What Makes Learners Overestimate Their Text Comprehension? The Impact of Learner Characteristics on Judgment Bias

Stefanie Golke[1] · Tim Steininger[1] · Jörg Wittwer[1]

Abstract

It is frequently assumed that learner characteristics (e.g., reading skill, self-perceptions, optimism) account for overestimations of text comprehension, which threaten learning success. However, previous findings are heterogenous. To circumvent a key problem of previous research, we considered cognitive, metacognitive, motivational, and personality characteristics of learners ($N = 255$) simultaneously with regard to their impact on the judgment biases in prediction and postdiction judgments about factual and inference questions. The main results for the factual questions showed that men, lower reading skill, working memory capacity, and topic knowledge, yet higher self-perceptions of cognitive and metacognitive capacities yielded stronger overestimations for prediction judgments. For inference questions, a lower reading skill, higher self-perceptions of metacognitive capacities, and a higher self-efficacy were related to stronger overestimations for prediction and postdiction judgments. A higher openness was a risk factor for stronger overestimations when making predictions for the inference questions. The findings demonstrate that learner characteristics are a relevant source of judgment bias, which should be incorporated explicitly in theories of judgment accuracy. At the same time, fewer learner characteristics were actually relevant than previous research suggests. Moreover, which learner characteristics impact judgment bias also depends on task requirements, such as factual versus inference questions.

Keywords Learner characteristics · Individual differences · Metacomprehension accuracy · Judgment bias · Overestimation · Underestimation

---

✉ Stefanie Golke
stefanie.golke@ezw.uni-freiburg.de

1   Department of Education, University of Freiburg, Freiburg, Germany

Learners are usually poor at accurately monitoring and judging their text comprehension (Maki & McGuire, 2002; Prinz et al., 2020a; Thiede et al., 2009). This poor accuracy, often reflected in overestimations but also in underestimations of comprehension, hampers learning (Dunlosky & Rawson, 2012). Therefore, knowledge of what causes poor metacomprehension accuracy is of major importance. In this context, characteristics of the learners (e.g., reading skill, self-beliefs, optimism) have repeatedly been in the focus of attention, which seems intuitively appealing. However, previous findings are heterogeneous and not overly robust (cf. Lin & Zabrucky, 1998). Methodological issues probably contributed to this lack of clarity. For example, studies have mostly been restricted to the examination of one or two learner characteristics simultaneously. Hence, interdependencies between (psychologically similar) learner characteristics across studies have mainly been disregarded. To circumvent these shortcomings and to gain reliable insights, the present study included, based on a literature review, multiple potentially relevant learner characteristics and determined their impact on metacomprehension accuracy when analyzed simultaneously. As accuracy measure, we investigated judgment bias (overestimation/underestimation) because individual differences are most likely to manifest in this measure (cf. Maki et al., 2005).

Metacomprehension Accuracy and Its Relevance for Learning

Successful learning from text is based on text comprehension and metacomprehension accuracy. Text comprehension is a complex process of constructing and mentally representing the meaning of a text. To this end, learners are required to extract the main idea units of the text, build connections (i.e., inferences) between text information within and across sentences and between text information and prior knowledge (van den Broek et al., 2002). Moreover, learners need to accurately monitor these cognitive processes and judge their level of comprehension, which is known as metacomprehension accuracy (Maki & McGuire, 2002). This metacognitive process is secondary to the cognitive processes. Thus, the less effectively and routinely learners carry out the cognitive processes during reading, the less cognitive capacity is available for metacomprehension (Nelson & Narens, 1990). Hence, factors that hinder the cognitive processes also hamper metacomprehension accuracy. These factors include a learner's skills and resources (e.g., reading skill, prior knowledge, reading motivation), text characteristics (e.g., text difficulty), and the reading situation (e.g., reading for comprehension vs. memory of facts; Zwaan & Rapp, 2006; van den Broek et al., 2002).

Metacomprehension accuracy is particularly relevant for learning because a higher compared to a lower accuracy enables a more effective regulation (e.g., rereading) and, hence, a higher text comprehension (Thiede et al., 2009). For example, overestimation entails that learners likely overlook underdeveloped aspects of their text representation, and hence, their efforts in regulating their understanding will very likely be insufficient (e.g., skip rereading, miss relevant parts during

rereading). Accordingly, poor accuracy and overestimation in particular, which are a widespread phenomenon (Lin & Zabrucky, 1998; Maki & McGuire, 2002; Prinz et al., 2020a; Thiede et al., 2009), have been shown to result in underachievement (Dunlosky & Rawson, 2012).

The Role of Learner Characteristics for Metacomprehension Accuracy

According to the cue-utilization framework (Koriat, 1997), metacomprehension accuracy depends on the cues that learners use to infer their level of comprehension because cues differ in their validity in indicating actual understanding. For metacomprehension, representation-based and heuristic cues are differentiated for prediction judgments (i.e., judgments made after reading but before answering comprehension questions; cf. Griffin et al., 2009). Representation-based cues emerge from monitoring the text processing (e.g., the coherence of text representation, ease of explanation, memorability of information). They are, hence, closely related to and rather valid indicators of a text representation. Yet, monitoring becomes less likely, the less effectively and routinely learners process a text at the cognitive level. Thus, regarding learner characteristics, poor cognitive skills and resources (i.e., reading skill, working memory capacity, prior knowledge), and reading-related motivational factors (e.g., self-efficacy) can be assumed to hinder metacomprehension accuracy.

In contrast, heuristic cues are available to learners whether or not a text has been read. These are factors that learners can access and believe to be relevant for judgments, including, with regard to learner characteristics, self-perceptions of cognitive or metacognitive capacities and motivational beliefs (e.g., self-perceived reading skill, self-assessed prior knowledge, self-efficacy). Learners can draw on these cues directly to make a judgment. For example, a high self-perceived reading skill might induce a learner to give a high judgment of comprehension, even though it might be an overestimation for the actual text. Therefore, heuristic cues are usually less valid than representation-based cues (Griffin et al., 2009).

Compared to prediction judgments, postdictions are judgments that learners make after having answered comprehension questions. This test experience provides additional judgment cues (e.g., feeling of confidence during answering, number and similarity of response alternatives). Apart from that, postdictions can theoretically be influenced by the same learner characteristics and cues as predictions.

Some of the learner characteristics investigated already, namely personality facets such as extraversion or optimism and gender (e.g., Agler et al., 2019; Buratti et al., 2013; Dahl et al., 2010; de Bruin et al., 2017), hardly influence text comprehension or appear to be less accessible to learners. These characteristics are, thus, conceptually distinct from representation-based and heuristic cues. Yet, they might involve bias-inducing cognitive styles, which are, nevertheless, compatible with the cue-utilization framework.

Findings on the Impact of Learner Characteristics on Metacomprehension Accuracy

As various learner characteristics may impact judgment accuracy, the focus in research has been broad: Prior studies addressed cognitive and metacognitive skills and resources, motivation, personality, and gender. In these studies, accuracy was determined as judgment bias (i.e., the signed difference between judgments and performance; overestimation/underestimation), absolute accuracy (i.e., the absolute deviation of judgments and performance), and/or relative accuracy (i.e., the differentiation between well and less well understood texts).

In the present study, we investigated judgment bias because it shows a good stability across tasks and over time (Kelemen et al., 2000) as opposed to relative accuracy. Hence, stable effects of individual differences are more likely to be detected for judgment bias (cf. Maki et al., 2005). The number of studies in this field is, however, limited. Thus, findings on absolute and relative accuracy can nevertheless provide helpful insights, albeit generalizations of effects across the different measures have to be drawn with caution. In the following overview of previous research, we therefore focused on findings about learner characteristics' effects on judgment bias, but also included some findings on absolute and relative accuracy. We also refer to studies that not strictly examined judgment accuracy for texts (i.e., metacomprehension accuracy) but some kind of written material (e.g., problem tasks) or knowledge tests. A summary of the reviewed previous findings is provided in Appendix 1 Table 5.

When interpreting effects of judgment bias, it needs to be considered that bias as a derived score can depend on differences in judgment magnitude and performance. This implies in its extreme that differences in bias between, for example, two groups of learners can completely result from performance differences (i.e., same judgment, yet higher/lower performance). In this case, it can be questioned whether the effect of bias results from metacognitive effects (Griffin et al., 2013, p. 26). Hence, effects of learner characteristics on bias should be interpreted alongside their effects on judgment magnitude and performance. This detailed information is often missing in previous studies though (see Appendix 1 Table 5).

Cognitive and Metacognitive Skills and Resources

For adults, *reading skill* mainly manifests in the proficiency in automatically conducting and strategically regulating the multiple cognitive processes involved in reading at the text level. Thus, a lower compared to a higher reading skill requires more working memory capacity for cognitive processes and therefore hampers monitoring (van den Broek et al., 2002; Veenman, 2016). Accordingly, Maki et al. (2005) and Golke and Wittwer (2017) showed, for texts with an age-appropriate text difficulty, that a lower reading skill leads to more overestimation of comprehension when making predictions, whereas higher reading skill results in less overestimation or rather accurate predictions. For postdictions, which are usually lower than predictions (Pierce & Smith, 2001), the extent of overestimation for lower reading skill decreased, while the high-skilled learners floated into underestimation. However,

the negative impact of a low reading skill on judgment bias might be mitigated by rereading texts as findings from Griffin et al. (2008) suggest.

Due to the resource-demanding processes of text comprehension, constraints in *working memory capacity* (WMC) can also account for poor monitoring. Accordingly, a lower compared to a higher WMC has been shown to hamper metacomprehension accuracy regarding relative accuracy for predictions after one-trial reading of a text (Griffin et al., 2008; Ikeda & Kitagami, 2012). Komori (2016) showed, though for word recall, that a lower compared to a higher WMC arrived at poorer absolute accuracy for postdictions, while WMC was irrelevant for judgment bias. Moreover, a lower compared to a higher WMC is linked to less attention control (Unsworth & McMillan, 2013), which contributes to poor monitoring. However, rereading a text can alleviate or, as Griffin et al. unveiled, eliminate the disadvantage of a low WMC.

Since monitoring as a metacognitive skill is a higher-order process, it has also been connected with *fluid intelligence* (Ohtani & Hisasaka, 2018; van der Stel & Veenman, 2014; Veenman et al., 2005). A higher compared to a lower fluid intelligence helps to process information more rapidly and carefully and to develop strategies prior to task completion (Sternberg, 1986). However, findings on its relationship with monitoring (not strictly metacomprehension) accuracy are very heterogenous. Ohtani and Hisasaka's (2018) meta-analysis yielded a small average effect. As it is based on effects aggregated across domains, accuracy measures, and predictions and postdictions, it has to be interpreted cautiously. Besides, reasoning ability is less relevant for text comprehension (Corso et al., 2016). It is, hence, unclear whether fluid intelligence has a substantial effect on metacomprehension accuracy, in particular when the conceptually related WMC (cf. Oberauer et al., 2005) is also considered.

Another potential impact on metacomprehension accuracy, directly or via its influence on text comprehension, is *prior knowledge*. It includes topic knowledge that enables cognitive processes of text comprehension (e.g., drawing inferences; Kendeou & O'Brien, 2016) and metacognitive knowledge about, for instance, reading strategies, comprehension standards, or monitoring activities (Veenman, 2016). Prior knowledge supports text comprehension (Griffin et al., 2009; Kendeou & O'Brien, 2016; Zabrucky et al., 2015), and it can partially compensate for lower WMC, reading skill, and general cognitive abilities (e.g., Hambrick & Engle, 2002; Miller et al., 2006). Accordingly, a lower compared to a higher topic knowledge has been linked to less accurate predictions in terms of absolute accuracy (Griffin et al., 2009) and relative accuracy (de Bruin et al., 2007). Griffin et al. also showed that it led to stronger underestimation but decreased and arrived at rather accurate judgments for higher topic knowledge. However, participants in Griffin et al.'s study, on average, produced underestimation, not the commonly found overestimation, which might be due to specifics of the study's material (e.g., detailed baseball knowledge for a psychology student sample). For materials that provoke overestimation, theoretically, lower prior knowledge might lead to stronger overestimations as it impairs text comprehension.

For postdiction accuracy, it is not topic knowledge that seems relevant (Schraw, 1997) but metacognitive knowledge (cf. Lin & Zabrucky, 1998). More specifically, Schraw (1997) unveiled that lower knowledge about monitoring strategies for taking tests was related to underestimation of the correctness of one's responses while

higher knowledge was linked to rather accurate judgments. Contrarily, metacognitive knowledge in its broad sense (i.e., declarative/procedural/conditional knowledge of cognition and knowledge about regulation) and with regard to learning situations in general instead of test taking situations was unrelated to judgment bias for postdictions (Schraw & Dennison, 1994). For prediction accuracy, evidence for the impact of metacognitive knowledge is scanty (Lin & Zabrucky, 1998).

In addition to the objective amount of prior (topic) knowledge, learners have been found to draw on their *self-perceived topic knowledge*, which reflects the familiarity with a domain or topic, when making predictions and postdictions (Ehrlinger & Dunning, 2003, Study 3; Glenberg et al., 1987; Griffin et al., 2009). As it is only a distal indicator of topic knowledge and not necessarily related to the actual text understanding, a higher self-perceived topic knowledge can theoretically be assumed to provoke overestimation. Findings from Ehrlinger and Dunning support this view. Contrarily, Griffin et al. observed that higher self-perceived topic knowledge resulted in less underestimating, rather accurate judgments, but the overall underestimation in their study limits the transferability of this finding. Generally, learners seem to make only limited use of domain/topic familiarity as judgment cue for text comprehension though (cf. Lin & Zabrucky, 1998).

Moreover, *epistemological beliefs* as a specific component of metacognitive knowledge (Winne & Hadwin, 1998) have been in the focus of interest. They are implicit beliefs about the nature of knowledge and the process of knowing. More sophisticated beliefs (i.e., knowledge is complex, uncertain, justified by evidence, and gained by rational inquiry) are related to more adaptive processing and hence a higher text comprehension than less sophisticated beliefs (i.e., knowledge is simple, certain, reflecting reality, and transmitted by authorities; Bråten et al., 2016). In line with this notion, prior studies found that more sophisticated epistemological beliefs resulted in better strategy selection during learning which has been interpreted as an indicator of enhanced monitoring (e.g., Mason et al., 2010; Stahl et al., 2006). Whether these findings allow to conclude that epistemological beliefs impact metacomprehension accuracy is up for debate though (Griffin et al., 2013). Thus, epistemological beliefs might theoretically impact metacomprehension accuracy in the way that more sophisticated beliefs support text comprehension and, hence, the generation of more representation-based cues as opposed to less sophisticated beliefs.

