



# The Role of Individual Differences in Sourcing: a Systematic Review

Øistein Anmarkrud<sup>1</sup> · Ivar Bråten<sup>1,4</sup> · Elena Florit<sup>2</sup> · Lucia Mason<sup>3</sup>

Accepted: 14 September 2021 / Published online: 29 October 2021  
© The Author(s) 2021

## Abstract

This article reviews how individual differences have been conceptualized and researched within the area of multiple document literacy, in particular the extent to which proposed relationships between individual differences and the multiple document literacy process of sourcing have been supported by the empirical research. The findings showed that although the majority of the individual differences included in theoretical models of multiple document literacy have been researched, the empirical backing of proposed relationships is rather ambiguous. Still, in-depth analyses of the most researched individual differences in relation to sourcing revealed some interesting and interpretable patterns. Further, the review suggested that relationships between individual differences and sourcing may vary not only with the way sourcing is measured but also with the domain or topic addressed in the reading materials. We discuss the current status of research on individual differences in the context of multiple document literacy with a focus on sourcing and suggest potential avenues for further clarifications.

**Keywords** Multiple document literacy · Sourcing · Individual differences · Systematic review

## Introduction

Within twenty-first century literacy research, sourcing is regarded as a hallmark of advanced literacy skills (Bråten, Stadtler, & Salmerón, 2018; Britt et al., 2013; Goldman & Brand-Gruwel, 2018; Magliano et al., 2018). In this article, following

---

✉ Ivar Bråten  
ivar.braten@ped.uio.no

<sup>1</sup> University of Oslo, Oslo, Norway

<sup>2</sup> University of Verona, Verona, Italy

<sup>3</sup> University of Padova, Padua, Italy

<sup>4</sup> Department of Education, University of Oslo, Blindern, P.O. Box 1092, 0317 Oslo, Norway

Bråten, Stadtler, and Salmerón (2018), we define sourcing as attending to, representing, evaluating, and using information about the sources of document content. Relevant source features may include the author, the document genre or type, the venue, and the place and date of document creation or update (Britt & Aglinskias, 2002). Presumably, considering such source features in the reading process helps individuals become critical readers and learners rather than passive consumers of information (Bråten & Braasch, 2017). Important insights have been gained regarding how sourcing is related to learning and comprehension and how it can be promoted through instructional interventions at different educational levels (for reviews, see Brante & Strømsø, 2018; Bråten, Stadtler, & Salmerón, 2018). Additionally, contextual influences on students' sourcing have been highlighted in recent years (Britt et al., 2018). In the current review, however, we focus on the role of individual difference factors in sourcing when people read to learn or comprehend document content.

Within literacy research, the process of sourcing has mainly been discussed by theoreticians interested in readers' ability to handle multiple documents, that is, multiple document literacy (Strømsø & Bråten, 2013). However, existing conceptualizations of multiple document literacy vary substantially with respect to the individual difference factors that they consider, and these factors are more or less directly related to the process of sourcing. Further, the inclusion of particular individual differences may sometimes seem unsubstantiated and even somewhat speculative, with an empirical backing that is far from clarified. On this backdrop, it seems highly pertinent to conduct a systematic review of the empirical grounding (or the lack of it) that prevailing assumptions about relationships between individual differences and sourcing have. Hopefully, such a review will not only interest researchers in the field due to its potential contribution to theoretical clarification and refinement, but also provide educators and policy makers with important insights into what it takes to be a critical reader and learner and how the teaching of critical literacy may have to be adapted to individual differences within the student population.

Although sourcing is but one component of multiple document literacy, with content integration considered another important aspect (Braasch et al., 2018), it has long been regarded as crucial by researchers in this area (Rouet, 2006). For example, sourcing may allow readers to prioritize information from competent, unbiased, vetted, and updated sources, which, in turn, may help them build a more appropriate, higher-quality mental representation of the issue discussed across documents (Delgado et al., 2020). In particular, sourcing may facilitate the integration of content across documents because it promotes understanding of the reasons for different views and perspectives on the same issue (e.g., because authors differ with respect to their competencies or motives; Bråten et al., 2019). However, research on sourcing has not only been reviewed (Brante & Strømsø, 2018) and compiled (Scharrer & Salmerón, 2016) by researchers in reading more recently; it has also been increasingly conducted and discussed in areas such as science education (Duncan et al., 2018; Sinatra & Lombardi, 2020) and civic reasoning and discourse (Chinn et al., 2021; McGrew, 2021). Still, no prior review has focused on the potential influences of diverse individual differences on this specific, yet indispensable aspect of multiple document literacy.

