Metacognitive Comprehension Monitoring: Cognitive Abilities Explain Performance Differences between Younger and Older Adults

Catharina Tibken[1,2], Tobias Richter[1,2], Wienke Wannagat[1]

[1]Department of Psychology, University of Würzburg

[2]Counseling Center for the Gifted and Talented, University of Würzburg

Accepted for publication in *Scientific Studies of Reading* (2023)

Author Note

Catharina Tibken https://orcid.org/0000-0003-4058-1332

Tobias Richter https://orcid.org/0000-0002-0467-9044

Wienke Wannagat https://orcid.org/0000-0002-1572-2474

The research reported in this article was supported by grants awarded by the Karg Foundation, the Faculty of Human Sciences of the University of Würzburg, and the Department of Psychology of the University of Würzburg.

Testing materials are available from the first and the second author upon request. The complete dataset and analysis scripts are available in the repository of the Open Science Framework (https://osf.io/q58ad/?view_only=debfaf4a57894e48a41fd9dafe77c9e5). The authors report there are no competing interests to declare.

Correspondence concerning this article should be addressed to Catharina Tibken, University of Würzburg, Department of Psychology, Röntgenring 10, 97070 Würzburg, Germany.

Email: catharina.tibken@uni-wuerzburg.de

Abstract

Purpose:

To understand complex expository text, readers often engage in metacognitive comprehension monitoring. Metacognitive monitoring is assumed to rely on basic cognitive abilities (working memory updating, short-term memory, verbal intelligence). These abilities decrease in later adulthood. We thus compared younger and older adults in their comprehension monitoring and examined whether performance differences are mediated by differences in cognitive abilities.

Methods:

Younger (n = 101; 18–29 years) and older adults (n = 108; 60–75 years) completed an inconsistency task to capture metacognitive comprehension monitoring, two tests of working memory updating (one based on a semantic and one on a formal criterion), a short-term memory test, and an indicator of verbal intelligence.

Results:

Older adults reported fewer inconsistencies than younger adults ($\beta = -.174$, $p = .009$). These differences were mediated by differences in working memory updating, short-term memory, and verbal intelligence. Working memory updating based on a semantic criterion, in contrast to working memory updating based on a formal criterion, was especially related to performance differences in the inconsistency task ($\beta = .299$).

Conclusion:

The present study extends previous results on the role of basic cognitive abilities for explaining differences between age groups in metacognitive comprehension monitoring in younger compared to older adults.

Keywords: comprehension monitoring, working memory updating, older adults, age effects

Metacognitive Comprehension Monitoring: Cognitive Abilities Explain Performance Differences Between Younger and Older Adults

The ability to comprehend written text is essential for coping with everyday life, for social participation, and for independent and self-determined living at any age. Age-associated declines in basic cognitive abilities such as processing speed or memory capacity in late adulthood (Salthouse, 1996) can also lead to difficulties in reading comprehension, especially with complex texts (De Beni et al., 2007). Such difficulties are particularly relevant when texts contain important information, for example, on health issues, or political discussions (e.g., Bann et al., 2006). In such cases, readers need to be aware of their difficulties in understanding (metacognitive comprehension monitoring) to be able to take countermeasures (metacognitive comprehension regulation), for example, by obtaining additional information or rereading passages that have been identified as difficult (Baker, 1989). Comprehension monitoring is thus an important metacognitive competence in the domain of reading comprehension in younger and older adults.

In the following, we first provide a brief introduction into (meta-)cognitive processes that are relevant for text comprehension, with a focus on comprehension monitoring in reading and learning from expository texts. Second, we provide an overview of research on how younger and older adults differ in their metacognitive comprehension monitoring performance, which raises the question of whether such differences can be explained by differences in basic cognitive abilities. Third, we discuss literature on the role of various basic cognitive abilities such as working memory updating and short-term memory capacity for metacognitive comprehension monitoring that might decline with age and discuss whether these abilities must be domain-specific for reading comprehension (semantic processing of verbal material). These considerations form the basis of our study, which examined performance differences between younger and older adults in metacognitive comprehension monitoring and the extent to which domain-specific basic cognitive abilities mediate the effect of age on metacognitive comprehension monitoring.

Metacognitive Monitoring in the Comprehension of (Expository) Texts

Text comprehension is a multifaceted cognitive process in which readers construct a coherent mental representation of the situation described in the text by integrating information from the text with their prior knowledge to construct a *situation model* (Kintsch, 1988; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). The construction of such a situation model is thought to involve both passive and active, reader-initiated processes in varying proportions, depending on a reader's reading abilities and current reading goals (van den Broek & Helder, 2017). For experienced readers like most adults who possess the relevant prior knowledge, the construction of a situation model is often a largely passive process that occurs without major cognitive effort. In his Construction-Integration Model, Kintsch (1988, 1998) described two alternating phases of text comprehension. In the construction phase, information from the text initially forms a rough representation of the content just read in working memory. This initial representation passively activates previously read information or world knowledge that is stored in short-term or long-term memory. This process of passively spreading activation of memory content is also referred to as *memory resonance* (Myers & O'Brien, 1998). The currently read information is integrated with the activated prior knowledge to form a coherent mental representation by, for example, deleting irrelevant information and validating the currently processed information against prior knowledge (Kintsch, 1998; O'Brien & Cook, 2016; Richter, 2015; Singer, 2006).

To explain individual differences in reading comprehension, component skills models have been proposed that decompose reading comprehension into several sub-skills. An influential model is the simple view of reading, which regards reading comprehension as a product of decoding and language comprehension (Hoover & Gough, 1990). More recent models such as the Direct-and-Inferential-Mediation-Model-of-Reading-Comprehension

(DIME; Cromley & Alzevedo, 2007) and the Direct-and-Indirect-Effect-Model-of-Reading (DIER; Kim, 2015) further divide these two primal skills into sub-skills. In studies on the DIER model with students from primary schools, comprehension monitoring has been shown to be a predictor of language/listening comprehension. Differences in comprehension monitoring itself could be explained by differences in working memory capacity that were mediated by general language ability such as vocabulary and grammar knowledge (e.g., Kim, 2017). In these studies, comprehension monitoring was operationalized with a task to detect inconsistencies in short narrative texts which are presumably particularly suitable for capturing passive, automated processes of comprehension monitoring (Kim, 2015, 2017) such as those described in general models of text comprehension with the term validation (e.g., Richter, 2015; Singer, 2013).

However, readers may become aware of reading processes if their *standards of coherence*, that is, their reading goals, are high and the text is complex in relation to the readers' abilities and prior knowledge (van den Broek et al., 1995; van den Broek & Helder, 2017). In such situations (e.g., when reading a text as an exam preparation), readers tend to consciously monitor their text comprehension to notice possible difficulties and resolve them through cognitive effort, for example, by actively drawing inferences and filling in blanks in the situation model. Expository texts often require high standards of coherence and considered more demanding than narrative texts (Mar et al., 2021; Wannagat, et al., 2022) because they presuppose specific prior knowledge, and are characterized by more infrequent and difficult words and high density of potentially important information. Therefore, readers may need to engage in strategic metacognitive monitoring in addition to routine validation to build up an adequate situation model of an expository text (Tibken, Richter, Wannagat, et al., 2022). Metacognitive comprehension monitoring can thus be seen as a subdomain of metacognition (e.g., Schneider et al., 2022). The broader term metacognition generally includes thinking about one's own cognitions (Flavell, 1979). In early studies, the term mainly referenced memory research, but was later extended to other areas such as reading comprehension (Schneider et al., 2022).