Taken together, empirical findings support the general view that low reading skill, WMC, and prior knowledge are related with poor metacomprehension accuracy. Specifically, a lower reading skill has led to stronger overestimations (for predictions and, to a smaller extent, for postdictions). Lower metacognitive knowledge about monitoring has been linked to underestimations for postdictions. Lower topic knowledge has been related to stronger underestimation for predictions, yet due to specifics of the prior research, it is unclear whether underestimation (or overestimation instead) would also emerge for other material. The impact of WMC and fluid intelligence (reasoning) on judgment bias for text comprehension is understudied. Theoretically, a lower WMC might more likely lead to overestimation, but the effect might be absent after rereading of texts. Lower fluid intelligence might weakly contribute to overestimations, if at all, especially when the other cognitive factors are considered simultaneously.

Motivation

Motivational characteristics can impact metacomprehension accuracy for two reasons. First, as text comprehension is effortful, especially for longer and difficult texts, it depends on learners' motivation (Guthrie & Klauda, 2016). It influences the constructed level of text representation that provides more or less representation-based cues. Second, motivation in terms of generalized self-perceptions of ability is available as heuristic cues. It is argued that learners draw on these self-representations because they are rather stable across texts and, hence, enable a less effortful and resource-demanding judgment as opposed to a situation-specific integration of text-, learner-, test-, and environment-related cues (Moore et al., 2005). Prior research on metacomprehension accuracy addressed self-perceptions of reading skill and monitoring ability, self-efficacy beliefs, and goal orientation.

Regarding *self-perceived reading skill*, Kwon and Linderholm (2014) found that, under control of actual reading skill, learners with a higher self-perceived reading skill gave higher predictions and postdictions of text comprehension than those with a lower self-perceived reading skill. This finding suggests, due to the general tendency toward overestimation, that a higher compared to a lower self-perceived reading skill bears a higher risk to overestimate one's text comprehension.

Additionally, the *self-perceived metacomprehension ability*, which is positively and moderately related to actual metacognitive knowledge (Schraw & Dennison, 1994), can also be relevant for the accuracy. Learners who ascribed themselves a low general metacomprehension ability showed stronger overestimations (Schraw & Dennison, 1994, Study 2) and poorer absolute accuracy (Schraw, 1994) for item-level postdictions than those with a high self-perceived metacomprehension ability.

A special role among the learner characteristics might be seen in *self-efficacy*. It refers to learners' beliefs about how capable they are to perform a certain upcoming task. Thus, comprehension/performance judgments and their accuracy have been construed as one dimension of self-efficacy (Zimmerman & Moylan, 2009). In line with this notion, Vössing and Stamov-Roßnagel (2016) found that lower self-efficacy beliefs for learning tasks (e.g., understanding ideas taught in a course/presented in a text) were related to lower predictions of comprehension than higher self-efficacy beliefs. Bouffard-Bouchard (1990) also observed lower item-level postdictions for problem solving tasks in a low self-efficacy group compared to a high self-efficacy group. Their data furthermore showed that judgments were mostly overconfident in the high self-efficacy group and rather equally often overconfident and underconfident in the low self-efficacy group.

Goal orientation, that is, the disposition to pursue a specific purpose when engaging with a task and to apply certain standards for evaluating task performance, has been assumed to extend to judgment accuracy (Ikeda et al., 2016; Kroll & Ford, 1992; Zhou, 2013). More specifically, performance orientation entails a learner's focus on their achievements relative to others. This can be expressed in displaying their abilities to project a positive self-image (i.e., performance-approach), which is believed to invite overestimations, or in hiding potential failures and deficits (i.e., performance-avoidance), which is assumed to result in less overestimation or, as self-defense, in underestimation. In contrast, mastery orientation is to strive for increasing one's knowledge and abilities and to use an intraindividual evaluation

of one's achievements. This should lead to more accurate judgments. Prior studies provide some support for these views. Learners with performance approach have been shown to give higher predictions than mastery-approach learners, while their performance on a text comprehension test was equal (Ikeda et al., 2016, Exp. 3). Moreover, stronger overestimations on predictions have been found for higher performance-approach and performance-avoidance orientations (Kroll & Ford, 1992; Zhou, 2013), which might, however, have mainly been driven by performance differences in Kroll and Ford's study (no details available in Zhou's study). For mastery orientation in general, no association with accuracy was evident (Kroll & Ford, 1992; Zhou, 2013), albeit Zhou who also distinguished the less commonly used mastery-avoidance orientation found that it yielded more overestimation. No significant relationships were found between goal orientations and postdiction accuracy (Zhou, 2013).

In conclusion, low self-perceived metacomprehension ability has been linked to stronger overestimations for item-level postdictions; research for predictions is missing. For self-efficacy, findings suggest that higher compared to lower self-efficacy might be related to stronger overestimations, but evidence is still limited. Furthermore, for self-perceived reading skill, it can be speculated that, when reading skill is equal, a higher compared to a lower self-perceived reading skill yields stronger overestimation. With respect to goal orientations, a higher performance orientation has been linked to stronger overestimation for predictions; for postdictions, no relationship has been found.

Personality

Although not explicitly discussed in theoretical models on metacomprehension accuracy, personality traits might function as bias-inducing cognitive styles. Previous research investigated the Big-Five personality facets, optimism, and narcissism. The latter is, however, not considered in the present study as nonclinical samples show rather low values of narcissism (i.e., possible bottom effects).

Personality is widely defined by the interrelated *Big-Five facets* extraversion, openness to experience, neuroticism, conscientiousness, and agreeableness (Rammstedt & John, 2005).[1] For metacomprehension, Agler et al. (2019) investigated relative accuracy and found that only a higher openness was linked to poorer accuracy of predictions. This effect became insignificant, however, when accuracy was regressed on all Big-Five facets simultaneously. No reliable relationships occurred for postdictions.

In addition, when judgment accuracy for postdictions in knowledge and ability tests (not text comprehension) was focused, a higher extraversion was rather consistently related to stronger overestimations (Buratti et al., 2013; Dahl et al., 2010; Pallier et al., 2002; Schaefer et al., 2004), but a relationship was not always evident (Händel et al., 2020). In addition, a higher openness was also found to be related

---

[1] The facets extraversion, openness to experience, conscientiousness, and agreeableness are positively interrelated, and neuroticism is negatively related to them (van der Linden et al., 2010).

to stronger overestimations in Buratti et al.'s study, but not in others (Dahl et al., 2010; Händel et al., 2020; Schaefer et al., 2004). Given that extraversion, which is associated with openness, overlaps with a more positive self-esteem and higher optimism (Dahl et al., 2010; Sharpe et al., 2011), these characteristics combined might result in overly confident judgments. Moreover, a higher conscientiousness has been related with more overestimation in some studies (Dahl et al., 2010; Händel et al., 2020, Study 1, for psychology exam), but not in others (Händel et al., 2020, Study 1, for math exam; Schaefer et al., 2004). The reported relationships were mostly small to moderate. Regarding their relationships with other learner characteristics, the Big-Five are most likely linked to self-efficacy and goal orientation (Judge & Ilies, 2002; Zweig & Webster, 2004), not to text comprehension (Agler et al., 2019).[2]

Dispositional optimism is the tendency to believe in positive outcomes in life (Sharpe et al., 2011) that becomes unrealistic when one expects more positive outcomes than could possibly be true (Shepperd et al., 2013). Thus, learners with a higher dispositional optimism might be particularly prone to overestimate their understanding (de Bruin et al., 2017; Kleitman et al., 2019). Händel et al. (2020) observed this type of relationship for performance judgments in a math test but not in a psychology test. De Bruin et al. (2017) also found no reliable connection to overestimation when learners predicted their exam score. Whether optimism impacts the more complex process of metacomprehension is unclear. Regarding its relationship with other learner characteristics, optimism appears to be unrelated with academic performance in general (Rand et al., 2011), positively linked to self-efficacy (Alarcon et al., 2013; Chemers et al., 2001), and presumably also to mastery and performance-approach goal orientation (Carver & Scheier, 2014).

Another factor that might contribute to overestimations is an individual's tendency to present themselves overly positive, as assessed via the tendency to *overclaim* knowledge (cf. Bensch et al., 2019; Bing et al., 2011). Overclaiming has been linked to overconfidence in the correctness of one's responses to a knowledge test (Bensch et al., 2019), yet its relationship with metacomprehension accuracy is unexplored. It is thinkable that a tendency toward a positive self-presentation explains some amount of overestimation which, interestingly, would not be attributable to poor metacomprehension during reading but to a subsequent (more or less conscious) adjustment when making the comprehension judgments.

In conclusion, evidence for an impact of personality (i.e., the Big-Five and optimism) on judgment accuracy is very mixed. For metacomprehension accuracy, no reliable effects of the Big-Five have been shown yet and studies on optimism are missing. As a currently unexplored facet, overclaiming as a self-enhancement tendency might further explain why some learners are more prone to overestimations than others.

---

[2] Findings on a weak relationship of fluid intelligence with the Big-Five probably reflect the impact of personality on test performance (e.g., higher cortical arousal of extroverts benefits performance on timed tasks), not an actual connection between both characteristics (Moutafi et al., 2006).

Gender

Although males are observed to be more overconfident than females in various domains (Lundeberg & Mohan, 2009), findings in the context of learning are mixed. Some findings support a weak gender effect (e.g., Buratti et al., 2013; Händel et al., 2020) and others discourage it (e.g., Ariel et al., 2018; Dahl et al., 2010; de Bruin et al., 2017; Händel et al., 2020; see Appendix 1 Table 5 for details). For metacomprehension accuracy, Stankov and Lee (2008) observed a small effect as males more likely overestimated the correctness of their responses (i.e., item-level postdictions) to text comprehension questions than females.

Importantly, although men might slightly outperform women in text inferencing, no effect of gender was found for text memory, knowledge integration, knowledge access (Hannon, 2014), and composite measures of text comprehension (Hannon, 2014; Hyde & McKinley, 1997). Moreover, males compared with females report to be less extraverted, open, agreeable, and neurotic (Moutafi et al., 2006); to adopt a weaker mastery but stronger performance goal orientation (D'Lima et al., 2014); and to have a lower self-perceived reading skill even as young adults (Marsh, 1989) and less sophisticated epistemological beliefs (Bendixen et al., 1998; King & Magun-Jackson, 2009).

Limitations of Previous Research and Open Questions

Which learner characteristics actually contribute to metacomprehension accuracy in terms of overestimation and underestimation of text comprehension cannot yet be answered clearly. Reasons are a scarcity of research for many of the learner characteristics and methodological discrepancies between the studies. More importantly though, most studies investigated only one, sometimes two learner characteristics simultaneously. This neglects the relationships between the various potentially relevant learner characteristics. Hence, findings from previous studies are insufficient to comprehensively explain the actual impact of learner characteristics on judgment bias. For example, a learner characteristic's relationship to bias might be substantial in terms of a bivariate correlation, but can become irrelevant when other characteristics are considered concurrently, or the impact of a learner characteristic on bias might be obscured (in bivariate data) until other learner characteristics are included. Both has been observed in Händel et al.'s (2020) study that is so far an exception in the previous research as they investigated the impact of multiple characteristics (gender, performance, motivation, and personality) on the judgment accuracy for knowledge tests in math and psychology. Yet, their findings are not fully transferable to metacomprehension accuracy as it is based on other cognitive and metacognitive processes and cues specific to text comprehension. To circumvent these shortcomings and to gain more reliable insights into the role of learner characteristics for metacomprehension accuracy, studies are needed that include several learner characteristics simultaneously.

Present Study

To address a major shortcoming of previous research, we identified, based on the literature review, multiple potentially relevant learner characteristics and examined their impact on metacomprehension accuracy in terms of judgment bias (overestimations/underestimations) of text comprehension. To do so, we used expository texts with highly similar features such as text difficulty because text features can also add to judgment bias (cf. Lin & Zabrucky, 1998), which was, however, not in the scope of the present study. We assessed accuracy with respect to two central comprehension processes: extracting main idea units (i.e., factual questions) and building semantic connections between and across sentences (i.e., inference questions). We did not assume differential effects of the learner characteristics on judgment bias for predictions versus postdictions (unless specified otherwise) and for factual versus inference questions as neither theory nor previous findings currently support this view.

We made the following assumptions about the impact of learner characteristics on judgment bias. For the cognitive factors, we expected in line with previous research (Golke & Wittwer 2017; Maki et al., 2005; for texts with age-appropriate text difficulty) that a lower reading skill would be associated with stronger overestimation. A lower WMC should either be related to stronger overestimation (cf. Griffin et al., 2008; Ikeda & Kitagami, 2012; Komori, 2016) or be irrelevant due to the rereading of texts (cf. Griffin et al., 2008) which was allowed in the present study. Moreover, a lower fluid intelligence might be related to stronger overestimation. However, previous findings are heterogeneous (Ohtani & Hisasaka, 2018), and an effect of fluid intelligence might be absent when the conceptually related WMC is considered as well. Regarding the knowledge-related learner characteristics, (a) we made no directed hypothesis about the impact of topic knowledge. Theoretical considerations suggested that a lower prior knowledge would result in stronger overestimations. Contrarily, Griffin et al. (2009) found stronger underestimation for lower topic knowledge, yet the specifics of their study limited the transferability of their finding to the present study. This inconclusive situation was basically the same for (b) self-perceived topic knowledge (topic familiarity). Due to Ehrlinger and Dunning's (2003, Study 3) support for the theoretically consistent assumption that a higher familiarity results in stronger overestimations, we adopted it for the present study. For (c) metacognitive knowledge about monitoring strategies, lower levels might result in stronger underestimation (Schraw, 1997) but probably only for postdiction judgments.

For the motivational factors, the few findings suggested that stronger overestimation can be expected for a higher self-perceived reading skill (Kwon & Linderholm, 2014), a lower self-perceived metacomprehension ability (Schraw, 1994), and, albeit for prediction judgments only, for a higher performance-approach and performance-avoidance goal orientation (Ikeda et al., 2016, Exp. 3; Kroll & Ford, 1992; Zhou, 2013). Moreover, we assumed that a higher self-efficacy for text comprehension would be associated with stronger overestimations while lower self-efficacy would be linked to less overestimation or to underestimation (Bouffard-Bouchard, 1990; Vössing & Stamov-Roßnagel, 2016).