## Conceptualizing Individual Differences within Multiple Document Literacy with a Focus on Sourcing

Thirty years ago, Wineburg (1991) described how expert historians differed from high school students when trying to make sense of the same set of documents on a particular historical event. One of the most salient differences noted by Wineburg (1991) was that the historians engaged heavily in a sense-making activity that he termed “sourcing,” involving that they not only studied the contents of the documents but also considered information about the documents themselves. Thus, before reading the content of a document, the historians were observed paying attention to document information such as the author, document type, and place and date of its creation to judge its evidentiary value. When further processing the document, such information was considered in the interpretation of a document’s content. The high school students, in contrast, were observed to generally disregard document information and simply rely on the textbook’s description of the historical event.

Wineburg’s (1991) landmark study can be considered ground zero within research on sourcing in the reading process (Goldman & Scardamalia, 2013). This study did not lead to a wave of studies concerning sourcing within literacy research in the 1990s, however, and hardly did so in the first decade of the new century.<sup>1</sup> Arguably, the next important step in this area of research was Perfetti and colleagues’ attempt to conceptualize how readers mentally represent multiple documents that deal with the same topic or issue (Perfetti et al., 1999). Thus, without any empirical basis except for Wineburg’s (1991) early study and their own preliminary empirical work (Britt et al., 1999; Perfetti et al., 1995; Rouet et al., 1996, 1997), these authors proposed a documents model framework in which attention to source information played a crucial role.

Perfetti et al. (1999) built on the influential construction-integration model of Kintsch (1988), which explains the comprehension of a single text, and proposed that readers mentally represent related semantic content from multiple documents as a “situations model,” that is, as an integrated understanding of the situation described across documents.<sup>2</sup> However, in addition to this content-based representation, they suggested that readers of multiple documents represent information about the source of each document and links between such information and important document content, as well as links between the sources of different documents. Taken together, the source-content links and the source-source links described above constitute the intertext model within the documents model framework, with this model presumably required to understand the contribution of each source to an integrated representation of the situation, to qualify or judge content information in light of its

<sup>1</sup> In fact, 93% of the studies included in the current review were published from 2010 to 2020.

<sup>2</sup> The term “situations model” signals the relationship between the documents model framework and Kintsch’s (1988) theory, as well as the original focus on reconciling different accounts of historical events and situations (Wineburg, 1991). Because the framework was extended to the comprehension of multiple documents in other domains, including the natural sciences, Britt and Rouet (2012) later suggested that the more general term “integrated mental model” is more appropriate for this aspect of multiple document literacy.

source, and to understand relationships between sources. No complete documents model can therefore be achieved without the situations model being supplemented and integrated with an intertext model (Perfetti et al., 1999).

In accordance with Wineburg's (1991) findings and their own preliminary work (Perfetti et al., 1995; Rouet et al., 1997), Perfetti et al. (1999) suggested that individual differences in disciplinary expertise, including "specific domain knowledge and trained experiences with texts in that domain" (p. 117), would influence readers' representation of source-content and source-source links (i.e., intertext model construction). Likewise, later conceptualizations building on and expanding the documents model framework have highlighted the potential importance of readers' domain, document, or disciplinary expertise for adaptive sourcing in multiple document literacy contexts (Britt et al., 2018; Rouet & Britt, 2011). In addition, these conceptualizations have featured a host of other individual differences presumably relevant to sourcing in the reading process.

Thus, in the multiple-document task-based relevance assessment and content extraction (MD-TRACE) model, which is a process model of multiple document use, sourcing is considered to be involved in the selection and processing of documents to build a complete documents model including source-content and source-source links (i.e., an intertext model; Rouet & Britt, 2011). In addition to knowledge about relevant source features (i.e., source knowledge) and document or disciplinary expertise, which are directly linked to sourcing in the model, Rouet and Britt (2011) highlighted the importance of "permanent internal resources" related to reading skills and strategies, prior knowledge about the content of the documents, and working memory/executive control. These reading skills and cognitive factors were not directly related to the subprocess of sourcing, however.