The inconsistency task can be considered an established, standard procedure for capturing passive and metacognitive processes of reading comprehension. It was first introduced by Baker (1979) and has since then been used frequently in many variants (e.g., Helder et al., 2016; O'Brien & Albrecht, 1992; Tibken, Richter, von der Linden, et al., 2022). In this task, participants read text that contains inconsistencies, often information that contradicts other information from the same text or participants' knowledge of the world. O'Brien and Albrecht (1992), for example, used short narrative texts that contained contradictions that could be recognized without much effort (e.g., inconsistencies between protagonists' values and their behavior). By analyzing reading times, the authors examined mainly passive processes of memory resonance. Other studies that used inconsistency tasks used more demanding expository texts and more complex inconsistencies (e.g., contradictions within the description of a physical phenomenon such as the lift of airplanes) within the same paradigm to examine metacognitive processes in text comprehension (e.g., Tibken, Richter, Wannagat, et al., 2022).

From the readers' behavior during and after reading the inconsistent text, different measures can be drawn that contain information about their comprehension monitoring. Online measures such as reading times of inconsistent compared to consistent text passages or eye-movements are collected during reading, whereas offline measures, such as indicating whether the previously read text was inconsistent or consistent, are collected after reading. Studies indicate that the two sorts of measures of the inconsistency task might capture different aspects of comprehension monitoring (e.g., Tibken, Richter, Wannagat, et al., 2022). In correspondence with the Two-Step Model of Validation (Maier & Richter, 2017), slowing down in reading time when reading inconsistent compared to consistent text might reflect the mere (mostly passive) detection of an inconsistency, whereas reporting an inconsistency after reading involves a conscious processing of inconsistent information to some extent and an attempt to resolve the inconsistency. Thus, the report of inconsistencies after reading is assumed be the more valid indicator of strategic and thus metacognitive comprehension monitoring (Tibken, Richter, Wannagat, et al., 2022).

Performance Differences in Metacognitive Comprehension Monitoring Between Younger and Older Adults

Metacognitive processes, such as metacognitive comprehension monitoring, might change not only during childhood but also in later adulthood, similar to other cognitive abilities (e.g., Salthouse, 2010). Studies indicate that in later adulthood, especially the processing of non-lexical information decreases, whereas the processing of lexical information and measures of crystallized intelligence remain relatively stable into late adulthood (e.g., Hale & Myerson, 1995). Thus, older readers, more than younger readers, tend to rely on context and discourse information, presumably to compensate for lower working memory capacity and processing speed (for an overview see Stine-Morrow & Radvansky, 2018). These considerations are supported by findings of older adults having more difficulties than younger adults in constructing a mental representation of the text surface and the textbase during reading, whereas many studies did not find any differences in the construction of a situation model (Radvansky & Dijkstra, 2007). This may be due to older adults drawing more on their (more extensive) prior knowledge to construct a situation model. Older readers, for example, were less likely than younger readers to reproduce errors presented in a text that conflicted with world knowledge in a subsequent query (Umanath & Marsh, 2012). They also benefited more from headings that helped them activate prior knowledge in their reading comprehension, especially when their working memory capacity was low (Miller et al., 2006). Although this behavior, commonly found in older readers, is an effective way of dealing with declining basic cognitive abilities and is usually useful in everyday life, there are situations where over-reliance on prior knowledge and contextual information can impede comprehension. Older adults might be more likely to build an incorrect situation model if contextual information is missing or prior knowledge does not fit the subject matter of the text. These comprehension difficulties occur especially when texts are complex because, for example, several storylines have to be followed (Noh & Stine-Morrow, 2009). Difficulties in comprehension might therefore occur in everyday life particularly with complex expository texts, where gaps in comprehension may not always be covered by appropriate prior knowledge. Given high standards of coherence of the readers, metacognitive monitoring of reading comprehension is therefore particularly important in these cases.

Only few studies have compared younger and older adults in metacognitive comprehension monitoring measured with the inconsistency task and the results of these studies are inconclusive. Zabrucky and Moore (1991) found no differences between younger and older adults in the verbal report of inconsistencies within short narrative text passages. Other studies found that younger adults were more likely than older adults to report an inconsistency in an expository text if the information, which at first glance seemed contradictory, was not resolved within the text by explaining further details (Zabrucky et al., 1987). Two other studies by Zabrucky and Moore (1994; 1999) with expository texts, each involving 20 younger and 20 older adults, found no differences between younger and older adults in the verbal report of inconsistencies, whereas the results of a more recent study with a larger sample of 80 participants in each of the two age groups indicated that younger adults performed better in reporting inconsistencies than older adults (Zabrucky et al., 2012). In summary, the results suggest that no major differences exist between younger and older adults in performance on the inconsistency task. However, the samples in many of these studies were small, that is, the statistical power may not have been sufficient to consistently detect possible differences between age groups.

We therefore assumed differences between younger and older adults based on the results of studies that employed the inconsistency task to investigate age-related differences in metacognitive comprehension monitoring in children and adolescents. These studies consistently showed that older students report more inconsistencies than younger students (e.g., Hacker, 1997; Helder et al., 2016; Tibken, Richter, Wannagat et al., 2022). An explanation for these findings might be the advanced abilities of older students in actively processing and encoding the inconsistencies in memory. A study examining metacognitive comprehension monitoring of students in Grades 6 and 8 found that the better performance of older compared to younger adolescents in the report of inconsistencies was related to underlying cognitive abilities (reading ability, working memory updating, short-term memory capacity, and general cognitive ability) that develop with age (Tibken, Richter, Wannagat, et al., 2022). In childhood and adolescence, an increase in short-term memory capacity and general cognitive abilities can be observed, whereas these abilities slowly decline again in older adulthood (Alloway & Alloway, 2013). These findings suggest a similar trajectory for metacognitive comprehension monitoring ability, as assessed with the inconsistency task, from middle to older adulthood. Younger adults might report more inconsistencies than older adults. In the following, the role of relevant basic cognitive abilities in metacognitive comprehension monitoring is explained in more detail.

The Role of (Domain-Specific) Cognitive Abilities in Metacognitive Comprehension Monitoring

Research has consistently shown that working memory capacity is a predictor of reading comprehension in general (e.g., Cain et al., 2004; Carretti et al., 2009). In particular, the dynamic component of working memory that involves the manipulation of content currently available in working memory seems to be relevant for reading comprehension, more so than the temporary storage of content in short-term memory (Butterfuss & Kendeou, 2018; Daneman & Merikle, 1996). From a theoretical perspective, the ability to update and thus actively control working memory content can be considered a prerequisite for forming a coherent mental representation of the situation described in a text. With each new piece of information, the readers must integrate that piece of information into their current situation model and update their situation model if the information does not fit into the previous mental representation (Kurby & Zacks, 2012). Butterfuss and Kendeou (2018) argued that difficulties in updating working memory content might result in too much irrelevant information in working memory, and because of limited working memory capacity, potentially relevant information from long-term memory becomes less effectively retrieved and processed. It is assumed that older adults have greater difficulty in suppressing information that has become irrelevant in their situational model (Stine-Morrow & Radvansky, 2018).

Studies indicate that not only reading comprehension but also comprehension monitoring is supported by readers' ability to update working memory content (de Bruïne et al., 2021; Tibken, Richter, Wannagat, et al., 2022). De Bruïne et al. (2021) found that performance on the inconsistency task was impaired when memory was loaded by a secondary task (digit or letter span task), suggesting that comprehension monitoring requires available working memory capacity. Tibken, Richter, Wannagat, et al. (2022) accordingly found that working memory updating was associated with performance in the inconsistency task in adolescents. The need to (actively) compare new information from the text with information stored in working and long-term memory is especially required when readers encounter complex expository texts and follow high standards of coherence. Readers then need to resolve inconsistencies or store them in memory for later regulatory activities (e.g., for collecting further information from other sources). Such strategic encoding of information is likely to rely strongly on working memory capacity also in older adults (Cherry et al., 2021).