Furthermore, we expected that the impact of personality on metacomprehension accuracy would be weak or negligible when the various personality facets are considered simultaneously (Agler et al., 2019). If, nevertheless, an effect occurred, it should be primarily a higher extraversion and/or openness that resulted in stronger overestimation (Buratti et al., 2013; Dahl et al., 2010; Händel et al., 2020, for psychology exam; Pallier et al., 2002; Schaefer et al., 2004). With respect to dispositional optimism, previous findings are ambivalent (de Bruin et al., 2017; Händel et al., 2020). Hence, if at all, we expected that a higher dispositional optimism would be related to stronger overestimation. We also assumed that a higher overclaiming is linked to a stronger overestimation (Bensch et al., 2019). Concerning gender, we expected that men would show a somewhat stronger overestimation than women (Stankov & Lee, 2008).

It is noteworthy that the assumptions mainly reflect findings on bivariate relationships. We did not expect that every effect should be present when all learner characteristics were considered concurrently. On the contrary, given the relationships between various learner characteristics within and across cognitive, motivational, and personality categories, including gender, we assumed that some of the hypothesized effects reflecting bivariate relationships would be diminished or mitigated then. Moreover, it has been shown that learners use indeed representation-based cues instead of solely drawing on heuristic cues (Lin & Zabrucky, 1998). Therefore, those learner characteristics that can directly impact text comprehension and hence representation-based cues might be more likely to play a role for judgment bias than motivational and personality factors which are more distal to text comprehension.

Method

Sample

In total, 270 students from various study programs (mainly educational science, teacher education, psychology, linguistics, business administration, law studies, natural sciences, and social sciences) participated in the first of two sessions. As 15 students missed the second session, the final sample was $N = 255$. Their mean age was 22.26 ($SD = 3.65$, $Mdn = 21.00$) years, 173 participants (68%) were female, and 233 participants (91%) were native speakers. Nonnatives were reasonably proficient in German due to entry requirements of their study programs. Moreover, 109 participants (43%) were first-semester students ($Mdn = 2$ semester). Students participated voluntarily and received money or course-credit for it.

Procedure

Participants attended two sessions (1 week apart) in groups of six or less. In session 1 (approx. 90 min), they were informed about the goal and procedure of the study before they gave their informed consent. Then, they completed the two computerized

tests (i.e., reading skill and WMC test). After that, they provided sociodemographic data and continued with the paper–pencil reasoning test (8-min limit). After a short break, they filled in the paper–pencil questionnaires (in the order of their presentation in the "Instruments" section). In session 2 (approx. 70 min), participants assessed their topic knowledge before they completed the topic knowledge test. Afterwards, they were informed about the number and nature of the to-be-read texts. They were instructed to read for understanding as their text comprehension would be tested afterwards. Then, participants read the first text (10-min limit). Reading time was fixed to control time-on-task and participants could not proceed before the time limit had expired. After text reading, participants received a description of the number and nature of the text comprehension questions. Then, they made their prediction judgments, answered the factual and inference questions, and provided their postdiction judgments. The procedure was repeated for the two other texts. Finally, participants received their course credit or money.

Materials

Texts

We used three expository texts (topics: the social-cognitive theory, the situated learning theory, and the attribution theory). They were comparable in length (630, 661, and 601 words, respectively) and readability (Flesch Reading Ease score 26, 31, and 38, respectively). The readability scores reflect rather difficult to read texts typical for higher education. The text order was identical for all participants as we were interested in evaluating individual differences in learner characteristics across tasks (cf. Carlson & Moses, 2001).

Text Comprehension Test

For each text, we used six factual questions and six inference questions. The factual questions asked for information explicitly stated in the texts. The inference questions referred to text information that had to be inferred from the given text information during reading. Both question types addressed central information of a text, yet the inference questions tested a higher level of understanding. Each question had a single-choice format with four response alternatives. As the questions addressed different parts of text information, as recommended for the assessment of metacomprehension (Wiley et al., 2005), internal consistency was low for the factual questions (Cronbach's $\alpha = 0.64$) and the inference questions (0.32).

Instruments

Reading Skill

The subtest "text comprehension" of the computerized German reading comprehension test for adults ELVES (Richter & van Holt, 2005) contained two

expository texts with 16 verification statements that focused on higher-order comprehension processes. The test score was an integrated value of performance correct and response latency (average Cronbach's α of 0.70, cf. Richter & van Holt, 2005).

Working Memory Capacity

In the computerized Reading-Span Task (Oswald et al., 2015; German version: Rummel et al., 2017, Cronbach's $\alpha = 0.85$), participants had to judge sentences as to whether or not they were sensible, each sentence followed by a letter to be recalled. Afterwards, the letters had to be identified in the correct order. The reading span score was, as recommended by Conway et al. (2005), the sum of letters recalled in the correct position. In line with common procedure, we filtered out scores of participants who misjudged 15% of all sentences or more, indicating that they, contrary to instruction, focused on letter memorization. Thus, the scores of 17 participants were omitted and treated as missing values.

Fluid Intelligence (Reasoning)

The nonverbal matrices subtest of the Intelligence Structure Test 2000 R (Amthauer et al., 2001; Cronbach's $\alpha = 0.71$) comprised 20 items and a time limit of 8 min. Each item presented five figures, of which one completed the logical structure of a given set of figures. Items were scored as correct or incorrect (1/0 points). A participant's sum of correct responses was converted into the age-dependent standard value.

Goal Orientation

The questionnaire SELLMO-St (Spinath et al., 2002) with 31 items and a 5-point scale (1 = *disagree completely*, 5 = *totally agree*) assessed goal orientation for university students in terms of mastery (8 items, Cronbach's $\alpha = 0.78$), performance-approach (7 items, $\alpha = 0.81$), performance-avoidance (8 items, $\alpha = 0.87$), and work-avoidance goal (8 items; not analyzed).[3]

Optimism

The Revised Life Orientation Test by Scheier et al., (1994, Study 2) measured dispositional optimism using six items about optimism (Cronbach's $\alpha = 0.80$) and four filler items, all on a 5-point scale (1 = *disagree completely*, 5 = *totally agree*).

---

[3] The subscales work-avoidance (goal orientation) and agreeableness (Big-Five) were not analyzed because they have not been investigated in previous studies or have not been shown to be relevant for metacomprehension accuracy. The subscales were, however, administered in order to contain the structure of the questionnaires.

Personality

The BFI-K, a short scale of the Big-Five Inventory in German (Rammstedt & John, 2005), included 21 items to assess extraversion (4 items; Cronbach's $\alpha = 0.82$), agreeableness (4 items; not analyzed)[1], conscientiousness (4 items; $\alpha = 0.72$), neuroticism (4 items; $\alpha = 0.79$), and openness (5 items; $\alpha = 0.74$) on a 5-point scale (1 = *disagree completely*, 5 = *totally agree*).

Overclaiming

The short form of the Overclaiming Questionnaire (Bing et al., 2011) measured self-enhancement as the tendency to inflate the amount of one's knowledge. It had 25 items and a 5-point scale (0 = *not familiar at all*, 4 = *very familiar*), of which 17 items were real terms (e.g., *myth*) and eight items were nonexistent terms (e.g., *meta-toxins*, Cronbach's $\alpha = 0.72$).

Self-Perceived Reading Skill

Based on Schraw (1994), we asked participants to assess their level of reading skill with a visual analogue (100 mm) item response format from 1 (*very low*) to 100 (*very high*). The term reading skill was explained, and the poles of the scale were labeled in the instruction. A participant's score was the mm value of their mark on the scale.

Self-Perceived Metacomprehension Ability

Following Schraw (1994), we asked participants to evaluate their ability to accurately monitor and assess their text comprehension using a visual analogue (100 mm) item response format from 1 (*very poor*) to 100 (*excellent*). The term metacomprehension ability was described, and the poles of the scale were labeled in the instruction. The position of the mark on the scale (in mm) represented the score.

General Monitoring Strategies

The General Monitoring Strategies Checklist (Schraw, 1997) included 10 items and a 5-point scale (1 = *disagree completely*, 5 = *totally agree*) to collect self-reported monitoring knowledge and strategies (e.g., *I stop and reread when I get confused*) when one has to demonstrate one's understanding or knowledge on a test. Cronbach's alpha was initially 0.63, but increased to 0.72 after excluding three items with a discrimination power below 0.30.

Self-Efficacy for Text Comprehension

We used six items of the Motivated Strategies for Learning Questionnaire (subscale self-efficacy for learning and performance; Pintrich et al., 1991) with a 5-point scale

(1 = *disagree completely*, 5 = *totally agree*). We changed their wording by referring to text comprehension instead of "learning in class". We instructed participants to answer the items with regard to reading an expository text about a rather new psychological topic about which they would have to answer comprehension questions. Cronbach's α was 0.81.

Epistemological Beliefs

The Epistemological Belief Questionnaire (Schommer, 1990; translated into German by Klopp, 2014) contained 63 items for five subscales (i.e., simple knowledge, omniscient authority, certain knowledge, innate ability, quick learning) and a 5-point response scale (1 = *disagree completely*, 5 = *totally agree*). Internal consistency (Cronbach's alpha) was very poor for all subscales, $Mdn = 0.32$, $M = 0.31$, Min = −0.10, Max = 0.59. An alternative factor structure proposed by Klopp showed the same problem. Therefore, epistemological beliefs were excluded from data analyses.

Self-Perceived Topic Knowledge

Participants indicated the amount of their prior knowledge of the topics of the three to-be-read texts (i.e., the social-cognitive theory, the situated learning theory, and the attribution theory; 3 items) on a 6-point response scale (1 = *nothing*, 6 = *a lot*). Cronbach's α was 0.86.

Topic Knowledge

Pre-existing knowledge about the three text topics was assessed with 12 questions, which had a single-choice format with four response alternatives and were scored as correct versus incorrect (1/0 points). Their content was not identical, neither verbatim nor paraphrased, with the questions of the text comprehension tests. Instead, the topic knowledge questions addressed general knowledge that would be helpful to understand the content of the texts. The questions tapped heterogenous aspects of the topics which is reflected in a low internal consistency: The initial Cronbach's α was 0.39; it increased to 0.51 after two items with a discrimination power near zero were excluded.

Metacognitive Judgments and Metacomprehension Accuracy

After reading a text, participants predicted how many of the comprehension test questions (values 0–6) they would answer correctly (i.e., prediction). After the test, they indicated how many questions (values 0–6) they thought they had answered

correctly (i.e., postdiction). Judgments were given separately for factual and inference questions.

Metacomprehension accuracy was calculated as judgment bias, which is the signed difference between a participant's judgment and their actual performance (negative values=underestimation, positive values=overestimation). Four outcome measures were calculated: the prediction bias and the postdiction bias of factual and inference questions, respectively.

Data Handling and Analysis

We used IBM SPSS Statistics version 27.0. Values of 2 *SD* above/below a variable's group mean were handled as outliers and replaced with this cutoff value. Missing data occurred in less than 1%. More specifically, (a) no more than 2% of data were missing on any questionnaire, (b) seven scores of the reading skill test were unavailable due to technical issues, and (c) the WMC score was missing once due to technical issues and was omitted for 17 participants due to violations of instruction (see the "Instruments" section). Missing data were replaced using the SPSS Multiple Imputation procedure.

To test the impact of the learner characteristics on metacomprehension accuracy, we conducted a hierarchical regression analysis (method: forward) for each of the four outcome measures (i.e., judgment bias of predictions/postdictions of factual/inference questions). Based on our assumption about the impact of learner characteristics when considered simultaneously (see the "Present Study" section) and similar to the analytical approach in Händel et al. (2020), we included the learner characteristics in three steps as predictors into the regression models: (1) gender because it has been shown to be related to several other cognitive, motivational and personality characteristics (see "Introduction"); (2) the objective measures of cognitive learner characteristics (i.e., reading skill, reasoning, WMC, and prior knowledge) as they can directly impact text comprehension and, hence, representation-based cues; (3) the subjective measures of self-perceptions of abilities and motivational and personality factors, whose impact on accuracy is mostly less substantiated by previous research (see "Introduction"). The statistical assumptions for the regression analyses were met (i.e., normal distribution of the outcome variables and residuals, linear relationships between predictors and outcome variables, homoscedasticity, and no multicollinearity).

To facilitate the interpretation of bias effects, we plotted the effect of each statistically significant learner characteristic on judgment bias showing whether high and low levels of the learner characteristic resulted in overestimation or underestimation (Fig. 1). Furthermore, we plotted the effects of each significant learner characteristic on judgment magnitude and performance separately to inspect whether its impact on bias was driven by influences on the one or the other (Figs. 2 and 3).

Results

Descriptive Statistics

The descriptive statistics for the outcome measures (see Table 1) show that the mean performance on both question types was on a medium level (i.e., no floor or ceiling effect). Thus, the mean judgment bias can be interpreted meaningfully. Moreover, judgment bias showed a substantial range in the extent of overestimation and underestimation with mostly a tendency toward overestimation. Table 2 presents the descriptive statistics for the learner characteristics and their bivariate correlations. Table 3 contains the bivariate correlations between the learner characteristics and the outcome measures (i.e., judgment magnitudes, performance, and judgment bias). It can be seen that a lot of the learner characteristics were significantly related to judgment magnitude, and a smaller number of them were (also) associated with performance. Those learner characteristics that were not related to the same extent to judgments and performance yielded a significant correlation with the bias measure. Overall, there were a few learner characteristics that showed a significant *bivariate* correlation with judgment bias. However, not all of them remained statistically significant for judgment bias when considered together in the regression analyses.

Regression Analyses

The regression analyses (see Table 4[4]) revealed that the statistically significant learner characteristics explained a moderate amount (cf. Cohen, 1988) of 17% and 14% of variance in the measures of prediction bias and a small amount of 11% and 12% of variance in the postdiction bias measures. Hence, learner characteristics were slightly more relevant for the prediction bias than for the postdiction bias. Which learner characteristics were significant for judgment bias differed between the factual and the inference questions.

The regression analysis for the prediction bias of factual questions revealed (see Table 4) that gender, WMC, topic knowledge, self-perceived topic knowledge, and self-perceived metacomprehension ability were statistically significant predictors. The impact of topic knowledge and reading skill was less straightforward, however: When initially included in step 2 of the analysis (see model 2 in Table 4), reading skill was statistically significant for the prediction bias, while topic knowledge was not. However, when self-perceived topic knowledge was included as

---

[4] For better readability, Table 4 contains only the beta-coefficients of the statistically significant learner characteristics as predictors for each judgment bias measure. In Appendix 2 Table 7, we provide the coefficients of all predictors, statistically significant and nonsignificant, for the judgment bias measures and, in addition, for judgment magnitude and performance as outcome measures.

significant predictor in the last step of the analysis, the weight of topic knowledge increased considerably (probably because some of its shared variance with self-perceived topic knowledge got cleared), revealing it as statistically significant, whereas the weight of reading skill decreased slightly falling below the level of significance.

