgements of learning (JOLs) are extremely accurate at predicting subsequent recall: the "delayed-JOL effect". *Psychological Science, 2*(4), 267-270.

Niessen, T.,Vermunt, J. D., Abma, T., Widdershoven, G., & van der Vleuten, C. (2004). On the nature and form of epistemologies: Revealing hidden assumptions through an analysis of instrument design. *European Journal of School Psychology, 2*(1-2), 39-64.

Pieschl, S. (2009). *Metacognitive calibration – An extended conceptualization and potential applications. Metacognition and Learning*, *4(1), 3-31.*

Pieschl, S., Stahl, E., & Bromme, R. (2013). Adaptation to context as core component of self-regulated learning. The example of complexity and epistemic beliefs. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies*, Springer International Handbooks of Education 26 (pp. 53-65). New York: Springer.

Rouet, J.-F. (2003). What was I looking for? The influence of task specificity and prior knowledge on students' search strategies in hypertext. *Interacting with Computers, 15*, 409-428.

Rouet, J.-F., Favart, M., Britt, M. A., & Perfetti, C. A. (1997). Studying and using multiple documents in history: Effects of discipline expertise. *Cognition and Instruction, 15*(1), 85-106.

Ryan, M. P. (1984). Monitoring text comprehension: Individual differences in epistemological standards. *Journal of Educational Psychology, 76*(2), 248-258.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*(3), 498-504.

Schommer, M. (1993). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology, 85*(3), 406-411.

Schommer, M., Crouse, A., & Rhodes, N. (1992). Epistemological beliefs and mathematical text comprehension: Believing it is simple doesn't make it so. *Journal of Educational Psychology, 84*(4), 435-443.

Schommer-Aikins, M., & Hutter, R. (2002). Epistemological beliefs and thinking about everyday controversial issues. *The Journal of Psychology, 136*(1), 5-20.

Schraw, G. (1995). Measures of feeling-of-knowing accuracy: A new look at an old problem. *Applied Cognitive Psychology, 9*, 321-332.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review, 7*(4), 351-371.

Stahl, E., & Bromme, R. (2007). CAEB. An instrument to measure connotative aspects of epistemological beliefs. *Learning and Instruction, 17*(6)*,* 773-785.

Stahl, E., Pieschl, S., & Bromme, R. (2006).Task complexity, epistemological beliefs and metacognitive calibration: An exploratory study. *Journal of Educational Computing Research, 35*(4), 319-338.

Strømsø, H. I., Bråten, I., & Samuelstuen, M. S. (2008). Dimensions of topic-specific epistemological beliefs as predictors of multiple text understanding. *Learning and Instruction*, *18*(6), 513-527.

Tu, Y.-W., Shih, M., & Tsai, C.-C. (2008). Eighth graders' web searching strategies and outcomes : The role of task type, web experience and epistemological beliefs. *Computers & Education*, *51*(3), 1142-1153.

Weingardt, K. R., Leonesio, R. J., & Loftus, E. F. (1994). Viewing eyewitness research from a metacognitive perspective. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition, knowing about knowing* (pp. 157-184). Cambridge: MIT Press.

Weinstock, M., & Cronin, M. A. (2003). The everyday production of knowledge: Individual differences in epistemological understanding and juror-reasoning skill. *Applied Cognitive Psychology*, 17, 161-181.

Windschitl, M., & Andre, T. (1998). Using computer simulations to enhance conceptual change: The roles of constructivist instruction and student epistemological beliefs. *Journal of Research in Science Teaching, 35*(2), 145-160.

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

Winne, P. H., & Jamieson-Noel, D. (2003).Self-regulating studying by objectives for learning: Students' reports compared to a model. *Contemporary Educational Psychology*, 28, 259-276.

AFFILIATIONS AND ACKNOWLEDGEMENT

Rainer Bromme
Department of Psychology,
University of Muenster

StephaniePieschl
Department of Psychology,
University of Muenster

Elmar Stahl
University of Education, Freiburg

This study was funded by the German Research Foundation (DFG, BR 1126/3-1). We thank our student research assistants of this project (in alphabetical order): Stephanie Adrian, Sabine Joachim, Mike Pillukat, and TorstenPorsch. The data in this paper were collected and analyzed as partial fulfillment of the second author's dissertation requirements.