The processes of updating working memory content required for reading comprehension and comprehension monitoring are thought to be domain-specific, at least from the end of primary school (Daneman & Merikle, 1996; Peng et al., 2018). Pelegrina et al. (2015) found that children with difficulties in reading comprehension showed more problems in a word updating task (recalling the smallest objects of an auditorily presented list) than in a number updating task (recalling the smallest numbers of an auditorily presented list). A meta-analysis arrived at similar results, suggesting stronger effects of the ability to update verbal working memory than visuospatial working memory content on reading comprehension in children and young adults (Caretti et al., 2009). A sentence span task, compared to a digit span task, was associated more strongly with comprehension monitoring (Cain et al., 2004). However, especially for older children and adolescents, results are ambiguous. A meta-analysis (Peng et al., 2018), for example, reports no differences between verbal and numeric updating tasks with regard to reading comprehension for children after Grade 4. A study with younger adults even found no effect of working memory updating with either verbal or numerical material on reading comprehension (Freed et al., 2017).

One explanation for this pattern of results might be that most of these studies only compared the predictive power of different learning materials and thus different content categories in working memory (e.g., words, numbers, figures) for reading comprehension or comprehension monitoring. However, to the best of our knowledge, different criteria according to which working memory content should be updated have not been examined to date. Pelegrina et al. (2015), for example, compared two semantic criteria (recall of the smallest objects vs. the smallest numbers of a list), whereas Cain et al. (2004) compared two formal criteria (recall of each the last word vs. number in a list of sentences/digits). Similarly, Freed et al. (2017) used several updating tasks that contained a semantic task (e.g., evaluating sentences according to their correctness), but in which the actual updating task was still based on a formal criterion (remember an additional word after each sentence). Although many studies suggest that the ability to update verbal information is more relevant to reading comprehension and comprehension monitoring than updating numeric and visuospatial working memory content, the updating criterion used (semantic vs. formal) could additionally determine the extent to which differences in comprehension monitoring can be explained by working memory capacity. In line with Butterfuss and Kendeou (2018), when comprehension monitoring includes detecting and resolving semantic inconsistencies within a complex text or with prior knowledge, semantic (rather than formal) updating might be involved because content not only has to be manipulated in working memory, but for the updating process, semantic information must also be retrieved from long-term memory. Working memory updating based on a semantic criterion (e.g., retrieving the size of different objects or numbers from long-term memory) might reflect this process more accurately than updating based on a formal criterion. The comparison between updating based on semantic vs. formal criteria might thus represent a further dimension of domain-specificity in addition to the dimension of working memory content.

In addition to working memory updating, short-term memory capacity is a relevant predictor of metacognitive comprehension monitoring (e.g., Tibken, Richter, Wannagat, et al., 2022), although short-term memory capacity seems less important for reading comprehension in general (e.g., Cain et al., 2004). Short-term memory capacity reflects the ability to store a larger number of items simultaneously in working memory for further processing. Metacognitive comprehension monitoring, in contrast to passive monitoring processes, is mostly required when reading complex (expository) texts that contain many pieces of information that must be stored and integrated successively to build a comprehensive situation model. Short-term memory capacity might thus be a limiting factor for working memory updating capacity because it limits the number of items that can be processed simultaneously in working memory (Tibken, Richter, Wannagat, et al., 2022). Short-term memory capacity is also necessary to store ambiguous text passages in order to retrieve further information later if the inconsistency cannot be resolved directly (Tibken, Richter, Wannagat, et al., 2022).

Apart from working memory updating and short-term memory capacity, verbal intelligence affects performance in reading comprehension and comprehension monitoring. Verbal abilities in general appear to be relevant for reading comprehension (Cain et al., 2004).

A large and well-connected vocabulary, for example, is beneficial to perceive subtle linguistic nuances in a text and thus to retrieve relevant prior knowledge to integrate it into the situation model (e.g., Kim, 2017). Accordingly, higher verbal ability scores have been shown to correlate with better reading comprehension in children and older adults (e.g., Payne et al., 2012; Reynolds & Turek, 2012). Presumably more than reading comprehension in general, metacognitive comprehension monitoring relies on verbal intelligence, which includes logical reasoning based on verbal material. The reason for assuming a close connection between comprehension monitoring and verbal intelligence is that readers need to pay explicit attention to the coherence of the text (e.g., by comparing and weighing verbal information from different parts of the text). Moreover, when readers' verbal processing speed is high, more cognitive resources are available for metacognitive processes during reading. This speed component of verbal intelligence, in particular, decreases in later adulthood, whereas vocabulary remains relatively stable (Salthouse, 2010). Additionally, higher intelligence has been shown to be associated with greater working memory capacity (Conway et al., 2003; Engle et al., 1999). Therefore, verbal intelligence might indirectly affect comprehension monitoring.

Rationale of the Present Study

In this study, we examined performance differences between younger and older adults in the inconsistency task as a measure of metacognitive monitoring in reading comprehension. We also examined the extent to that these differences between age groups are mediated by differences in domain-specific underlying basic cognitive abilities. Previous research indicates that building a situation model and metacognitively monitoring the involved processes rely on basic cognitive abilities, especially on working memory updating (Butterfuss & Kendeou, 2018; Tibken, Richter, Wannagat, et al., 2022). Further cognitive abilities that support working memory updating and are relevant for (actively) storing ambiguous text passages in memory for later resolution are short-term memory capacity and intelligence (Cain et al., 2004; Tibken, Richter, Wannagat, et al., 2022).

Only few studies with relatively small sample sizes have examined performance differences between younger and older adults in the inconsistency task with ambiguous results (e.g., Zabrucky & Moore, 1991, 1994; Zabrucky et al., 2012). However, studies with children and adolescents have shown consistently that older students outperformed younger students regarding the report of inconsistencies in the inconsistency task (e.g., Helder et al., 2016). Thus, we also expected age differences between younger and older adults. A study that explained performance in the report of inconsistencies with underlying cognitive abilities such as working memory updating in adolescents suggests that the decline in basic cognitive abilities in late adulthood might also be associated with a lower performance in the inconsistency task (Tibken, Richter, Wannagat, et al., 2022).

Previous results are inconclusive as to the extent to which individual differences in updating verbal material in working memory explain more variance in comprehension monitoring than, for example, individual differences in updating numerical material in working memory (Cain et al., 2004; Caretti et al., 2009; Freed et al., 2017; Pelegrina et al., 2015; Peng et al., 2018). In addition to the type of information to be held in working memory, the type of updating process (semantic vs. formal) might thus also be relevant. Against this background, in a preliminary analysis, we compared working memory updating based on a semantic criterion with verbal material with working memory updating based on a formal criterion with numerical material to strengthen the argument for including a verbal updating measure with a semantic updating criterion in our main analyses. The critical factor was how strongly these two types of updating tasks were associated with performance in the inconsistency task.

We addressed our research questions by examining younger (18 to 29 years) and older (60 to 75 years) adults.

After the preliminary analysis regarding the two updating measures, we tested the following hypotheses:

1. Younger adults perform better in metacognitive monitoring in the domain of reading comprehension (report of inconsistencies in the inconsistency task) than older adults.
2. The effect of age group on the performance in the inconsistency task is mediated by differences in underlying cognitive abilities. More specifically, the differences in performance between younger and older adults in the report of inconsistencies can be explained by differences in working memory updating, which (in part) relies on short-term memory capacity and verbal intelligence, which are both assumed to decrease with age.