Regarding the question whether high and low levels of the significant learner characteristics resulted in overestimation or underestimation, Fig. 1 shows that, in line with expectations, males, lower WMC, and lower topic knowledge, and, contrary to expectations, higher self-perceived topic knowledge and self-perceived metacomprehension ability were associated with stronger overestimations while their counterparts showed hardly any overestimation. Furthermore, to unravel how the identified learner characteristics impacted judgment bias, Fig. 2 presents their relationship with judgment magnitude and performance for the factual questions.[5] It can be seen that gender and self-perceived metacomprehension ability mainly influenced the magnitude of the prediction judgments, but hardly performance. In contrast, WMC, actual, and self-perceived topic knowledge had a significant effect on performance and a slight (albeit not statistically significant) influence on judgment magnitude.

For the postdiction bias of the factual questions, which slightly tended toward underestimation on the group level (see Table 1), underestimation was prevalent for females, higher reading skill, lower self-perceived topic knowledge, and lower self-perceived metacomprehension ability (see Fig. 1). Conversely, males, participants with lower reading skill, and higher self-perceptions of topic knowledge and of metacomprehension ability were hardly overestimating anymore. As Fig. 2 shows in greater detail, the effects of gender and self-perceived metacomprehension ability on postdiction bias resulted mainly from differences in judgment magnitude, not in performance. Self-perceived topic knowledge was related to inverse effects on both judgment magnitude and performance, that is, participants with higher self-perceived topic knowledge provided higher judgments but lower performance than participants with lower self-perceived knowledge, although the spread between judgment and performance was greater for the latter. Furthermore, the bias-inducing effect of reading skill was mainly due to the fact that participants with higher reading skill outperformed those with lower reading skill, while their judgments were almost equal.

Regarding the inference questions, the results for the prediction bias and postdiction bias were similar. As expected, a lower reading skill was associated with stronger overestimations than higher reading skill (see Fig. 1). A higher amount of

---

[5] The beta-coefficients for the relationships of the learner characteristics with judgment magnitude and performance, respectively, that are graphically displayed in Figs. 2 and 3 are provided as numerical values in Appendix 2 Table 6.

Table 1 Means (and standard deviations) for the outcome measures (N=255)

Variables	Factual questions	Inference questions
Text comprehension	60.22 (16.45)	49.83 (12.05)
Predictions[a] (magnitude)	64.61 (13.04)	59.83 (12.94)
Postdictions[a] (magnitude)	58.43 (14.67)	54.65 (13.28)
Prediction bias	+4.49 (17.62)	+9.87 (16.25)
Postdiction bias	−1.78 (16.35)	+5.01 (16.49)

Numbers are percent values

[a]Judgments were collected as absolute numbers but transformed into percent values for consistency of presentation

self-reported general monitoring strategies was related to stronger overestimations compared to lower levels of this learner characteristic, which, against our expectation, still showed some overestimation. Likewise, in line with expectations, a higher compared with a lower self-efficacy and, although only for the prediction bias, a higher compared to a lower openness contributed to stronger overestimations. As Fig. 3 breaks down, the effect of openness on the prediction bias was mainly based on differences in judgment magnitude, not performance. However, the effects of reading skill, self-efficacy, and self-reported general monitoring strategies on the prediction bias and postdiction bias were driven by divergent effects on both judgment magnitude and performance, albeit statistically significant differences in judgments and performance, respectively, occurred only for self-efficacy.[5] It is also noteworthy that gender, although it was a significant predictor in the first step of the regression analyses, became nonsignificant after motivational and personality factors were included in the analyses, indicating that the latter factors carry the true relationship between gender and judgment bias for inference questions.

Discussion

We examined which learner characteristics contribute to overestimation and underestimations of text comprehension. Based on a literature review and theory, we investigated multiple learner characteristics concurrently to consider interdependencies between them. In doing so, we circumvented a major gap in previous research and the present study is the first of its kind in the field of metacomprehension. Since our texts resulted, on average, in medium performance, overestimations and underestimations are not the pure by-product and, hence, not a statistical artifact of very hard or very easy tasks (cf. Juslin et al., 2000). Thus, the judgment bias and the impact of learner characteristics on it can be interpreted meaningfully.

Who Is at Risk to Overestimate or Underestimate Their Text Comprehension?

We found a general tendency toward overestimation, except for the postdictions of the factual questions. Which learner characteristics substantially contributed to

Table 2 Means and standard deviations for the learner characteristics and their bivariate correlations (N=255)

Variables	M	SD	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19 [a]
1 Reading skill	15.62	5.20	**.19**	**.12**	.11	-.07	-.03	**.34**	.04	.01	.11	.01	**.22**	-.11	-.07	-.01	.01	.07	**-.15**	.02
2 WMC	54.58	10.72		**.17**	-.05	.05	-.11	**.31**	**.19**	-.01	.05	.06	.10	-.08	**-.15**	-.12	.02	.02	.02	**-.13**
3 Reasoning	101.15	8.21			.03	-.01	-.03	**.22**	.06	.11	**.14**	.04	**.15**	-.04	-.08	-.06	.03	-.13	.00	-.02
4 Topic knowledge	4.01	1.91				**.13**	**.49**	.09	-.01	.12	.06	.03	**.14**	-.01	-.04	**.13**	-.07	-.09	.07	**.15**
5 Gms	3.57	0.52					**.17**	**.30**	**.35**	**.34**	**.21**	-.10	**.36**	**.16**	.11	**.37**	**-.24**	**.17**	**.31**	**-.13**
6 Self TK	1.74	1.05						-.05	.01	**.14**	-.02	-.09	**.14**	**.15**	.01	**.16**	-.12	-.07	.04	.07
7 Self RS	65.28	16.08							**.34**	**.15**	**.16**	-.05	**.53**	.12	-.04	**.14**	**-.23**	**.15**	**.17**	**-.13**
8 Self meta	60.29	18.19								-.03	.10	-.07	**.25**	.12	**.14**	.12	**-.30**	.03	**.23**	**-.17**
9 Mastery goal	4.47	0.39									**.19**	-.11	**.22**	.07	.07	**.35**	-.14	**.23**	**.24**	**.15**
10 Perf-approach goal	3.14	0.70										**.54**	**.16**	-.01	.06	**.17**	**.16**	.05	.01	.01
11 Perf-avoid goal	2.33	0.77											**-.13**	-.07	-.10	-.11	**.36**	-.09	**-.26**	.08
12 Self-efficacy	3.50	0.58												**.17**	-.06	**.19**	**-.29**	**.13**	**.23**	**-.23**
13 Overclaiming	0.86	0.54													**.18**	.08	-.10	.09	**.12**	-.01
14 Extraversion	3.67	0.86														**.13**	**-.16**	.06	**.27**	.06
15 Conscientiousness	3.64	0.69															**-.18**	.04	**.29**	**.21**
16 Neuroticism	3.20	0.90																-.05	**-.52**	**.30**
17 Openness	4.05	0.69																	**.15**	-.07
18 Optimism	3.44	0.72																		-.09
19 Gender [a] (0=male)																				

Values in bold represent p<.05

Gms general monitoring strategies, Self TK self-perceived topic knowledge, Self RS self-perceived reading skill, Self meta self-perceived general metacomprehension ability, Perf-approach goal performance-approach goal orientation, Perf-avoid goal performance-avoidance goal orientation

[a]Values represent η^2 coefficients

Table 3 Bivariate correlations between learner characteristics and the outcome measures (upper part of table) and bivariate correlations between the outcome measures (lower part of table)

Variables	Judgment magnitude of				Performance		Judgment bias for			
	Predictions		Postdictions				Predictions		Postdictions	
	Fa (1)	Inf (2)	Fa (3)	Inf (4)	Fa (5)	Inf (6)	Fa (7)	Inf (8)	Fa (9)	Inf (10)
Reading skill	.09	−.04	.06	−.04	**.23**	**.15**	**−.16**	**−.14**	**−.17**	**−.16**
WMC	**.18**	**.13**	**.23**	.11	**.30**	**.13**	**−.15**	.01	−.09	.00
Reasoning	**.14**	**.14**	**.14**	.11	**.17**	**.13**	−.05	.00	−.03	.00
Topic knowledge	.10	.06	**.20**	**.18**	**.20**	**.20**	−.11	−.10	−.03	−.01
Gms	**.16**	**.27**	**.15**	**.25**	.01	−.06	.11	**.26**	.12	**.26**
Self TK	.12	**.15**	**.14**	**.19**	−.07	.03	**.15**	.10	**.18**	**.14**
Self RS	**.29**	**.26**	**.28**	**.24**	**.29**	**.16**	−.07	.09	−.03	.06
Self meta	**.26**	**.33**	**.19**	**.32**	.03	.09	**.17**	**.20**	**.16**	**.19**
Mastery goal	.09	**.17**	**.15**	.11	.06	.04	.01	.10	.07	.05
Perf-approach goal	.02	.04	−.03	.05	−.01	−.03	.03	.05	−.02	.07
Perf-avoid goal	−.07	−.08	−.05	−.03	.02	−.01	−.07	−.06	−.09	−.03
Self-efficacy	**.39**	**.38**	**.35**	**.38**	**.16**	**.14**	**.13**	**.20**	**.16**	**.20**
Overclaiming	.08	.11	.00	.07	−.12	−.10	**.18**	**.16**	**.13**	**.13**
Extraversion	−.07	.03	−.07	.01	**−.13**	−.07	.07	.08	.08	.07
Conscientiousness	.06	.09	.02	.06	−.07	−.03	.10	.10	.09	.09
Neuroticism	**−.22**	**−.27**	**−.20**	**−.29**	.00	−.06	**−.16**	**−.17**	**−.19**	**−.20**
Openness	**.15**	**.23**	.12	**.14**	.03	−.04	.07	**.20**	.06	**.13**
Optimism	**.18**	**.32**	**.15**	**.29**	.04	.03	.08	**.23**	.09	**.23**
Gender[a] (0 = male)	**−.29**	**−.26**	**−.18**	**−.26**	.02	−.09	**−.22**	**−.13**	**−.18**	**−.15**
(1)		**.77**	**.77**	**.70**	**.30**	**.17**	**.45**	**.49**	**.39**	**.43**
(2)			**.63**	**.77**	**.15**	**.16**	**.42**	**.67**	**.41**	**.50**
(3)				**.73**	**.46**	**.24**	**.15**	**.32**	**.44**	**.40**
(4)					**.24**	**.19**	**.29**	**.47**	**.41**	**.66**
(5)						**.40**	**−.71**	**−.18**	**−.59**	−.11
(6)							**−.24**	**−.62**	**−.18**	**−.60**
(7)								**.52**	**.85**	**.41**
(8)									**.47**	**.84**
(9)										**.47**

Values in bold represent $p < .05$

Gms general monitoring strategies, *Self TK* self-perceived topic knowledge, *Self RS* self-perceived reading skill, *Self meta* self-perceived general metacomprehension ability, *Perf-approach goal* performance-approach goal orientation, *Perf-avoid goal* performance-avoidance goal orientation, *Fa/Inf* factual/inference questions

[a]Values represent η^2 coefficients

Table 4 Beta-coefficients (B values) of statistically significant predictors from the hierarchical regression analyses ($N = 255$)

Predictor variable	Model 1	Model 2	Model 3
Prediction bias factual questions			
Gender (0 = male)	− .22*** (− 8.44)	− .24*** (− 9.14)	− .21** (− 7.74)
Reading skill		− .13*	− .11
WMC		− .16*	− .18**
Topic knowledge		− .07	− .19**
Self-perceived topic knowledge			.23**
Self-perceived metacomprehension ability			.17**
ΔR^2 (total R^2)	.05*** (.05***)	.05** (.10***)	.07*** (.17***)
Postdiction bias factual questions			
Gender (0 = male)	− .18** (− 6.21)	− .17** (− 6.08)	− .17** (− 5.75)
Reading skill		− .17**	− .17**
Self-perceived topic knowledge			.19**
Self-perceived metacomprehension ability			.14*
ΔR^2 (total R^2)	.03** (.03**)	.03** (.06***)	.05** (.11***)
Prediction bias inference questions			
Gender (0 = male)	− .13* (− 4.54)	− .13* (− 4.44)	− .06 (− 2.17)
Reading skill		− .14*	− .17**
General monitoring strategies			.16*
Openness (Big-5)			.16**
Self-efficacy			.14*
ΔR^2 (total R^2)	.02* (.02*)	.02* (.04*)	.10*** (.14***)
Postdiction bias inference questions			
Gender (0 = male)	− .15* (− 5.17)	− .14* (− 5.05)	− .09 (− 2.99)
Reading skill		− .15*	− .17**
General monitoring strategies			.18**
Self-efficacy			.15*
ΔR^2 (total R^2)	.02* (.02*)	.02* (.04**)	.07*** (.12***)

For better readability, only the statistically significant predictors are presented. The B values for gender allow a direct interpretation of females' versus males' judgment bias