More recently, Britt et al. (2018; see also, Rouet et al., 2017) proposed a model of reading as problem solving (RESOLV). In particular, this model further elaborates and differentiates the first processing steps of the MD-TRACE model in suggesting that individuals construct mental representations of the reading context as well as the reading task, with these representations, in turn, guiding their processing of the documents (including documents model construction). In constructing these representations and further processing the documents, readers are assumed to draw on a broad array of internal personal resources, including not only reading skills and strategies (including metacognitive, self-regulatory skills), prior knowledge, and working memory/executive control, but also resources related to motivation and personality. Thus, achievement goal orientations (mastery, performance), individual interest, task values (attainment, utility, intrinsic), and self-concept of ability (self-efficacy, expectancy for success, perceived competence) were described as motivational factors most relevant to reading as problem solving, along with personality factors such as conscientiousness, openness to experience, need for cognition, and growth versus fixed mindset.<sup>3</sup> Finally, the RESOLV model acknowledged that readers' beliefs, both regarding the topic discussed in the documents (i.e., topic beliefs) and regarding the nature of knowledge and the process of knowing (i.e., epistemic

<sup>3</sup> Growth versus fixed mindset was categorized as a motivational factor by Britt et al. (2018).

beliefs), are likely to influence reading as problem solving. Although the RESOLV model is distinguished by its broad discussion of the potential role of individual differences in reading, cognition, motivation, personality, and beliefs, it remains unclear whether or to what extent each of these factors is associated with reading as problem solving, in particular, not to speak of the subprocess of sourcing involved in documents model construction. Arguably, most of these individual difference factors seem as relevant to learning and comprehension in general as they are to reading as problem solving and the specific process of sourcing.

List and Alexander (2019) introduced an integrated framework of multiple texts (IF-MT) based on several previous conceptualizations within multiple document literacy (including the documents model framework, the MD-TRACE, and the RESOLV), most notably on their own cognitive affective engagement model of multiple source use (CAEM; List & Alexander, 2017). In this framework, List and Alexander (2019) described three stages of multiple document use: a preparation stage in which readers establish a stance toward the reading task, an execution stage in which they strategically process the documents, and a production stage in which they construct a task product. In terms of individual differences, the preparation stage is especially important. In this stage, readers are assumed to adopt a particular stance or orientation to task completion that involves a combination of their individual interests and attitudes concerning the domain or topic with their pre-established habits in dealing with multiple documents in terms of content integration and source evaluation (i.e., their multiple document proficiency/expertise). Specifically, when individual interest is low, attitudes weak, and multiple document proficiency low, readers are assumed to adopt a disengaged stance. Conversely, when individual interest is high, attitudes strong, and multiple document proficiency high, readers are assumed to adopt a critical analytic stance. Regarding the two stance profiles that fall in between, readers with high individual interest, strong attitudes, and low multiple document proficiency are assumed to adopt an affectively engaged stance, and readers with low individual interest, weak attitudes, and high multiple document proficiency are assumed to adopt an evaluative stance. Most importantly, these four stance profiles are also assumed to differ in terms of their sourcing behavior.

Thus, readers adopting evaluative and critical analytic stances to task completion are assumed to frequently engage in sourcing when working with multiple documents. However, only readers adopting a critical analytic stance are likely to judge the credibility of content information in light of the sources and also use source information in trying to reconcile textual conflicts, as well as to remember both content and source information in an organized and integrated way. In comparison, readers adopting an evaluative stance are likely to routinely access source information in judging content credibility and to accurately recall source information and source-content links, but these readers are not assumed to use source information in the service of meaning making and integration because of their low engagement (i.e., low interest and weak attitudes) in the topic or the task. Finally, both disengaged and affectively engaged readers are assumed to engage infrequently in sourcing (List & Alexander, 2017).

In addition to highlighting the importance of individual differences in readers' interest, attitudes, and multiple document proficiency/expertise for adaptive

sourcing, List and Alexander (2019) proposed that prior knowledge is likely to influence sourcing by supporting the adoption of a critical analytic stance. Similarly, epistemic beliefs about justification for knowing, in particular about appropriate sources of knowledge and methods for justifying knowledge claims, were assumed to influence sourcing by supporting the adoption of evaluative and critical analytic stances to task completion (List & Alexander, 2019).

Although not highlighted as an individual factor in their framework, List and Alexander (2019) discussed several forms of strategic competence of potential importance for readers' multiple source use, including behavioral, cognitive (intra-textual, intertextual), and metacognitive and regulatory strategies. Such strategies are described in great detail in Cho and colleagues' (Afflerbach & Cho, 2009; Cho & Afflerbach, 2017; Cho et al., 2018) taxonomies of constructively responsive reading comprehension strategies (CRRCS) in reading multiple and digital texts (including hypertext).

The CAEM (List & Alexander, 2017) and the IF-MT (List & Alexander, 2019) are distinguished from other models of multiple document literacy by specifying how individual difference factors—in particular, profiles of individual differences based on individual interest, attitudes, and multiple document proficiency—are related to the subprocess of sourcing. Thus far, these assumptions must be regarded as hypotheses that are fairly loosely grounded in empirical evidence, however.