Method

Power and Required Sample Size

We conducted a power analysis with G*Power (Faul et al., 2009). We assumed small to medium effect sizes ($f^2 = .09$) for the incremental effects of working memory updating skills, short-term memory capacity, and verbal intelligence on the performance in the inconsistency task (based on results reported by Tibken, Richter, Wannagat, et al., 2022). The power analysis suggested a sample size of 97 participants in each age group to detect these effects (at $\alpha = .05$ and $1-\beta = .90$).

Sample

We examined 209 adults from two different age groups (102 males, 107 females) who had given their informed consent. The 101 younger adults were between 18 and 29 years old ($M = 22.36$ years, $SD = 2.34$), the 108 older adults were between 60 and 75 years old ($M = 65.42$ years; $SD = 4.61$). The gender distribution was not significantly different between the age groups, $\chi^2(1) = 0.56$, $p = .490$. In the younger age group, 19.80% of the participants held a university degree and 54.46% were university students at the time of data collection. In the older age group, 38.89% of the participants held a university degree. In both age groups, the proportion of students or participants with a university degree was thus higher than the average of the respective age groups in Germany (Statistisches Bundesamt, 2022). Similar, the percentage of participants with a Gymnasium degree (school leaving certificate of the academic-track high school) was higher than in the respective age groups in Germany in both age groups (91% of the younger adults compared to 40% of this age group in Germany and 52% of the older adults compared to 15% of this age group in Germany; Bundesministerium für Forschung und Bildung, 2023). All participants mastered the German language at native speaker level and none of them reported any diagnosis of learning or reading disability nor any documented sensory or neurological impairment relevant to this study. The participants received 15 euros for their participation.

Materials and Test Instruments

As a reliability measure of the scales, we provide McDonald's ω in the following because it estimates the internal consistency of dichotomous scales better than Cronbach's α (Trizano-Hermosilla & Alvarado, 2016).

Inconsistency Task

We used the inconsistency task introduced by Tibken, Richter, von der Linden, et al. (2022) to capture metacognitive monitoring in the domain of reading comprehension. For this task, participants read nine short expository texts (92 to 98 words) on a computer screen. Each text described a sequence of events, for example, a craft instruction or a description of a natural phenomenon. The texts covered various content domains to minimize effects of prior knowledge on task performance.

We created an inconsistent and a consistent version of each text that differed only in one word. The critical word resulted in two pieces of information contradicting each other within the inconsistent text version. Thus, no specific prior knowledge was required to detect the inconsistencies. For example, a craft instruction described how to fold a paper such that it appeared to become smaller (consistent version) or larger (inconsistent version) as a result of the folding. The inconsistencies were located at different positions in the texts so that participants could not form expectations about where an inconsistency might occur.

The frequencies (based on the DeReWo database; Institut für deutsche Sprache, 2009) of the critical words in German were not significantly different between the consistent and inconsistent text versions in our study, $t(16) = 0.20$, $p = .848$. Thus, assuming that word frequencies in German texts indicate familiarity, the critical words can be considered to be equally familiar to adult German native speakers. A sample text is in Appendix 1. All texts are available in the Open Science Framework repository (https://osf.io/q58ad/?view_only=debfaf4a57894e48a41fd9dafe77c9e5).

The task began with an inconsistent sample text to familiarize participants with the procedure of the task at the computer. Participants were instructed that inconsistencies could occur in the texts and that they would be asked about inconsistencies at the end of each text. If a participant failed to recognize the inconsistency in the sample text, an automatic correction was displayed on the screen with a brief explanation to ensure that all participants understood what was meant by inconsistency. After the sample text, the participants read the nine texts successively. The texts were presented in random order for each participant. Also randomly, six texts were presented in the inconsistent version and three texts in the consistent version. Participants read the sentences successively on a computer screen while the rest of the text was masked out (self-paced reading). After each text, participants indicated whether they had detected an inconsistency. If yes, the entire text was presented again, and participants marked the sentence in which they thought the inconsistency had occurred. The number of detected inconsistencies served as an indicator of the extent to which inconsistencies were (consciously) processed and stored in memory.

Working Memory Updating

We used two measures of working memory updating: a task to measure updating based on a semantic criterion and with verbal material and a task to measure updating based on a formal criterion and with numeric material.

Working memory updating based on a semantic criterion was measured with a word-updating task (the same as in Tibken et al., 2022), adapted from Palladino et al. (2001) and Pelegrina et al. (2015). Participants were presented auditorily with six word lists, each containing 12 words that named objects of the same shape but of different sizes (e.g., elongated objects such as ”rice grain”, “toothpick”, or “pencil”). After the presentation of each list, the participants were asked to recall the three smallest objects. Given that participants had no foreknowledge of whether subsequent objects in the list would be smaller or larger, they needed to memorize the current smallest objects and compare them to the size of each new object and replace one of the currently stored objects with a new, smaller one if necessary (i.e., update their working memory based on a semantic criterion). The lists required between two and four updating operations. An updating operation means that out of the three currently stored smallest objects one object had to be replaced by a new one. Participants could obtain one point for each correctly recalled object (max. 18 points). In a pilot study, we ensured that the different objects' sizes were reliably distinguishable. In the sample of the present study, McDonald's ω based on the six word lists was .72.

Updating based on a formal criterion was measured with a task that was similar to an *n*-back task (see Jaeggi et al., 2010). Participants were presented auditorily with 11 lists that contained between 6 and 13 digits (numbers 1–9). The participants' task was, for each list, to name the n^{th} last digit, progressing from the second last digit in the first two lists to the sixth last digit in the final list. Thus, to name the correct digit, participants needed to continuously replace their working memory content with the digit that currently held the n^{th} last position of the currently presented digits (i.e., update their working memory based on a formal criterion).

For each correctly named digit, participants obtained one point (max. 11 points). In the sample of the present study, McDonald's ω based on the 11 lists was .56.

Short-Term Memory Capacity

Short-term memory capacity was assessed with a digit span task, adapted from the task in the Wechsler Adult Intelligence Scale – Fourth Edition (Wechsler, 2008; subtest digit span forward). The participants listened to 16 lists of digits. After each list, they were asked to recall the presented digits in their exact order of presentation. The first list consisted of three digits. In the course of the task, the digit spans became progressively longer, with the final list consisting of 10 digits. For each correctly recalled list, participants received one point (max. 16 points). In the sample of the present study, McDonald's ω based on the 16 lists was .73.

Verbal Intelligence

We used the subtest Wortanalogien (word analogies) of the Kognitiver Fähigkeitstest 4–12 (Cognitive Ability Test for Grade 4 to12; Heller & Perleth, 2000) as an indicator of verbal intelligence. This paper-and-pencil test consisted of 44 tasks in which participants had to select the one word (out of five words) that was related to a particular word in the same way that a particular pair of words were related to each other (e.g., *swimming* is to *water* as *flying* is to *?*). The test had a time limit of 7 minutes. For each correct answer, the participants received one point (max. 44 points). The test thus included both a vocabulary and a processing speed component as well as logical reasoning based on verbal material. Originally, the test was intended for students from Grade 4 to 12 (approximately corresponding to an age of 10 to 18 years). Given the heterogeneity within our sample, we considered this test suitable because it presumably differentiates within a broad range of abilities. Furthermore, the test took approximately 10 min to complete and thus provided us with an efficient option to obtain an indicator of verbal intelligence. In the sample of the present study, McDonald's ω based on the 44 items was .85.