*$p < .05$; ** $p < .01$; *** $p < .001$

overestimation, or underestimation, differed mainly between factual and inference questions. For factual questions, we found that *men* provided considerably higher prediction and postdiction judgments than *women*, while their performance was equal. Consequently, men overestimated their comprehension for predictions of factual questions more strongly than women who were rather accurate. When making postdictions, men hardly overestimated and women underestimated their comprehension. These findings do not necessarily mean that men have a better metacomprehension accuracy in this situation because men and women have lowered their postdictions to an equal extent. It is common that learners drop their judgments from predictions to postdictions as a reaction to the test experience (Pierce & Smith,
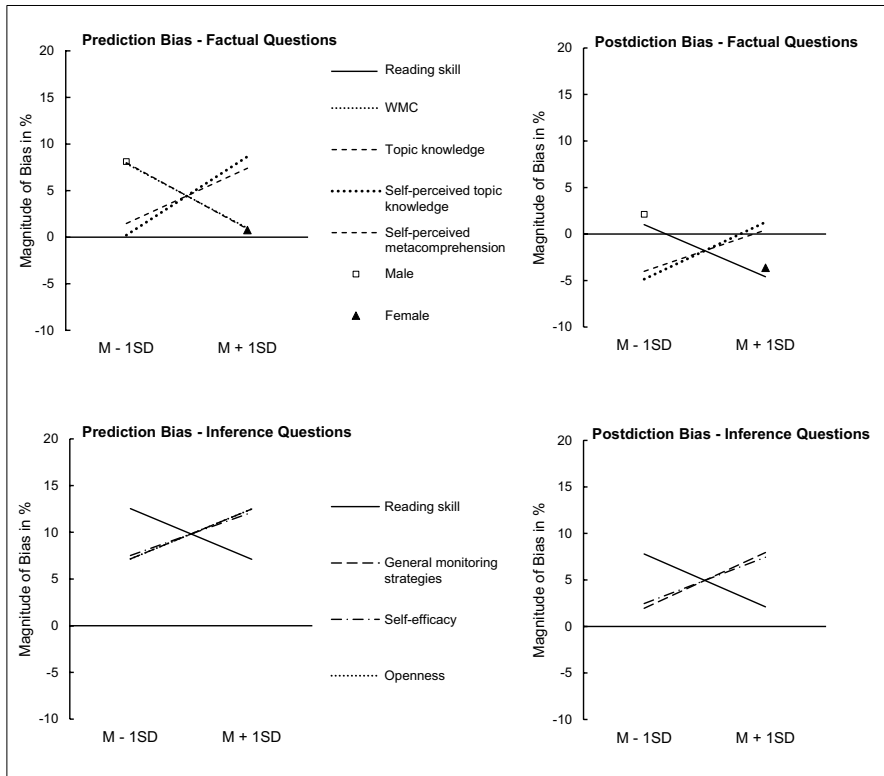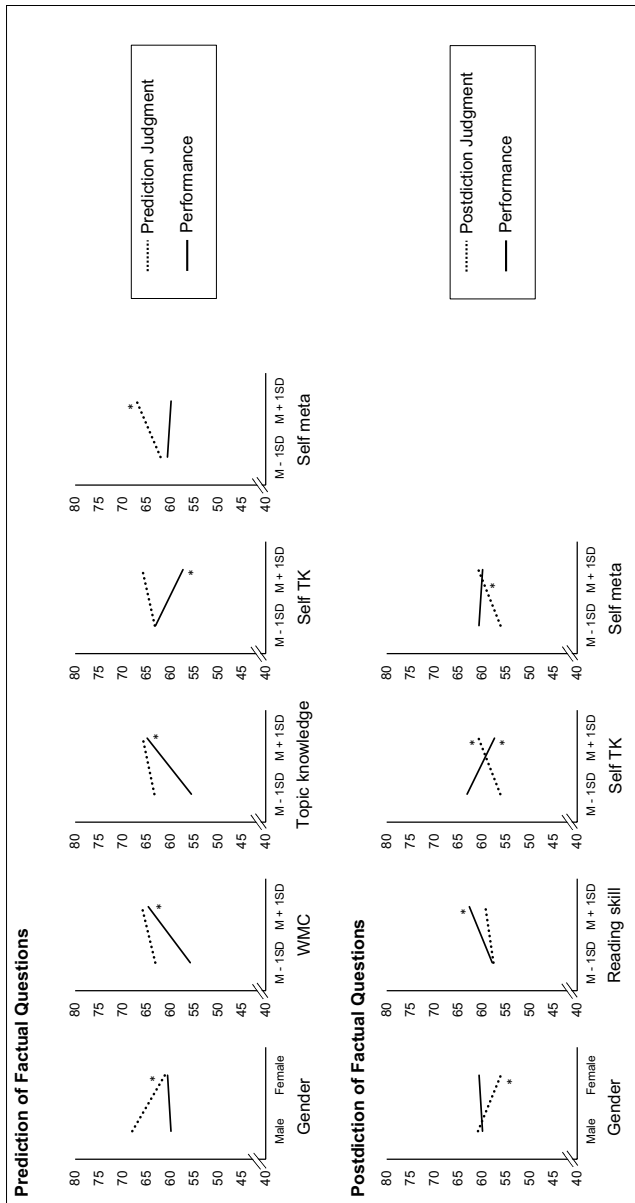
Note. Values are based on the unstandardized regression weights.

Fig. 1 Extent of overestimation/underestimation at the low ($M - 1\ SD$) and high level ($M + 1\ SD$) of a statistically significant learner characteristic, with the other significant factors set at their mean

2001). Hence, this finding may also reflect a general metacognitive heuristic rather than a gender-specific difference. For the inference questions, a gender effect disappeared after self-efficacy and the self-assessed amount of monitoring strategies were considered. Hence, the latter two factors accounted for men's stronger overestimations for inference questions. Thus, our findings add to previous studies that either did (e.g., Buratti et al., 2013; Stankov & Lee, 2008) or did not find a gender effect (e.g., Ariel et al., 2018; de Bruin et al., 2017), which altogether underline the relevance of contextual influences such as subject matter for the effect to occur.

A rather consistent finding was that a low *reading skill* contributed to overestimations, in particular for inference questions. For factual questions, a lower compared to a higher reading skill was also, albeit only marginally significantly, related to stronger overestimations for predictions. For the postdiction bias of factual questions, that, in general, tended toward underestimation, a higher reading skill resulted in underestimation and a lower reading skill in rather accurate judgments. This was, however, the result of differences in performance, not in judgment magnitude. Contrarily, for the inference questions, the stronger overestimations occurred because

Fig. 2 Impact of the statistically significant learner characteristics on judgment magnitude and performance (both in mean percent) for factual questions at the low ($M − 1$ SD) and high level ($M + 1$ SD) of a learner characteristic, with the other significant factors set at their mean

Note. Values are based on the unstandardized regression weights. Self TK = self-perceived topic knowledge; Self meta = Self-perceived general metacomprehension ability.

*p-value of the regression weight $< .05$.

Fig. 3 Impact of the statistically significant learner characteristics on judgment magnitude and performance (both in mean percent) for inference questions at the low $(M − 1 SD)$ and high level $(M + 1 SD)$ of a learner characteristic, with the other significant factors set at their mean
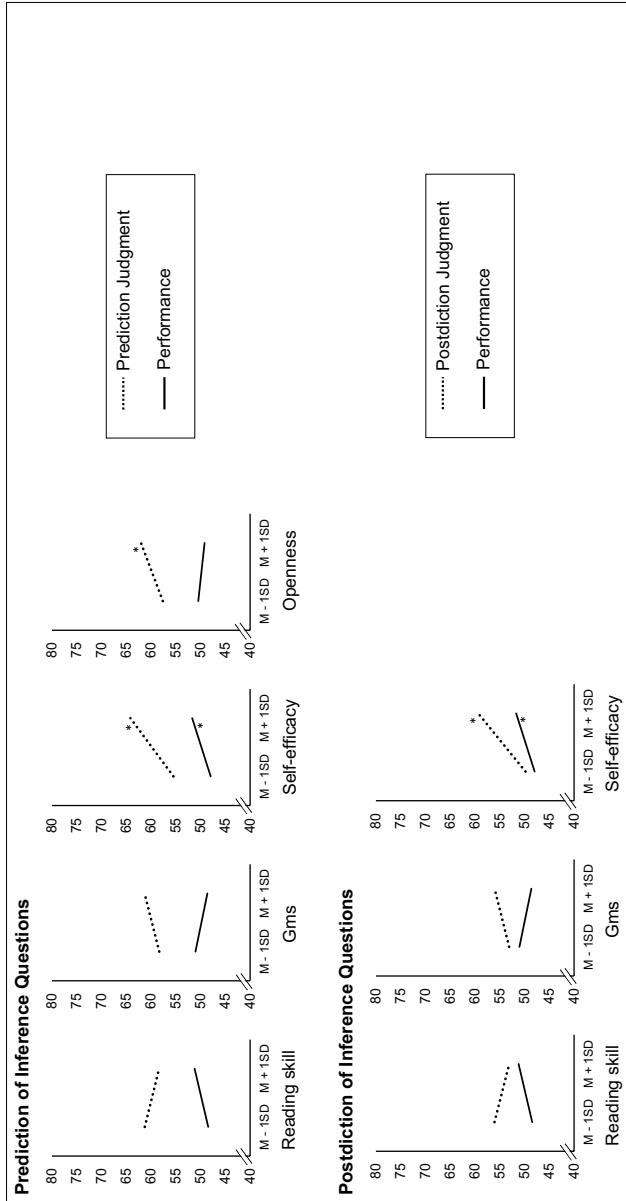
Note. Values are based on the unstandardized regression weights. Gms = General monitoring strategies.

*p-value of the regression weight is below .05.

participants with a lower reading skill provided slightly higher judgments and lower performance (both not statistically significant though) than those with higher reading skill. In all, these findings are compatible with prior studies using texts with age-appropriate text difficulty (Golke & Wittwer, 2017; Maki et al., 2005) and with theories on text comprehension and the cue-utilization framework on metacomprehension (cf. Griffin et al., 2009). Concretely, deficits in reading skill entail that learners construct poorer mental text representations (i.e., fewer, more incorrect, and less integrated information) for which they also have to invest more cognitive resources. Hence, they have fewer resources available for metacomprehension compared to learners with a higher reading skill. This includes less awareness of representation-based cues which therefore cannot be used for valid comprehension judgments and, relevant for postdictions, a poorer mental representation also hampers evaluating the correctness of one's responses to comprehension questions (Griffin et al., 2009; Koriat, 1997). In this case, learners with a lower reading skill are seduced to overestimate rather than to underestimate their comprehension.

Importantly, the detrimental effect of lower reading skill (in particular for judgment bias of inference questions) occurred although participants could reread each text. Rereading normally improves comprehension as it alleviates constraints in WMC which enhanced metacomprehension in terms of relative accuracy in Griffin et al.'s (2008) study. However, in the present study, either rereading as a regulation strategy was insufficient to alleviate poorer reading skills for the text materials or the less-skilled participants did not reread (they were not obliged to do so as opposed to Griffin et al.'s study). The latter interpretation is supported by the fact that learners with lower reading skills often hold erroneous beliefs about what text comprehension involves and when comprehension is attained (i.e., standards of coherence, cf. van den Broek et al., 2011). Furthermore, it is worth noting that participants in the present study were university students. Thus, even among learners with a comparatively high level of reading skill, differences in this ability affect judgment bias noticeably. For learners with poorer reading skills, this negative effect should be even more pronounced.

Moreover, a lower topic knowledge contributed to stronger overestimations for the prediction of factual questions. Similar to the role of reading skill, having little topic knowledge impedes the construction of a mental text representation (cf. Kendeou & O'Brien, 2016; van den Broek et al., 2002). Accordingly, we found that lower compared to higher topic knowledge was related to a substantially lower performance on the factual questions. Its positive effect on the magnitude of the prediction judgments was, however, only slight (and statistically nonsignificant). Hence, learners with low topic knowledge might have had some sense of their lower understanding, but they failed to adjust their prediction judgments properly compared to learners with high topic knowledge. The same pattern was found for the prediction judgments of the inference questions, albeit topic knowledge did not prevail its impact on the prediction bias over other learner characteristics. For the postdictions, however, learners with lower and higher topic knowledge were better able to match their judgments to their lower and higher performance, respectively; hence, no effect on postdictions bias occurred, neither for the factual nor the inference questions.

Theoretically, a low *WMC* should contribute to overestimations due to its limiting effect on comprehension (cf. van den Broek et al., 2002) and, hence, metacomprehension. For measures other than judgment bias, a detrimental effect has already been shown (Griffin et al., 2008; Ikeda & Kitagami, 2012; Unsworth & McMillan, 2013). Rereading of texts, however, can alleviate this effect (Griffin et al., 2008). Although rereading was allowed in the present study, we found that lower WMC did result in stronger overestimations, but only for the prediction bias of factual questions. This suggests, first, that lower-WMC participants not necessarily reread the texts; hence, lower WMC took its toll. Second, the factual questions asked for explicit text information. Hence, participants with a higher compared with a lower WMC could retain more of the facts which they believed to have gathered when making the prediction judgment and use it to answer the questions. The bias-reducing effect of a higher WMC was mainly driven by a higher performance, but also by slightly (though statistically nonsignificant) higher predictions. Hence, the deviation between prediction judgments and performance was smaller for higher compared to lower WMC. However, for the inference questions, solely a better memorization of facts due to a higher WMC was not sufficient as they asked for information that had to be inferred. For these questions, reading skill is more predictive. Likewise, the postdiction bias for both question types depended on reading skill rather than WMC alone. The absent effect of WMC on the postdiction bias measures is in line with Komori's (2016) findings on item-specific postdictions for a word recall task.

Over and above gender and the cognitive factors, a higher *self-perceived topic knowledge* contributed to overestimations for the predictions of factual questions while a lower level was related to rather accurate judgments. This effect occurred because a higher compared to a lower self-perceived topic knowledge resulted in slightly higher judgments and considerably lower performance. As the postdiction judgments decreased, participants with a higher self-perceived topic knowledge were rather accurate while those with a lower level were underconfident. These findings correspond with prior research: A higher domain/topic familiarity gives learners a heightened sense of confidence because familiarity concerns the knowledge about a general topic (which is vague), while test questions address a certain text (Lin & Zabrucky, 1998). Our finding of the positive relationship between familiarity and judgment magnitude is in line with previous studies (Ehrlinger & Dunning, 2003, Exp. 3; Glenberg et al., 1987, Exp. 3 + 4; Griffin et al., 2009). It differs, however, with respect to performance–in these previous studies, higher familiarity did not result in lower performance, but in higher performance or no differences at all. It can be speculated that the self-perceived knowledge was generally very low in the present study. Hence, knowing to have no or hardly any knowledge about a topic might be less harmful for text comprehension than believing to have some knowledge about it.

Furthermore, a higher compared to a lower *self-perceived metacomprehension ability* contributed to stronger overestimations on the predictions of the factual questions due to higher judgments, yet equal performance. This finding indicates that learners have an unrealistically high self-perception of their monitoring ability or they have underdeveloped beliefs about what accurate monitoring entails. This interpretation is supported by the fact that self-perceived metacomprehension ability

was statistically unrelated to actual reading skill. When making postdictions for the factual questions, the participants with a higher self-perceived metacomprehension ability were rather accurate while the self-described lower-ability participants were underconfident. This might again be simply a by-product of the general tendency to lower one's postdiction judgments. For the inference questions, the self-perceived metacomprehension ability was eventually not predictive for the bias, although their bivariate correlations suggested so. Instead, the self-reported amount of general monitoring strategies and knowledge, which seems to be more specific to inference questions than metacomprehension ability in general, became relevant. Interestingly, in Schraw and Dennison's (1994) study, a higher self-perceived metacomprehension ability was also associated with higher judgments (in their case, item-specific post-dictions), but with an even higher performance on text comprehension questions as well. This resulted in stronger overestimations for lower self-perceived metacomprehension ability in their study. Differences in text materials might explain these contradicting findings for performance, yet due to missing details in the description of their materials, this remains an open question.