Procedure

All tests were administered individually in a single appointment. After answering a demographic questionnaire, participants completed the different tasks, starting with the working memory updating task based on a formal criterion, followed by the inconsistency task. After a 5-min break, participants completed the working memory updating task based on a semantic criterion, then the subtest to assess verbal intelligence, and finally the test of short-term memory capacity. All tasks, except the verbal intelligence subtest, were presented on the computer. The participants wore headphones for the auditorily presented tasks (the two working memory updating tasks and the test of short-term memory capacity). The volume of auditory stimulus presentation was individually adjusted to the hearing performance of each participant beforehand.

Results

Table 1 provides an overview of the descriptive statistics of all variables relevant for the research questions of this study (separately for the total sample, younger adults, and older adults). Bivariate correlations between all measures, the complete data set, and the analysis scripts are available in the repository of the Open Science Framework (https://osf.io/q58ad/?view_only=debfaf4a57894e48a41fd9dafe77c9e5).

To test our hypotheses, we conducted path analyses with Mplus 7 (Muthén & Muthén, 2012). This approach allows for the analysis of mediation effects and thus, based on theoretical considerations, the explanation of differences between age groups in metacognitive comprehension monitoring through differences in basic cognitive abilities between the two age groups (Hypothesis 2). In addition to direct effects between single variables, for example, between age group and short-term memory, conclusions can be drawn about indirect effects, for example, to what extent the effect of age group on the number of reported inconsistencies are mediated by differences in memory performance. All variables were included in the models as manifest variables. The dataset was complete with respect to all variables of interest and the testing of each subject was independent (no nested data structure). The Type I error probability was set at .05 in all significance tests (two-tailed) unless otherwise stated. We report standardized coefficients for all path analyses.

Preliminary Analysis: Relevance of Working Memory Updating Measures for Metacognitive Comprehension Monitoring

As a preliminary analysis, we examined the domain-specificity of the relevance of working memory updating for metacognitive monitoring in the domain of reading comprehension. More specifically, we compared the relevance of working memory updating based on a semantic criterion with working memory updating based on a formal criterion for comprehension monitoring measured via the number of detected inconsistencies in an inconsistency task. The model is depicted in Figure 1. Working memory updating based on a semantic criterion and working memory updating based on a formal criterion were moderately correlated ($\beta = .229$, $p < .001$). Working memory updating based on a semantic criterion was significantly related to the number of detected inconsistencies ($\beta = .299$, $p < .001$), whereas we found no evidence that working memory updating based on a formal criterion was related to the number of detected inconsistencies ($\beta = .027$, $p = .691$). The model explained 9.4% of the variance in the number of detected inconsistencies[1].

Differences between Younger and Older Adults in Metacognitive Comprehension Monitoring

The main goal of this study was to analyze performance differences in comprehension monitoring between younger and older adults and to examine whether these differences are mediated through differences in (domain-specific) cognitive abilities. In a first step, we

[1] Because the updating tasks differed in their reliabilities, we added an attenuation correction to the model in a further step to rule out the possibility that the findings could be due to lower reliability in the *n*-back task. In this approach, the measurement errors of the two predictors are explicitly modelled based on their reliabilities and variances (e.g., DeShon, 1998). This model produced the same pattern of results for the effects of working memory updating based on a semantic criterion ($\beta = .355$, $p < .001$) and working memory updating based on a formal criterion ($\beta = .014$, $p = .893$) on the number of detected inconsistencies as the uncorrected model. Therefore, we conclude that the findings regarding the two updating tasks are not due to reliability issues.

examined whether younger adults detected more inconsistencies in the inconsistency task than older adults (Hypothesis 1). In the following models, age group was effect-coded (-1: younger adults, 1: older adults). The results showed a moderate effect of age group on the number of detected inconsistencies (β = -.174, p = .009) with younger adults detecting more inconsistencies (M = 3.41, SD = 1.48) than older adults (M = 2.86, SD = 1.60).

The Role of (Domain-Specific) Cognitive Abilities

In the next step, we analyzed whether the effect of age group on the performance in the inconsistency task was mediated by differences in (domain-specific) cognitive abilities (working memory updating based on semantic criterion, short-term memory capacity, and verbal intelligence; Hypothesis 2). To this end, we estimated a model with the number of detected inconsistencies as criterion and age group as predictor. We hypothesized that the effect of age group would be mediated by verbal intelligence and short-term memory capacity, which in turn should have a direct effect on the number of detected inconsistencies and an indirect effect via working memory updating based on a semantic criterion. The results of the model estimation can be found in Figure 2.

The two age group differed significantly in verbal intelligence (β = -.396, p < .001) and short-term memory capacity (β = -.295, p < .001), with younger adults showing better performance in these two abilities. We also found direct positive effects of verbal intelligence (β = .316, p < .001) and short-term memory capacity (β = .119, p = .043; one-tailed) on the number of detected inconsistencies. In addition, we found positive effects of verbal intelligence (β = .363, p < .001) and short-term memory capacity (β = .195, p = .003) on working memory updating based on a semantic criterion. The assumed positive effect of working memory updating based on a semantic criterion on the number of detected inconsistencies was also significant (β = .126, p = .036; one-tailed).

We found no direct effect of age group on the number of detected inconsistencies when considering working memory updating based on a semantic criterion, short-term memory capacity and, verbal intelligence in the model ($\beta = .013$, $p = .853$). Instead, the results showed a moderate total indirect effect of age group on the number of detected inconsistencies ($\beta = -.187$, $p < .001$). The effect of age group on the number of detected inconsistencies was thus entirely mediated via (domain-specific) cognitive abilities that decrease with age. This model explained 20.1% of variance in the number of detected inconsistencies.

Discussion

This study had two objectives. First, we compared younger (16–29 years) and older adults (60–75 years) in their comprehension monitoring of expository texts (Hypothesis 1). Second, we examined whether working memory updating, short-term memory capacity, and verbal intelligence mediated differences in performance between these two age groups (Hypothesis 2). In addition, as a preliminary analysis, we compared the effect of working memory updating with verbal material based on a semantic criterion on comprehension monitoring with the effect of working memory updating with numerical material based on a formal criterion.

In correspondence with Hypothesis 1, younger adults were able to report more inconsistencies than older adults. These findings are in line with findings by Zabrucky et al. (2012) who, like the present study, examined a larger sample and who found that younger adults showed better abilities in reporting inconsistencies after reading than older adults. This seems to apply at least to metacognitive comprehension monitoring (as compared to automated monitoring processes), which becomes more relevant when reading complex texts such as the ones used in this study. Furthermore, previous research has suggested an increase of metacognitive monitoring abilities across childhood and adolescence (e.g., Helder et al., 2016, Tibken, Richter, Wannagat, et al., 2022). Thus, the current study contributes to the understanding of metacognitive monitoring development across the lifespan.

In correspondence with Hypothesis 2, the effect of age group on the number of detected inconsistencies was mediated by basic cognitive abilities (working memory updating, short-term memory capacity, and verbal intelligence), all of which decrease with age (e.g., Salthouse, 2010). In our model, lower verbal intelligence and short-term memory capacity in older adults were associated with a lower performance in working memory updating and consequently with a lower performance in metacognitive comprehension monitoring. Comparable to adolescents (Tibken, Richter, Wannagat, et al., 2022), differences in performance between younger and older adults in number of reported inconsistencies were fully mediated by differences in basic cognitive abilities. The variance in number of reported inconsistencies that could be explained by basic cognitive abilities was also quite similar to the results of adolescents (approx. 20% of variance). These finding suggest that the development of metacognitive comprehension monitoring during later adulthood is opposite to that in childhood and adolescence, albeit with subtle differences.

Our results extend previous findings by providing evidence that the effect of short-term memory, as the static component of working memory, on metacognitive comprehension monitoring is partially mediated by the dynamic component, as was suggested by Butterfuss and Kendeou (2018) with regard to reading comprehension in general. An explanation could be that during the continuous updating of the situation model in the reading process, information that has become irrelevant has to be suppressed by the readers (Caretti et al., 2004). With insufficient suppression, the capacity for further updating processes could quickly be exhausted if the capacity of the readers' short-term memory is low. Research on children with reading difficulties came to similar conclusions (Swanson et al., 2006). Swanson et al. (2006) argued that working memory capacity is determined by both short-term memory (storage) and updating processes. The storage component in their study was not specific to verbal material and reading comprehension. The authors thus propose a basic storage capacity of working memory that is domain-general in addition to reading-specific aspects. Our study expands this argumentation to adults with unimpaired reading skills. Short-term memory capacity might be particularly relevant in comprehension monitoring when the demands exceed the readers' abilities. This might be the case with older adults, who often have a relatively low short-term memory capacity, and for demanding and complex texts with a lot of information to be updated and integrated.

In addition, the findings of this study add to the knowledge of reading comprehension in older adults the aspect of metacognitive comprehension monitoring that is particularly relevant when reading complex texts with high standards of coherence. Metacognitive monitoring is particularly susceptible to decline in older age because it relies more strongly than automated aspects of reading on basic cognitive abilities and cannot be compensated for by drawing on contextual information alone (Stine-Morrow & Radvansky, 2018).

In contrast to findings with adolescents (Tibken, Richter, Wannagat, et al., 2022), short-term memory and verbal intelligence had not only indirect effects on the number of reported inconsistencies via working memory updating but also direct effects. The direct effect of verbal intelligence on metacognitive comprehension monitoring was especially substantial, whereas the direct effect of short-term memory capacity was rather small. A possible explanation might lie in the higher heterogeneity of the sample in the current study. Our sample included all types of educational degrees, but a higher percentage of participants had higher education degrees than in the German population of the corresponding age groups. In contrast, the adolescents in the study by Tibken, Richter, Wannagat, et al. (2022) were all attending the academic track of secondary school during the study and the variation in age was much smaller. Thus, the individual differences in verbal intelligence were supposedly larger in the present study. In addition, the percentage of participants with a gymnasium degree was higher in the younger age group than in the older age group, so that differences between the age groups in their educational level cannot be completely ruled out. However, the percentages in relation to each other corresponded to the population shares in the respective age group, so that relevant distortions of the results are unlikely.

Another explanation could be that participants with high verbal intelligence might be better able to compensate for an age-related decrease in basic cognitive abilities and thus reading comprehension, given that verbal intelligence as part of crystallized intelligence is also an indicator of world knowledge and thus increases the likelihood of having relevant contextual information even for complex topics (see Stine-Morrow & Radvansky, 2018). A study with older adults showed that the more time older participants regularly spent in reading activities, the better they were able to compensate for a decline in their working memory on a sentence recall test (Payne et al., 2012). Thus, verbal intelligence and education might be more important for reading comprehension and comprehension monitoring in older adults than in children and adolescents.

The differential findings for our two updating tasks in the preliminary analysis raise interesting questions for future research on the debate on whether the ability to update verbal content in working memory is more relevant for reading comprehension and metacognitive comprehension monitoring than the ability to update numerical or visuospatial content (e.g., Cain et al., 2004; Caretti et al., 2009; Pelegrina et al., 2015). Our preliminary analysis revealed a small to medium effect of working memory updating with verbal material based on a semantic criterion on comprehension monitoring but no significant effect of working memory updating with numerical material based on a formal criterion. Thus, the present study indicates that the type of content (verbal vs. numerical) might only be one relevant dimension when examining the effect of working memory updating on comprehension monitoring or reading comprehension. The size of the effect may also be affected by the updating criterion (semantic vs. formal) of the task. The lack of predictive power of the updating measure with numerical material based on a formal criterion in our study may suggest that not only the ability to update verbal content is more important to comprehension monitoring than the updating of numerical content but the ability to update working memory based on a semantic criterion might be more important than updating based on a formal criterion. In the present study, we included the two extreme forms (verbal-semantic and numeric-formal) and therefore cannot differentiate the extent to which the two dimensions play a role in reading comprehension and comprehension monitoring. Future research should use tasks that dissociate the updating criterion (semantic vs. formal) and the content dimension (verbal vs. numerical) to shed light on the empirical validity of these considerations. More precisely, such a study should examine updating of verbal content based on a formal criterion and updating of numeric content based on a semantic criterion, for example, by assessing the four possible combinations of the two two-level dimensions simultaneously in order to compare the respective effects on reading comprehension and comprehension monitoring.

Along with the different updating criteria, the two updating measures in the present study also differed in whether the updating process within working memory required access to information stored in long-term memory. To update working memory based on a semantic criterion, participants needed to retrieve the size of each newly presented object from their long-term memory and compare the size of the new object with the objects currently stored in working memory. These processes might resemble the updating processes involved in reading comprehension more closely (compared to those of a formal updating task) because of the involvement of knowledge stored in long-term memory (Butterfuss & Kendeou, 2018). The use of updating tasks that draw on little knowledge stored in long-term memory might have led to an underestimation of the importance of updating for reading comprehension, especially of complex expository texts, in previous studies. Again, we believe that these questions deserve more attention in future research.

Furthermore, an interesting objective for future research would be to expand the research on comprehension monitoring to include a lifespan perspective because numerous studies have examined and compared different but distinct age groups. A broader cross-sectional sample would help to extend conclusions about the development of metacognitive comprehension monitoring to the lifespan.

It would also be interesting to compare the findings of this study on the inconsistency task with developments in other metacognitive measures such as calibration tasks. In calibration tasks, participants' own evaluation of their comprehension of a text is compared to their actual performance in a task measuring their comprehension, by for example asking them after having read a text how many out of a predefined number of questions about the text content they think they can answer correctly (e.g., Soto et al., 2023). Thus, in contrast to an inconsistency task, calibration measures not only consider objective performance but also include the readers' subjective perspective. Considering performance in inconsistency and calibration tasks simultaneously could provide a more holistic picture of changes in metacognitive processes in reading comprehension across the lifespan.

In addition, variations of the inconsistency task might provide more insight into differences between younger and older adults (e.g., narrative instead of expository texts). It might also be beneficial to use a greater number of consistent and inconsistent texts in future research. In the current study, the texts were randomly assigned to subjects in their consistent and inconsistent versions, resulting in a nearly balanced distribution of inconsistent and consistent version of each text across subjects.

In conclusion, this study yields findings on differences between younger and older adults in metacognitive comprehension monitoring and thus extends previous results on differences between age groups and the role of basic cognitive abilities in comprehension monitoring from childhood and adolescence into adulthood. Based on these findings, future research could also consider the criterion underlying the updating process as another relevant dimension apart from the content to be updated.

References

Alloway, T. P., & Alloway, R. G. (2013). Working memory across the lifespan: A cross-sectional approach. *Journal of Cognitive Psychology, 25*(1), 84–93. https://doi.org/ggbs8j

Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. *Journal of Reading Behavior, 11(*4), 365–374. https://doi.org/10.1080/10862967909547342

Baker, L. (1989). Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review, 1*(1), 3–38. http://dx.doi.org/10.1007/BF01326548

Bann, C. M., Bayen, U. J., McCormack, L. A., & Uhrig, J. D. (2006). Effects of reading habits, reading comprehension, and memory beliefs on older adults' knowledge about medicare. *Journal of Applied Gerontology, 25*(1), 49–64. https://doi.org/10.1177/0733464805283627

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. http://dx.doi.org/10.18637/jss.v067.i01.