With respect to the question which learner characteristics' impacted judgment bias for the inference questions, reading skill was already discussed as one source of overestimation. Moreover, a higher self-reported amount of *general monitoring strategies and knowledge* was associated with stronger overestimations of both predictions and postdiction judgments for these questions as the self-ascribed more knowledgeable learners tended toward higher judgments and lower performance. Since Schraw (1997) found a reverse effect (lower knowledge, more underestimation), it suggests that, in the present study, the questionnaire assessed an inflated sense of monitoring skills and knowledge, which is a common threat in assessing strategy knowledge (cf. Veenman, 2016). Its medium correlation with self-perceived metacomprehension ability, which activated overestimations of the factual questions, supports this explanation. Hence, future research should consider more objective (e.g., behavioral) measures of monitoring knowledge and strategies.

Furthermore, we found for the inference questions that a higher *self-efficacy for text comprehension* resulted in higher performance and even more so in higher judgments. Consequently, higher self-efficacy contributed to stronger overestimations for predictions and postdictions than lower self-efficacy, which is in line with our expectation and prior research (Bouffard-Bouchard, 1990). This bias-inducing effect might seem counter-intuitive given that a higher self-efficacy promotes performance (Zimmerman & Moylan, 2009) and should, hence, decrease overestimations. However, as Zimmerman and Moylan state, self-efficacy as a learner's beliefs about how capable they are to perform a certain task implies performance judgments and their accuracy. Given that learners usually tend toward overestimation (e.g., Lin & Zabrucky, 1998), (higher) self-efficacy beliefs might in turn be generally overoptimistic. Since overestimations are, however, likely to produce underachievement in later testing situations (Dunlosky & Rawson, 2012), future research should disentangle the pros and cons of a higher self-efficacy in the various phases of self-regulated learning.

Moreover, learners with a higher *openness* arrived at higher predictions, yet lower performance on inference questions, hence stronger overestimation, than those with lower openness. This finding corresponds to Agler et al.'s (2019) finding on

relative metacomprehension accuracy, albeit in their study the effect disappeared when all Big-Five facets were considered concurrently. It might be speculated that openness is particularly predictive for bias of higher-order comprehension questions such as inference questions (Agler et al. did not differentiate between factual/inference questions): Inference questions address higher-order processes (e.g., causal relations between/across sentences) and are, therefore, usually more complex and difficult than factual questions. Correspondingly, learners in the present study gave lower judgments and performed poorer on the inference questions. The complexity and difficulty of these questions might heighten learners' uncertainty about their performance on them. In this unclear situation, participants with a higher openness, who have also been found to experience less self-doubt (Buratti et al., 2013), show more overestimation for predictions (not postdictions) than less open participants.

In sum, learners with a low reading skill and, when controlled for cognitive capabilities, higher self-perceptions/beliefs about their topic knowledge and capabilities in understanding (self-efficacy) and monitoring texts (self-perceived metacomprehension ability) are particularly at risk to overestimate their comprehension of expository texts, especially when these risk factors coexist. A low WMC, a low topic knowledge, a higher openness, and being male can also contribute to overestimations, depending on the task. Thus, relatively few of the various learner characteristics that have been studied (mostly separately) in prior work emerged as significant predictors for judgment bias when considered simultaneously. This finding corresponds with Händel et al.'s (2020) study for judgment accuracy on knowledge tests.

Which Learner Characteristics Have No Substantial Effect on Judgment Bias?

Various learner characteristics turned out to be negligible for judgment bias of text comprehension. The absent effects of *goal orientations*, in particular the performance-approach and performance-avoidance goal orientation, contradict previous findings by Kroll and Ford (1992) and Zhou (2013). Our finding might, thus, suggest that previously found effects of goal orientation are rather attributable to other learner characteristics which are more closely related to text comprehension, for instance, self-efficacy.

The findings on the other learner characteristics that were insignificant for judgment bias are less surprising since evidence from prior research has only been weak or limited. This includes *reasoning ability* (cf. Ohtani & Hisasaka, 2018). Thus, the intellectual capability of drawing logical inferences per se does not help the accuracy of comprehension judgments, at least for (above) average levels of reasoning of university students. Instead, reading skill as the ability underlying text comprehension is essential.

Furthermore, the *self-perceived reading skill* had no significant impact on judgment bias, although Kwon and Linderholm (2014) suggested so. We found, just as Kwon and Linderholm did, a moderate relationship between self-perceived and actual reading skill. Yet, the present findings suggest that self-efficacy as the task-specific beliefs about an upcoming text comprehension task and self-perceived

metacomprehension ability are more relevant to judgment bias than the generalized beliefs about one's reading skill.

Moreover, our findings showed that a learner's tendency to present themselves overly positive (as measured via *overclaiming* knowledge) is irrelevant or, compared to other learner characteristics, negligible for bias. This finding underlines the accumulated evidence that poor metacomprehension accuracy mainly results from a poor choice of judgments cues (cf. Thiede et al., 2009), not from self-elevation, at least in low-stake test situations.

In addition, we found that *dispositional optimism* as the belief in positive outcomes in life did not significantly add to explain poor metacomprehension accuracy. This finding corresponds with previous research that found no reliable relationship between bias measures and optimism (de Bruin et al., 2017; Händel et al., 2020 regarding the psychology test), except for Händel et al.'s study on a math test. Thus, learners in general are not strongly inclined to be misled by their optimistic disposition.

With regard to the Big-Five personality, we found that, beyond openness, the other examined facets (i.e., extraversion, neuroticism, and conscientiousness) had no substantial impact on the judgment bias of text comprehension. This finding mainly replicates previous studies that analyzed personality facets simultaneously and found either no effect (Agler et al., 2019; Händel et al., 2020 for math test) or an effect of a single facet (extraversion in Schaefer et al., 2004; conscientiousness in Händel et al., 2020 for psychology test). Our finding also does not conflict with previous studies (e.g., Buratti et al., 2013; Dahl et al., 2010; Pallier et al., 2002) that reported an effect of one or the other Big-Five facet on bias, but based on bivariate correlations only. Thus, the impact of personality facets on judgment bias (for text comprehension) seems to be weak. However, it is thinkable that the impact of personality facets (e.g., neuroticism) depends on affordances of the learning situation (e.g., high-/low-stake tests, text difficulty, domain).

Theoretical and Practical Implications

The cue-utilization framework (Griffin et al., 2009; Koriat, 1997) acknowledges a potential impact of learner characteristics on accuracy. Yet, given the empirical evidence, it needs to be developed with explicit assumptions about which learner characteristics impact accuracy. These should include the role of cognitive, motivational, and personality factors and whether they impact the emergence of representation-based cues, act as heuristics cues, or function as a bias-inducing cognitive style. Since task requirements such as task complexity (e.g., factual vs. inference questions) seem to trigger the influence of different learner characteristics, the theoretical framework should also elaborate on the conditions under which learner characteristics influence accuracy.

Most of the learner characteristics we found to contribute to overestimations are modifiable through instruction or training. These include reading skill (also beliefs about what different levels of text understanding mean) and self-perceptions of one's abilities or knowledge. Such instructional support is naturally part of the education of younger students, but–as the present study showed–also university students would benefit from such support. Learner characteristics are, however, not the only impact

on poor metacomprehension accuracy. Text and task characteristic and unsystematic influences (such as guessing) are also relevant. Hence, students should be provided with interventions such as delayed keyword generation or summarization that have been shown to be effective for de-biasing metacognitive judgments (e.g., Prinz et al., 2020b; Thiede et al., 2009).

Limitations and Future Research

As metacomprehension accuracy depends on the learner, the text, and the test material, the present findings might not be generalized to other text genres or types of tasks. As discussed, the type of task (factual vs. inference questions) was relevant in the present study and future research might focus on other material as well. Moreover, the present findings might not be generalized to samples with, for example, clinically relevant high levels of neuroticism or narcissism, or to learners with severely low levels of cognitive capacities or self-perceptions.

To assess reasoning ability, we used only a single test instead of a thorough diagnostic. However, as no individual diagnoses were intended, and for the sake of the practicability of the study, it seems to be an acceptable solution. Likewise, the topic knowledge test did not include multiple items for each aspect. Hence, internal consistency of the topic knowledge items was low. However, the results on the topic knowledge test were plausible given that almost half of the sample were at the beginning of their first semester. Additionally, internal consistency for the inference questions of the text comprehension tests was low. This can be explained by the fact that, as recommended in metacomprehension research (cf. Wiley et al., 2005), the questions covered various main aspects of the texts. The questions barely overlapped with each other; hence, internal consistency was low. Hence, to substantiate the present findings, replication studies are needed.

Regarding the data analyses, epistemological beliefs could not be modeled. This problem has been reported before (cf. Klopp, 2014) indicating flaws in the construct itself. The fact that the impact of epistemological beliefs was not investigated might however have only marginally affected the overall results (if at all) because, following Griffin et al. (2013), epistemological beliefs might shape the reading goals and activities rather than the judgments or their accuracy.

The findings of the present study should encourage further studies that aim to replicate the present findings and transfer them to other text materials. In a next step, studies should aim to advance the understanding of the interplay between the learner characteristics. Although interdependencies between learner characteristics are taken into account in multiple regression analyses, they cannot unveil the specific pattern of their interplay. To investigate how relevant learner characteristics interact (e.g., mediations) to predict overestimations and underestimations of text comprehension, other analytical approaches (e.g., path models) and more research on the relationships between cognitive, motivational, and personality factors are needed. Moreover, longitudinal studies might be fruitful to investigate the dynamic interplay between metacomprehension accuracy, performance, and learner characteristics, of which self-efficacy as a motivational characteristic seems particularly interesting, given the present findings on its counteracting effects on performance and judgment bias.

Appendix 1

Table 5 Summary of reviewed empirical findings on the impact of learner characteristics on accuracy

Study	Judgment type	Performance type	Effect of learner characteristic on accuracy[c]	Relationship of learner characteristic with	
				Judgment magnitude	Performance
Reading skill					
Golke & Wittwer (2017)	Pre, post	Text: fact	For age-appropriate text difficulty (cohesion): - Lower skill with stronger overestimation than higher skill For easy texts (high cohesion): no effect of skill on judgment bias	n.a.	n.a.
Griffin et al. (2008, Exp. 1)	Pre	Text: inf	- Read-once: low skill with poorer relative accuracy than high skill - Reread: no effect of skill on relative accuracy	n.a.	n.a.
Maki et al. (2005)	Pre, post	Text: fact; "more conceptual information"	For age-appropriate text readability ('hard texts' in original study): - Low skill: overestimation on predictions, not biased on postdictions - Medium, high skill: not biased on predictions, underestimation on postdictions - No effect of skill on relative accuracy For easy-to-read texts ('revised text'): no effect of skill on judgment bias and relative accuracy	Predictions: high, low > medium skill[d] Postdictions: high> low > medium skill[d]	High > medium > low skill[d]
WMC					
Griffin et al. (2008, Exp. 2)	Pre	Text: inf	- Read-once: higher WMC with higher relative accuracy than lower WMC - Reread: no effect of WMC on relative accuracy	n.a.	n.a.
Ikeda and Kitagami (2012)	Pre	Text: inf	- For intact texts: high WMC with higher relative accuracy than low WMC	=	+
Komori (2016)	Post (per item)	Word recall	- High WMC with better absolute accuracy than low WMC - High and low WMC: underestimation to equal extent Note: very high performance in high WMC group	+	+

Table 5 (continued)

Study	Judgment type	Performance type	Effect of learner characteristic on accuracy[c]	Relationship of learner characteristic with	
				Judgment magnitude	Performance
Reasoning ability					
Ohtani and Hisasaka (2018) *Meta-analysis*			- Higher reasoning with rather better accuracy than lower reasoning; Note: effects aggregated across domains, accuracy measures, and prediction/postdictions	n.a.	+
Topic knowledge					
De Bruin et al. (2007)	Pre	Text[b]	- Lower knowledge (novices) with poorer relative accuracy than higher levels (experts)	+	+
Griffin et al. (2009)	Pre	Text: fact, inf	- Higher knowledge with better absolute accuracy than lower levels; - Lower knowledge with stronger underestimation than higher levels; - No effect of knowledge on relative accuracy	+	+
Metacognitive knowledge (GMS = general monitoring strategies)					
Schraw (1997)	Post (per item)	Text (and others)	- Low knowledge of GMS for taking tests: underestimation; - Medium, high knowledge of GMS: not significantly biased	High > medium > low GMS[d]	Low > high > medium GMS[d]
Schraw and Dennison (1994, Exp. 2)	Post (per item)	Text	- Knowledge about cognition and regulation: no effect on bias	+	+
Self-perceived topic knowledge (domain/topic familiarity)					
Ehrlinger and Dunning (2003), Exp. 3	Post (per item)	Knowledge (geography)	- n.a.	Under control of performance: +	=
Glenberg et al. (1987, Exp. 3+4)	Pre	Text: recall, inf	- n.a.	+	Recall: + Inf: =
Griffin et al. (2009)	Pre	Text: fact, inf	- Lower topic familiarity with more underestimation, higher topic familiarity less underestimation/rather accurate; - Lower familiarity with poorer absolute accuracy than higher familiarity; - No effect on relative accuracy	+	+

Table 5 (continued)

Study	Judgment type	Performance type	Effect of learner characteristic on accuracy[c]	Relationship of learner characteristic with	
				Judgment magnitude	Performance
Maki and Serra (1992, Exp. 3)	Pre	Text	- Feedback about topic knowledge prior to reading and giving judgments enhanced relative accuracy compared to a nonfeedback condition	n.a.	n.a.
Self-perceived reading skill					
Kwon and Linderholm (2014)	Pre, post	Text: fact, inf	- n.a.	Under control of reading skill: +	n.a.
Self-perceived metacomprehension ability (self-meta)					
Schraw (1994)	Post (per item)	Text	- Low self-meta with poorer absolute accuracy than medium and high self-meta	High > medium, low self-meta	High > medium, low self-meta
Schraw and Dennison (1994, Exp. 2)	Post (per item)	Text	- Lower self-meta with stronger overestimation than higher self-meta	+	+
Self-efficacy					
Bouffard-Bouchard (1990)	Post (per item)	Problem solving	Based on visual plots: - Judgments and performance more consistently related to each other in the high self-efficacy group than in the low self-efficacy group - High self-efficacy group gave more often overconfident judgments - Low self-efficacy group gave rather equally often over-confident and underconfident judgments	+	No. completed problems: + No. correct solutions: =
Händel et al. (2020, Study 1)	Post, overall and per item	Knowledge test (exam) in math and psychology	Under control for gender, high-school grade, and further motivational and personality factors as predictors: - No effect of self-efficacy on bias in math and psychology exam	Under control of … (see left) and test performance: =	Bivariate correlation: +
Vössing and Stamov-Roßnagel (2016)	Pre	Text	- n.a.	+	=
Goal orientation					
Ikeda et al. (2016, Exp. 3)	Pre	Text: fact, inf	- n.a.	Performance-approach > mastery-approach	=

Table 5 (continued)

Study	Judgment type	Performance type	Effect of learner characteristic on accuracy[c]	Relationship of learner characteristic with	
				Judgment magnitude	Performance
Kroll and Ford (1992)	Pre	Text	- Higher performance-orientation with stronger overestimation than lower levels - Mastery-orientation: no effect on bias	= =	- +
Zhou (2013)	Pre, post	Text: fact, inf	- Higher performance-approach, performance-avoidance and mastery-avoidance with stronger overestimation for predictions than lower levels - Mastery-approach: no effect on bias for predictions - No effect of any goal orientation on bias for postdictions	n.a.	n.a.
Personality					
Agler et al. (2019)	Pre, post	Text	- All Big-5 facets analyzed simultaneously: no effect of any facet on relative accuracy	When analyzed simultaneously: Openness + Extraversion - Else =	When analyzed simultaneously: All facets =
Buratti et al. (2013)	Post (per item)	General knowledge	Bivariate correlations: - Higher openness and extraversion/narcissism (composite) with stronger overestimation than lower levels - Higher agreeableness with better absolute accuracy than lower levels	Multiple regression and under control of age, gender, test performance: Openness + Extrav./narcs. + Agreeableness = Conscientiousness =	Bivariate correlations: Openness + Else =
Dahl et al. (2010)	Post (per item)	General knowledge	Bivariate correlations: - Higher extraversion and conscientiousness with stronger overestimation than lower levels - Other Big-5 facets: no effect on bias	Under control of age, yrs education, performance: Neuroticism - Openness + Else =	Under control of age, yrs education: Neuroticism - Openness + Extraversion - Else =
Händel et al. (2020, Study 1)	Post, overall and per item	Knowledge test (exam) in math and psychology	Under control for gender, high-school grade, and further motivational and personality factors as predictors: - No effect of Big-5 facets on bias for any postdictions in math - Higher conscientiousness with stronger overestimation (lower conscientiousness: rather underestimation), other facets no effect on bias, for overall postdictions in psychology	Under control of gender, grade, exam performance: = Conscientiousness + Else =	Bivariate correlation: = =

Table 5 (continued)

Study	Judgment type	Performance type	Effect of learner characteristic on accuracy[c]	Relationship of learner characteristic with	
				Judgment magnitude	Performance
Pallier et al. (2002, Exp. 2)	Post (per item)	Composite of knowledge and ability tests	- Higher extraversion more overestimation than lower extraversion	=	n.a.
Schaefer et al. (2004)	Post (per item)	General knowledge	Under control of the other personality facets (partial correlations): - Higher extraversion with stronger overestimation than lower extraversion - Other Big-5 facets: no effect on bias	Under control of other personality facets: Extraversion + Openness + Conscientiousness + Else =	Under control of other personality facets: = + = =
Dispositional optimism					
De Bruin et al. (2017)	Pre[a]	Knowledge (exam)	- Under control for gender, language: no effect of optimism on bias	n.a.	n.a.
Händel et al. (2020, Study 1)	Post, overall and per item	Knowledge test (exam) in math and psychology	Under control for gender, high-school grade and further motivational and personality factors as predictors: - No effect of optimism on bias for item-specific postdictions in math - Higher optimism with stronger overestimation (lower optimism: rather underestimation) for overall postdictions in math - No effect of optimism on bias for overall postdictions in psychology	Under control of gender, grade, exam performance: = + =	Bivariate correlation: = =
Overclaiming					
Bensch et al. (2019)	Post (per item)	General knowledge	- Higher overclaiming with stronger overestimation than lower overclaiming	n.a.	n.a.
Gender					
Ariel et al. (2018)	Pre, post	Ability tests (e.g., spatial relations)	- No effect on bias - For 1 out of 3 tests: men with poorer relative accuracy than women	Men > women for most ability tests	Men > women for most ability tests
Buratti et al. (2013)	Post (per item)	General knowledge	- Men with stronger overestimation than women	Men > women	Men > women
Dahl et al. (2010)	Post (per item)	General knowledge	- No effect on bias and absolute accuracy	Men > women	Men > women
De Bruin et al. (2017)	Pre[a]	Knowledge (exam)	- No effect on bias	n.a.	n.a.

Table 5 (continued)

Study	Judgment type	Performance type	Effect of learner characteristic on accuracy[c]	Relationship of learner characteristic with	
				Judgment magnitude	Performance
Händel et al. (2020)	Post, overall and per item; Pre	Knowledge test (exam) in math and psychology	Study 1: - Men with stronger overestimation than women for overall postdictions in math - No effect on bias for overall postdictions in psychology and for item-specific postdictions in math Study 2: - No effect on bias for overall pre- and postdictions in psychology		
Ring et al. (2016)	Post	Cognitive reflection	- No effect on bias	Men> women	Men > women
Stankov and Lee (2008)	Post	Text (in L2)	- Men with stronger overestimation than women	=	Men < women

Pre prediction, *Post* postdiction, *n.a.* information not available

[a] Of course test grades from 0 to 10

[b] Computerized presentation of chess rules

[c] Studies may include further conditions with experimental manipulations, e.g., intervention or texts with deleted words, which were, however, not used in the present study

[d] Information based on descriptive statistics (means) from the study

Appendix 2

Table 6 Beta-coefficients from multiple regression analyses with learner characteristics as predictors and magnitude of prediction and postdiction judgments as well as performance as outcome measures ($N = 255$)

Predictor variable	Prediction (magnitude)	Performance	Postdiction (magnitude)
Factual questions			
Gender (0 = male)	− 0.26***	0.02	− 0.16**
Reading skill	n.t	0.15*	0.06
WMC	0.11	0.27***	n.t
Topic knowledge	0.09	0.28***	n.t
Self-perceived topic knowledge	0.11	− 0.18**	0.16*
Self-perceived metacomprehension ability	0.19**	− 0.02	0.16*
Total R^2	0.17***	0.18***	0.09***
Inference questions			
Reading skill	− 0.11	0.12	− 0.11
General monitoring strategies	0.12	− 0.10	0.11
Openness (Big-5)	0.17**	− 0.05	n.t
Self-efficacy	0.34***	0.16*	0.36***
Total R^2	0.20***	0.05*	0.17***

This table provides the beta-coefficients that are graphically displayed in Figs. 2 and 3. We included those learner characteristics as predictors in each multiple regression (method: enter) that have been identified as statistically significant for the respective judgment bias measures as reported in Table 4

n.t. not tested

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Table 7 Beta-coefficients from the hierarchical regression analyses with learner characteristics as predictors and magnitude of prediction and postdiction judgments, performance, as well as prediction bias and postdiction bias as outcome measures ($N=255$)

Predictor variable	Pre	Post	Perf	Pre bias[b]	Post bias[b]
Factual questions					
Step 1:					
Gender (0 = male)	− 0.21**	− 0.12	0.04	− 0.21**	− 0.17**
Step 2:					
Reading skill	− 0.02	− 0.06	0.10	− 0.11	− 0.17**
WMC	0.08	0.20**	0.22***	− 0.18**	− 0.09
Reasoning	0.10	0.06	0.07	0.00	− 0.01
Topic knowledge	0.11*	0.19**	0.27***	− 0.19**	− 0.10
Step 3:					
General monitoring strategies	− 0.05	0.00	− 0.05	0.00	0.01
Self-perceived topic knowledge	0.08	0.05	− 0.17*	0.23**	0.19**
Self-perceived reading skill	0.01	0.07	0.16*	− 0.06	− 0.04
Self-perceived metacomprehension ability	0.14*	0.07	− 0.08	0.17**	0.14*
Mastery goal	0.03	0.10	0.03	0.03	0.08
Performance-approach goal	− 0.07	− 0.09	− 0.08	0.04	− 0.01
Performance-avoidance goal	− 0.01	− 0.03	− 0.01	. − 0.02	− 0.04
Self-efficacy	0.25***	0.28***	0.03	0.07	0.12
Overclaiming	0.02	− 0.03	− 0.09	0.11	0.07
Extraversion	− 0.05	− 0.01	− 0.07	0.02	0.05
Conscientiousness	0.05	− 0.01	− 0.07	0.09	0.08
Neuroticism	− 0.04	− 0.10	0.03	− 0.01	− 0.09
Openness	0.12*	0.09	0.01	0.07	0.07
Optimism	0.05	0.07	0.02	0.00	0.01
Total R^{2a}	0.25***	0.20***	0.20***	0.17***	0.11***
Inference questions					
Step 1:					
Gender (0 = male)	− 0.17**	− 0.19**	− 0.11	− 0.06	− 0.09
Step 2:					
Reading skill	− 0.09	− 0.10	0.11	− 0.17**	− 0.17**
WMC	0.04	0.03	0.14*	0.01	− 0.01
Reasoning	0.13*	0.07	0.11	0.03	0.00
Topic knowledge	− 0.01	0.16**	0.20**	− 0.11	− 0.02
Step 3:					
General monitoring strategies	0.01	0.00	− 0.10	0.16*	0.18**
Self-perceived topic knowledge	0.14**	0.12	− 0.08	0.06	0.10
Self-perceived reading skill	− 0.01	− 0.01	0.12	− 0.02	− 0.05
Self-perceived metacomprehension ability	0.20***	0.19**	0.06	0.12	0.10
Mastery goal	0.04	0.04	0.02	− 0.01	− 0.04
Performance-approach goal	− 0.03	− 0.01	− 0.05	0.01	0.03
Performance-avoidance goal	0.05	0.07	− 0.02	− 0.01	0.02
Self-efficacy	0.18**	0.22***	0.10	0.14*	0.15*

Table 7 (continued)

Predictor variable	Pre	Post	Perf	Pre bias[b]	Post bias[b]
Overclaiming	-0.01	-0.02	-0.08	0.08	0.06
Extraversion	-0.04	-0.05	-0.04	0.05	0.05
Conscientiousness	-0.01	-0.04	-0.04	0.02	0.01
Neuroticism	0.02	-0.02	-0.05	-0.07	-0.10
Openness	0.18**	0.09	-0.03	0.16**	0.09
Optimism	0.19**	0.17**	0.01	0.11	0.12
Total R^{2a}	0.32***	0.27***	0.06**	0.14***	0.12***

The beta-coefficients resulted from the 3-step hierarchical regression analyses (method: forward) outlined in the "Data Handling and Analysis" section. In these analyses, statistically nonsignificant predictors are excluded from the model. However, the beta-coefficients of the nonsignificant predictors are also included in this table to provide complete information. *Pre/Post* magnitude of prediction/postdiction judgment, *Perf* performance, *Pre/Post Bias* Bias of predictions/postdictions

[a]The total R^2 of the statistically significant predictors

[b]The values of the statistically significant predictors correspond with the values in Table 4

*$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Declarations

Conflict of Interest The authors declare no competing interests.

References

Agler, L.-M. L., Noguchi, K., & Alfsen, L. K. (2019). Personality traits as predictors of reading comprehension and metacomprehension accuracy. *Current Psychology*, 1–10.https://doi.org/10.1007/s12144-019-00439-y

Alarcon, G. M., Bowling, N. A., & Khazon, S. (2013). Great expectations: A meta-analytic examination of optimism and hope. *Personality and Individual Differences, 54*(7), 821–827. https://doi.org/10.1016/j.paid.2012.12.004

Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (2001). *Intelligenz-Struktur-Test 2000 R [Intelligence-structure-test 2000 R]*. Hogrefe.

Ariel, R., Lembeck, N. A., Moffat, S., & Hertzog, C. (2018). Are there sex differences in confidence and metacognitive monitoring accuracy for everyday, academic, and psychometrically measured spatial ability? *Intelligence, 70*, 42–51. https://doi.org/10.1016/j.intell.2018.08.001

Bendixen, L. D., Schraw, G., & Dunkle, M. E. (1998). Epistemic beliefs and moral reasoning. *The Journal of Psychology, 132*(2), 187–200. https://doi.org/10.1080/00223989809599158

Bensch, D., Paulhus, D. L., Stankov, L., & Ziegler, M. (2019). Teasing apart overclaiming, overconfidence, and socially desirable responding. *Assessment, 26*(3), 351–363. https://doi.org/10.1177/1073191117700268

Bing, M. N., Kluemper, D., Davison, H. K., Taylor, S., & Novicevic, M. (2011). Overclaiming as a measure of faking. *Organizational Behavior and Human Decision Processes, 116*(1), 148–162. https://doi.org/10.1016/j.obhdp.2011.05.006

Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *The Journal of Social Psychology, 130*(3), 353–363. https://doi.org/10.1080/00224545.1990.9924591

Bråten, I., Strømsø, H. I., & Ferguson, L. E. (2016). The role of epistemic beliefs in the comprehension of single and multiple texts. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text and context* (pp. 67–79). Routledge.