Borella, E., Ghisletta, P., & de Ribaupierre, A. (2011). Age differences in text processing: The role of working memory, inhibition, and processing speed. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 66*(3), 311–320. https://doi.org/d8t3p4

Bundesministerium für Bildung und Forschung (2023, January 17). Anteil der Studienberechtigten an der altersspezifischen Bevölkerung (Studienberechtigtenquote) nach Art der Hochschulreife [Proportion of persons with higher education entrance qualification in the age-specific population (rate of persons with higher education entrance qualification) by type of higher education entrance qualification]. https://www.datenportal.bmbf.de/portal/de/Tabelle-2.5.85.html

Butterfuss, R., & Kendeou, P. (2018). The role of executive functions in reading comprehension. *Educational Psychology Review, 30*(3), 801–826. https://doi.org/gjjk4w

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31-42. http://dx.doi.org/10.1037/0022-0663.96.1.31

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences, 19*(2), 245-251. https://doi.org/ds4grq

Carretti, B., Cornoldi, C., De Beni, R., & Palladino, P. (2004). What happens to information to be suppressed in working-memory tasks? Short and long term effects. *The Quarterly Journal of Experimental Psychology Section A, 57*(6), 1059–1084. https://doi.org/10.1080/02724980343000684.

Cherry, K. E., Elliott, E. M., Golob, E. J., Brown, J. S., Kim, S., & Jazwinski, S. M. (2021). Strategic encoding and retrieval processes in verbal recall among middle-aged and older adults. *British Journal of Developmental Psychology, 39*(2), 252–268. https://doi.org/10.1111/bjdp.12349

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7*(12), 547–552. https://doi.org/c8jjx5

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 99*(2), 311–325. https://doi.org/10.1037/0022-0663.99.2.311

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422–433. https://doi.org/c659r8

De Beni, R., Borella, E., & Carretti, B. (2007). Reading comprehension in aging: The role of working memory and metacomprehension. *Aging, Neuropsychology, and Cognition, 14*(2), 189–212. https://doi.org/10.1080/13825580500229213

de Bruïne, A., Jolles, D., & van den Broek, P. (2021). Minding the load or loading the mind: The effect of manipulating working memory on coherence monitoring. *Journal of Memory and Language, 118,* 104212. https://doi.org/10.1016/j.jml.2020.104212

DeShon, R. P. (1998). A cautionary note on measurement error corrections in structural equation models. *Psychological Methods, 3*(4), 412–423. https://doi.org/10.1037/1082-989X.3.4.412

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309–331. https://doi.org/ddp3gg

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906–911. http://dx.doi.org/10.1037/0003-066X.34.10.906

Freed, E. M., Hamilton, S. T., & Long, D. L. (2017). Comprehension in proficient readers: The nature of individual variation. *Journal of Memory and Language, 97*, 135–153. https://doi.org/10.1016/j.jml.2017.07.008

Hacker, D. J. (1997). Comprehension monitoring of written discourse across early-to-middle adolescence. *Reading and Writing: An Interdisciplinary Journal, 9*(3), 207–240. http://dx.doi.org/10.1023/A:1007989901667

Hale, S., & Myerson, J. (1996). Experimental evidence for differential slowing in the lexical and nonlexical domains. *Aging, Neuropsychology, and Cognition, 3*(2), 154-165. https://doi.org/10.1080/13825589608256621

Heller, K. & Perlth, C. (2000). *Kognitiver Fähigkeitstest 4–12* [Cognitive ability test for grade 4 to12]. Hogrefe.

Helder, A., van Leijenhorst, L., & van den Broek, P. (2016). Coherence monitoring by good and poor comprehenders in elementary school: Comparing offline and online measures. *Learning and Individual Differences, 48*, 17–23. http://dx.doi.org/10.1016/j.lindif.2016.02.008

Hoover, W. A., & Gough, P. B. (1990). The Simple View of Reading. *Reading and Writing: An Interdisciplinary Journal, 2*, 127-160. http://dx.doi.org/10.1007/BF00401799

Institut für deutsche Sprache. (2009). *DeReWo. Korpusbasierte Wortlisten* [DeReWo. Corpus-based word lists]. Technical Report IDS-KL-2009-02.

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning. Implications for training and transfer. *Intelligence, 38*(6), 625–635. https://doi.org/10.1016/j.intell.2010.09.001

Kim, Y.-S. G. (2015). Developmental, component-based model of reading fluency: An investigation of predictors of word-reading fluency, text-reading fluency, and reading comprehension. *Reading Research Quarterly, 50*(4), 459–481. https://doi.org/10.1002/rrq.107

Kim, Y.-S. G. (2017). Why the simple view of reading is not simplistic: Unpacking component skills of reading using a direct and indirect effect model of reading (DIER). *Scientific Studies of Reading, 21*(4), 310–333. https://doi.org/10.1080/10888438.2017.1291643

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163–182. http://dx.doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kurby, C. A., & Zacks, J. M. (2012). Starting from scratch and building brick by brick in comprehension. *Memory & Cognition, 40*(5), 812–826. https://doi.org/fxnxgw

Mar, R. A., Li, J., Nguyen, A. T., & Ta, C. P. (2021). Memory and comprehension of narrative versus expository texts: A meta-analysis. *Psychonomic Bulletin & Review*, 1–18. https://doi.org/f359

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

Muthen, L. K., & Muthen, B. (2012). *Mplus user's guide* (7th ed.). Muthen & Muthen.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes, 26*(2-3), 131–157. https://doi.org/10.1080/01638539809545042

Noh, S. R., & Stine-Morrow, E. A. L. (2009). Age differences in tracking characters during narrative comprehension. *Memory & Cognition, 37*(6), 769–778. https://doi.org/10.3758/MC.37.6.769

O'Brien, E. J., & Albrecht, J. E. (1992). Comprehension strategies in the development of a mental model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(4), 777.

https://doi.org/10.1037/0278-7393.18.4.777

O'Brien, E. J., & Cook, A. E. (2016). Separating the activation, integration, and validation components of reading. In B. H. Ross (Ed.), *The psychology of learning and motivation; the psychology of learning and motivation* (pp. 249–276). Elsevier.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition, 29*(2), 344–354. https://doi.org/dhrqds

Payne, B. R., Gao, X., Noh, S. R., Anderson, C. J., & Stine-Morrow, E. (2012). The effects of print exposure on sentence processing and memory in older adults: Evidence for efficiency and reserve. Aging, Neuropsychology, and Cognition, *19*(1-2), 122–149. https://doi.org/10.1080/13825585.2011.628376

Pelegrina, S., Capodieci, A., Carretti, B., & Cornoldi, C. (2015). Magnitude representation and working memory updating in children with arithmetic and reading comprehension disabilities. *Journal of Learning Disabilities, 48*(6), 658–668. https://doi.org/10.1177/0022219414527480

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H., Dardick, W., & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin, 144*, 48–76. http://dx.doi.org/10.1037/bul0000124

Radvansky, G. A., & Dijkstra, K. (2007). Aging and situation model processing. *Psychonomic Bulletin & Review, 14*(6), 1027–1042. https://doi.org/10.3758/BF03193088

Reynolds, M. R., & Turek, J. J. (2012). A dynamic developmental link between verbal comprehension-knowledge (gc) and reading comprehension: Verbal comprehension-knowledge drives positive change in reading comprehension. *Journal of School Psychology, 50*(6), 841–863. https://doi.org/10.1016/j.jsp.2012.07.002

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes, 52*(5–6), 337–355. https://doi.org/10.1080/0163853X.2015.1025665

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information:

A two-step model of validation. *Educational Psychologist, 52*(3), 148–166. https://doi.org/10.1080/00461520.2017.1322968

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Schneider, W., Tibken, C., & Richter, T. (2022). The development of metacognitive knowledge from childhood to young adulthood: Major trends and educational implications. In J. Lockman (Ed.), *Advances in Child Development and Behavior* (Vol. 63). Elsevier.

Singer, M. (2006). Verification of text ideas during reading. *Journal of Memory and Language, 54*, 574–591. https://doi.org/10.1016/j.jml.2005.11.003

Singer, M. (2013). Validation in reading comprehension. *Current Directions in Psychological Science, 22*(5), 361–366. https://doi.org/10.1177/0963721413495236

Soto, C., Gutierrez de Blume, A. P., Rebolledo, V., Rodríguez, F., Palma, D., & Gutiérrez, F. (2023). Metacognitive monitoring skills of reading comprehension and writing between proficient and poor readers. *Metacognition and Learning, 18*(1), 113-134.

Statistisches Bundesamt (2022, January 24). Anteil der Personen mit Hochschulabschluss nach Altersgruppen in Deutschland in den Jahren 2009, 2014 und 2019 [Proportion of people with a university degree by age group in Germany in 2009, 2014, and 2019]. https://de.statista.com/statistik/daten/studie/1227976/umfrage/anteil-der-personen-mit-hochschulabschluss-nach-altersgruppen/

Stine-Morrow, E. A. L., & Radvansky, G. A. (2018). Discourse processing and development through the adult lifespan. In M. F. Schober, D. N. Rapp, & M. A. Britt (Eds.), *The Routledge handbook of discourse processes* (pp. 247–268). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781315687384-14

Swanson, H. L., Howard, C. B., & Sáez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities, 39*(3), 252–269. https://doi.org/10.1177/00222194060390030501

Tibken, C., Richter, T., von der Linden, N., Schmiedeler, S., & Schneider, W. (2022). The role of metacognitive competences in the development of school achievement among gifted adolescents. *Child Development, 93*(1), 117–133. https://doi.org/10.1111/cdev.13640

Tibken, C., Richter, T., Wannagat, W., Schmiedeler, S., von der Linden, N., & Schneider, W. (2022). Measuring comprehension monitoring with the inconsistency task in adolescents: Stability, associations with reading comprehension skills, and differences between grade levels. *Discourse Processes, 59*(5–6), 439–461. https://doi.org/10.1080/0163853X.2022.2073736

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology, 7*, 769. https://doi.org/10.3389/fpsyg.2016.00769

Umanath, S., & Marsh, E. J. (2012). Aging and the memorial consequences of catching contradictions with prior knowledge. *Psychology and Aging, 27*(4), 1033. http://dx.doi.org/10.1037/a0027242

van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes, 54*(5–6), 360–372. http://dx.doi.org/10.1080/0163853X.2017.1306677

van den Broek, P. W., Risden, K., & Husebye-Hartman, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 353–373). Lawrence Erlbaum Associates, Inc.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.

Wannagat, W., Steinicke, V., Tibken, C., & Nieding, G. (2022). Same topic, different genre: Elementary school children's mental representations of information embedded in narrative and expository texts. *Learning and Instruction, 80*, 1–10. https://doi.org/jhst

Wassenburg, S. I., Bos, L. T., de Koning, B. B., & van der Schoot, M. (2015). Effects of an inconsistency-detection training aimed at improving comprehension monitoring in primary school children. *Discourse Processes, 52*(5–6), 463–488. https://doi.org/gtck

Wechsler, D. (2014). Wechsler intelligence scale for children (5th ed., WISC–V). Pearson.

Zabrucky, K., & Moore, D. (1991). Effects of skill on standards used by younger and older adults to evaluate understanding. *Reading Psychology, 12*(2), 147–158. https://doi.org/fvnbd5

Zabrucky, K., & Moore, D. (1994). Contributions of working memory and evaluation and regulation of understanding to adults' recall of texts. *Journal of Gerontology, 49*(5), 201–212. https://doi.org/10.1093/geronj/49.5.P201

Zabrucky, K. M., & Moore, D. (1999). Influence of text genre on adults' monitoring of understanding and recall. Educational Gerontology, 25(8), 691–710. https://doi.org/bd4jdm

Zabrucky, K. M., Moore, D., & Agler, L. M. L. (2012). Metacomprehension across the Lifespan: Influence of education and instructional support. *Procedia-Social and Behavioral Sciences, 46*, 601–604. https://doi.org/10.1016/j.sbspro.2012.05.170

Zabrucky, K., Moore, D., & Schultz, N. R. (1987). Evaluation of comprehension in young and old adults. *Developmental Psychology, 23*(1), 39–43. https://doi.org/10.1037/0012-1649.23.1.39

Zabrucky, K., & Ratner, H. H. (1986). Children's comprehension monitoring and recall of inconsistent stories. *Child Development, 57*(6), 1401-1418. https://doi.org/10.2307/1130419

Zargar, E., Adams, A. M., & Connor, C. M. (2019). The relations between children's comprehension monitoring and their reading comprehension and vocabulary knowledge: An eye-movement study. *Reading and Writing: An Interdisciplinary Journal 33*(3). 511–545. https://doi.org/gtcm

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162 –185. https://doi.org/cntx9t

Table 1

Means and Standard Deviations of Study Variables for the Total Sample, Younger Adults, and Older Adults

	Total sample (N = 209)	Younger adults (n = 101)	Older adults (n = 108)
Number of detected inconsistencies	3.12 (1.56)	3.41 (1.48)	2.86 (1.60)
Working memory updating based on a semantic criterion	15.31 (1.87)	15.72 (1.37)	14.93 (2.18)
Working memory updating based on a formal criterion	6.28 (1.89)	6.78 (1.68)	5.81 (1.97)
Short-term memory capacity	7.30 (2.18)	7.96 (2.35)	6.68 (1.81)
Verbal intelligence	21.94 (7.11)	24.76 (4.49)	19.30 (8.05)

Note. Number of detected inconsistencies (self-constructed inconsistency task), working memory updating based on a semantic criterion (self-constructed task), working memory updating based on a formal criterion (self-constructed *n*-back task), short-term memory (digit span forward of the WISC-V), indicator of verbal intelligence (word analogies of the KFT 4–12).

Appendix 1: Text Example of the Inconsistency Task

English Translation

Craft instruction for a paper star

Take a square sheet of paper and fold it into a triangle. For this, fold one half diagonally on the other. Then fold again one half on the other. This creates a triangle again, but a smaller (= consistent)/larger (= inconsistent) one. Now cut different slits, holes and curves into the paper as you like. All layers of the paper have to be exactly on top of each other. Pay attention that the serrations and incisions do not run completely over the closed side of the triangle, otherwise the star will fall apart in the end. Open the paper completely and the star is finished.

Original Text (in German)

Bastelanleitung für einen Papierstern

Nimm ein quadratisches Blatt Papier in die Hand und falte es zum Dreieck. Dafür wird eine Hälfte diagonal auf die andere gefaltet. Falte dann noch einmal die eine Hälfte auf die andere. So entsteht wiederum ein Dreieck, aber ein kleineres (= konsistent)/größeres (= inkonsistent). Schneide jetzt nach Belieben verschiedene Schlitze, Löcher und Rundungen in das Papier. Dabei liegen alle Schichten des Papiers immer genau aufeinander. Pass auf, dass die Zacken und Einschnitte nicht komplett über die geschlossene Seite des Dreieckes verlaufen, sonst fällt der Stern am Ende auseinander. Klappe das Papier wieder vollständig auf und der Stern ist fertig.