Buratti, S., Allwood, C. M., & Kleitman, S. (2013). First- and second-order metacognitive judgments of semantic memory reports: The influence of personality traits and cognitive styles. *Metacognition and Learning, 8*, 79–102. https://doi.org/10.1007/s11409-013-9096-5

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032–1053. https://doi.org/10.1111/1467-8624.00333

Carver, C. S., & Scheier, M. F. (2014). Dispositional optimism. *Trends in Cognitive Sciences, 18*(6), 293–299. https://doi.org/10.1016/j.tics.2014.02.003

Chemers, M. M., Hu, L. T., & Garcia, B. F. (2001). Academic self-efficacy and first year college student performance and adjustment. *Journal of Educational Psychology, 93*(1), 55–64. https://doi.org/10.1037/0022-0663.93.1.55

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786. https://doi.org/10.3758/BF03196772

Corso, H. V., Cromley, J. G., Sperb, T., & Salles, J. F. (2016). Modeling the relationship among reading comprehension, intelligence, socioeconomic status, and neuropsychological functions: The mediating role of executive functions. *Psychology & Neuroscience, 9*(1), 32. https://doi.org/10.1037/pne0000036

Dahl, M., Allwood, C. M., Rennemark, M., & Hagberg, B. (2010). The relation between personality and the realism in confidence judgements in older adults. *European Journal of Ageing, 7*, 283–291. https://doi.org/10.1007/s10433-010-0164-2

de Bruin, A. B. H., Rikers, R. M., & Schmidt, H. G. (2007). Improving metacomprehension accuracy and self-regulation in cognitive skill acquisition: The effect of learner expertise. *European Journal of Cognitive Psychology, 19*(4–5), 671–688. https://doi.org/10.1080/09541440701326204

de Bruin, A. B. H., Kok, E. M., Lobbestael, J., & de Grip, A. (2017). The impact of an online tool for monitoring and regulating learning at university: Overconfidence, learning strategy, and personality. *Metacognition and Learning, 12*, 21–43. https://doi.org/10.1007/s11409-016-9159-5

D'Lima, G. M., Winsler, A., & Kitsantas, A. (2014). Ethnic and gender differences in first-year college students' goal orientation, self-efficacy, and extrinsic and intrinsic motivation. *The Journal of Educational Research, 107*(5), 341–356. https://doi.org/10.1080/00220671.2013.823366

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

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. https://doi.org/10.1037/0022-3514.84.1.5

Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General, 116*(2), 119–136. https://doi.org/10.1037/0096-3445.116.2.119

Golke, S., & Wittwer, J. (2017). High-performing readers underestimate their text comprehension: Artifact or psychological reality? In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 2108–2113). Cognitive Science Society.

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition, 36*, 93–103. https://doi.org/10.3758/MC.36.1.93

Griffin, T. D., Jee, B. D., & Wiley, J. (2009). The effects of domain knowledge on metacomprehension accuracy. *Memory & Cognition, 37*(7), 1001–1013. https://doi.org/10.3758/MC.37.7.1001

Griffin, T. D., Wiley, J., & Salas, C. R. (2013). Supporting effective self-regulated learning: The critical role of monitoring. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 19–34). Springer. https://doi.org/10.1007/978-1-4419-5546-3_2

Guthrie, J. T., & Klauda, S. L. (2016). Engagement and motivational processes in reading. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text and context* (pp. 41–53). Routledge.

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology, 44*, 339–387. https://doi.org/10.1006/cogp.2001.0769

Händel, M., de Bruin, A. B. H., & Dresel, M. (2020). Individual differences in local and global metacognitive judgments. *Metacognition and Learning, 15*, 51–75. https://doi.org/10.1007/s11409-020-09220-0

Hannon, B. (2014). Are there gender differences in the cognitive components of adult reading comprehension? *Learning and Individual Differences, 32*, 69–79. https://doi.org/10.1016/j.lindif.2014.03.017

Hyde, J. S., & McKinley, N. M. (1997). Gender differences in cognition: Results from meta-analyses. In P. J. Caplan, M. Crawford, J. S. Hyde, & J. R. E. Richardson (Eds.), *Gender differences in human cognition* (pp. 30–51). Oxford University Press.

Ikeda, K., & Kitagami, S. (2012). The effect of working memory capacity and mental effort on monitoring accuracy in text comprehension. *Psychologia, 55*(3), 184–193. https://doi.org/10.2117/psysoc.2012.184

Ikeda, K., Yue, C. L., Murayama, K., & Castel, A. D. (2016). Achievement goals affect metacognitive judgments. *Motivation Science, 2*(4), 199–219. https://doi.org/10.1037/mot0000047

Judge, T. A., & Ilies, R. (2002). Relationship of personality to performance motivation: A meta-analytic review. *Journal of Applied Psychology, 87*(4), 797–807. https://doi.org/10.1037/0021-9010.87.4.797

Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard–easy effect. *Psychological Review, 107*(2), 384–396. https://doi.org/10.1037/0033-295X.107.2.384

Kelemen, W. L., Frost, P. J., Weaver, C. A., & III. (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. *Memory & Cognition, 28*, 92–107. https://doi.org/10.3758/BF03211579

Kendeou, P., & O'Brien, E. J. (2016). Prior knowledge: Acquisition and revision. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 151–163). Routledge. https://doi.org/10.4324/9780203075562

King, B. A., & Magun-Jackson, S. (2009). Epistemological beliefs of engineering students. *Journal of Technology Studies, 35*(2), 56–64.

Kleitman, S. (2008). *Metacognition in the rationality debate: Self-confidence and its calibration*. VDM Verlag Dr.

Kleitman, S., Hui, J. S., & Jiang, Y. (2019). Confidence to spare: Individual differences in cognitive and metacognitive arrogance and competence. *Metacognition and Learning, 14*, 479–508. https://doi.org/10.1007/s11409-019-09210-x

Klopp, E. (2014). *Die Struktur epistemologischer Überzeugungen: Empirische und theoretische Analysen* [The structure of epistemological beliefs: Empirical and theoretical analyses] (Doctoral dissertation, Saarland University). https://publikationen.sulb.uni-saarland.de/handle/20.500.11880/23435

Komori, M. (2016). Effects of working memory capacity on metacognitive monitoring: A study of group differences using a listening span test. *Frontiers in Psychology, 7*, 285. https://doi.org/10.3389/fpsyg.2016.00285

Koriat, A. (1997). Monitoring one's knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Kroll, M. D., & Ford, M. L. (1992). The illusion of knowing, error detection, and motivational orientations. *Contemporary Educational Psychology, 17*(4), 371–378. https://doi.org/10.1016/0361-476X(92)90075-A

Kwon, H., & Linderholm, T. (2014). Effects of self-perception of reading skill on absolute accuracy of metacomprehension judgements. *Current Psychology, 33*(1), 73–88. https://doi.org/10.1007/s12144-013-9198-x

Lin, L. M., & Zabrucky, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology, 23*(4), 345–391. https://doi.org/10.1006/ceps.1998.0972

Lundeberg, M., & Mohan, L. (2009). Context matters: Gender and cross-cultural differences in confidence. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 221–239). Routledge.

Maki, R. H. (1998). Metacomprehension of text: Influence of absolute confidence level on bias and accuracy. In D. L. Medin (Ed.), *The psychology of learning and motivation*. (Vol. 38, pp. 223–248). Academic Press.

Maki, R. H., & McGuire, M. J. (2002). Metacognition for text: Findings and implications for education. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (pp. 39–67). Cambridge University Press.

Maki, R. H., & Serra, M. (1992). Role of practice tests in the accuracy of test predictions on text material. *Journal of Educational Psychology, 84*(2), 200–210. https://doi.org/10.1037/0022-0663.84.2.200

Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology, 97*(4), 723–731. https://doi.org/10.1037/0022-0663.97.4.723

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology, 81*(3), 417–430. https://doi.org/10.1037/0022-0663.81.3.417

Mason, L., Boldrin, A., & Ariasi, N. (2010). Epistemic metacognition in context: Evaluating and learning online information. *Metacognition and Learning, 5*(1), 67–90. https://doi.org/10.1007/s11409-009-9048-2

Miller, L. M. S., Cohen, J. A., & Wingfield, A. (2006). Contextual knowledge reduces demands on working memory during reading. *Memory & Cognition, 34*(6), 1355–1367. https://doi.org/10.3758/BF03193277

Moore, D., Lin-Agler, L. M., & Zabrucky, K. M. (2005). A source of metacomprehension inaccuracy. *Reading Psychology, 26*(3), 251–265. https://doi.org/10.1080/02702710590962578

Moutafi, J., Furnham, A., & Crump, J. (2006). What facets of openness and conscientiousness predict fluid intelligence score? *Learning and Individual Differences, 16*(1), 31–42. https://doi.org/10.1016/j.lindif.2005.06.003

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). Academic Press.

Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H. M. (2005). Working memory and intelligence – their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*(1), 61–65. https://doi.org/10.1037/0033-2909.131.1.61

Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: A meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning, 13*, 179–212. https://doi.org/10.1007/s11409-018-9183-8

O'Reilly, T., Sabatini, J., & Wang, Z. (2019). What you don't know won't hurt you, unless you don't know you're wrong. *Reading Psychology, 40*(7), 638–677. https://doi.org/10.1080/02702711.2019.1658668

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods, 47*(4), 1343–1355. https://doi.org/10.3758/s13428-014-0543-2

Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. *The Journal of General Psychology, 129*, 257–299. https://doi.org/10.1080/00221300209602099

Pierce, B. H., & Smith, S. M. (2001). The postdiction superiority effect in metacomprehension of text. *Memory & Cognition, 29*(1), 62–67. https://doi.org/10.3758/BF03195741

Pintrich, P. R., Smith, D., Garcia, T., & McKeachie, W. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. The University of Michigan.

Prinz, A., Golke, S., & Wittwer, J. (2020a). How accurately can learners discriminate their comprehension of texts? A comprehensive meta-analysis on relative metacomprehension accuracy and

influencing factors. *Educational Research Review, 31*, 100358. https://doi.org/10.1016/j.edurev.2020.100358

Prinz, A., Golke, S., & Wittwer, J. (2020b). To what extent do situation-model-approach interventions improve relative metacomprehension accuracy? Meta-analytic insights. *Educational Psychology Review, 32*(4), 917-949. https://doi.org/10.1007/s10648-020-09558-6

Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K): Entwicklung und Validierung eines ökonomischen Inventars zur Erfassung der fünf Faktoren der Persönlichkeit [Short scale of the Big Five Inventory]. *Diagnostica, 51*(4), 195–206. https://doi.org/10.1026/0012-1924.51.4.195

Rand, K. L., Martin, A. D., & Shea, A. M. (2011). Hope, but not optimism, predicts academic performance of law students beyond previous academic achievement. *Journal of Research in Personality, 45*(6), 683–686. https://doi.org/10.1016/j.jrp.2011.08.004

Richter, T., & van Holt, N. (2005). ELVES: Ein computergestütztes Diagnostikum zur Erfassung der Effizienz von Teilprozessen des Leseverstehens [A computer-based instrument to assess efficiency of reading processes]. *Diagnostica, 51*(4), 169–182. https://doi.org/10.1026/0012-1924.51.4.169

Rummel, J., Steindorf, L., Marevic, I., & Danner, D. (2017). A validation study of the German complex-span tasks and some general considerations on task translation procedures in cognitive psychology. *European Journal of Psychological Assessment*, 1–12. https://doi.org/10.1027/1015-5759/a000444

Schaefer, P. S., Williams, C. C., Goodie, A. S., & Campbell, W. K. (2004). Overconfidence and the big five. *Journal of Research in Personality, 38*, 473–480. https://doi.org/10.1016/j.jrp.2003.09.010

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology, 67*(6), 1063–1078. https://doi.org/10.1037/0022-3514.67.6.1063

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*, 498–504. https://doi.org/10.1037/0022-0663.82.3.498

Schraw, G. (1994). The effect of metacognitive knowledge on local and global monitoring. *Contemporary Educational Psychology, 19*(2), 143–154. https://doi.org/10.1006/ceps.1994.1013

Schraw, G. (1997). The effect of generalized metacognitive knowledge on test performance and confidence judgments. *The Journal of Experimental Education, 65*(2), 135–146. https://doi.org/10.1080/00220973.1997.9943788

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Sharpe, J. P., Martin, N. R., & Roth, K. A. (2011). Optimism and the big five factors of personality: Beyond neuroticism and extraversion. *Personality and Individual Differences, 51*, 946–951. https://doi.org/10.1016/j.paid.2011.07.033

Shepperd, J. A., Klein, W. M. P., Waters, E. A., & Weinstein, N. D. (2013). Taking stock of unrealistic optimism. *Perspectives on Psychological Science, 8*(4), 395–411. https://doi.org/10.1177/1745691613485247

Spinath, B., Stiensmeier-Pelster, J., Schöne, C., & Dickhäuser, O. (2002). *Die Skalen zur Erfassung von Lern- und Leistungsmotivation (SELLMO) [Measurement scales for learning and performance motivation]*. Hogrefe.

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. https://doi.org/10.2190/1266-0413-387K-7J51

Stankov, L., & Lee, J. (2008). Confidence and cognitive test performance. *Journal of Educational Psychology, 100*(4), 961. https://doi.org/10.1037/a0012546

Sternberg, R. (1986). Inside intelligence: Cognitive science enables us to go beyond intelligence tests and understand how the human mind solves problems. *American Scientist, 74*(2), 137–143.

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). Routledge.

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 832–842. https://doi.org/10.1037/a0029669

van den Broek, P. W., Bohn-Gettler, C. M., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. T.

McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 123–139). IAP Information Age Publishing.

van den Broek, P. W., Virtue, S., Everson, M. G., Tzeng, Y., & Sung, Y.-C. (2002). Comprehension and memory of science texts: Inferential processes and the construction of a mental representation. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 131–154). Routledge.

van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality, 44*(3), 315–327. https://doi.org/10.1016/j.jrp.2010.03.003

van der Stel, M., & Veenman, M. V. J. (2014). Metacognitive skills and intellectual ability of young adolescents: A longitudinal study from a developmental perspective. *European Journal of Psychology of Education, 29*, 117–137. https://doi.org/10.1007/s10212-013-0190-5

Veenman, M. V. J. (2016). Metacognition. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 26–40). Routledge. https://doi.org/10.4324/9780203075562

Veenman, M. V. J., Kok, R., & Blöte, A. W. (2005). The relation between intellectual and metacognitive skills in early adolescence. *Instructional Science, 33*, 193–211. https://doi.org/10.1007/s11251-004-2274-8

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. https://doi.org/10.1016/j.chb.2015.07.066

Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *The Journal of General Psychology, 132*(4), 408–428. https://doi.org/10.3200/GENP.132.4.408-428

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

Zabrucky, K. M., Moore, D., Agler, L. M. L., & Cummings, A. M. (2015). Students' metacomprehension knowledge: Components that predict comprehension performance. *Reading Psychology, 36*(7), 627–642. https://doi.org/10.1080/02702711.2014.950536

Zhou, M. (2013). University student's goal profiles and metacomprehension accuracy. *Educational Psychology: An International Journal of Experimental Educational Psychology, 33*(1), 1–13. https://doi.org/10.1080/01443410.2012.730325

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299–315). Routledge.

Zwaan, R. A., & Rapp, D. N. (2006). Discourse comprehension. In M. J. Traxler, & M. A. Gernsbacher (Eds.). *Handbook of psycholinguistics* (2nd ed., pp. 725–764). Academic Press.

Zweig, D., & Webster, J. (2004). What are we measuring? An examination of the relationships between the big-five personality traits, goal orientation, and performance intentions. *Personality and Individual Differences, 36*(7), 1693–1708. https://doi.org/10.1016/j.paid.2003.07.010