Finally, more specialized theoretical models that focus on situations in which different sources disagree also have relevance to the issue of individual differences in sourcing. One such model is the discrepancy-induced source comprehension (D-ISC) model (Braasch & Bråten, 2017; Braasch et al., 2012; Bråten & Braasch, 2018). Essentially, this model states that, when readers encounter conflicting information about a particular topic or issue and thereby experience a break in situational coherence, they may strategically pay attention to source information (i.e., who said what) in an effort to understand the conflict and restore coherent understanding of document content. Recently, the D-ISC model has been extended to contexts in which the content of a document contradicts readers' pre-existing beliefs about a particular topic or issue (Braasch & Bråten, 2017; Bråten & Braasch, 2018). In such contexts, document content will be considered less plausible and individuals may therefore seek support from source information to try to make sense of the content (Bråten & Braasch, 2018). This explanation is consistent with the two-step model of validation by Richter and Maier (2017, 2018). According to these authors, when readers detect a conflict between content information and their prior topic beliefs, they may take strategic action and engage in elaborative processing of conflicting information, given that they are motivated and cognitively capable of doing so. Presumably, such elaborative processing of conflicting information may also include attention to the source of that information. To what extent perceived conflicts between content information and readers' prior beliefs about the topic influence the process of sourcing, in particular, is an issue for further research, however.

While the D-ISC model seems to be consistent with an emphasis on individual differences in readers' beliefs about the topic of the documents, the content-source integration (CSI) model described by Stadtler and Bromme (2014), which also focuses on the reading of conflicting information, brings to the forefront individual

differences in prior knowledge. According to the CSI model, readers may try to resolve conflicts encountered in documents by asking themselves “what is true?,” in which case they direct their efforts toward validating textual claims in light of their prior knowledge or toward evaluating the quality of the argumentative reasoning underlying the claims. However, when readers do not possess sufficient prior knowledge or argumentative reasoning skills to evaluate textual claims directly, they may resort to the strategy of asking themselves “who to believe?,” in which case they pay attention to and evaluate the sources (e.g., the authors) presenting the claims.

Interestingly, then, the CSI model of Stadler and Bromme (2014) seems to suggest that less prior knowledge may be associated with more sourcing, whereas other conceptualizations, such as the frameworks proposed by List and Alexander (2017, 2019), seem to suggest that more prior knowledge will increase sourcing, for example through its contribution to the adoption of a critical analytic stance to task completion. Needless to say, further clarification of whether prior knowledge actually is positively or negatively related to readers’ sourcing (if at all) seems to be an important aim for a systematic review of the role individual differences play in sourcing.

## The Present Review

In summary, a range of individual difference factors have been highlighted within conceptualizations of multiple document representation and use during the last decade. These individual differences can be categorized into reading skills and strategies, cognitive factors, motivation and engagement, beliefs, personality, and expertise. As indicated in Table 1, different conceptualizations have emphasized the importance of these individual difference factors for multiple document representation and use, including for the subprocess of sourcing, to varying degrees. It is currently not clear, however, to what extent these individual difference factors actually have been included in research on sourcing, nor is it clear to what extent proposed relationships between these individual difference factors and sourcing have received empirical backing.

Therefore, we set out to review the empirical work investigating associations between individual differences and sourcing in the reading process. The review was restricted to studies published between 1991 and 2020, that is, the period from the publication of Wineburg’s (1991) groundbreaking study to the present. As noted previously, we adopted a broad definition of sourcing as attending to, representing, evaluating, and using information about the sources of document content, such as the author, publication, or type of document (Bråten, Stadler, & Salmerón, 2018). By building on this definition, we expected to identify studies that operationalized and measured sourcing in different ways, varying from noting and remembering source information to evaluating the credibility of sources and including source information in task products (e.g., in essays). Hopefully, this approach would also give us the opportunity to explore whether any associations between individual difference variables and sourcing might vary with the way sourcing was measured. Finally, given that we expected to identify studies

**Table 1** Individual differences highlighted in conceptualizations of multiple document representation and use

Individual differences	Conceptualizations
<i>Reading skills and strategies</i>	MD-TRACE, RESOLV, IF-MT, CRRCS
Word-level reading skills	
Comprehension	
Comprehension strategies	
<i>Cognitive constructs</i>	MD-TRACE, RESOLV, IF-MT, D-ISC, CSI
Prior knowledge	
Working memory/executive function	
Argumentative reasoning	
<i>Motivation and engagement</i>	RESOLV, IF-MT
Achievement goals	
Task values	
Interest	
Attitudes	
Self-concept of ability	
<i>Personality</i>	RESOLV
Conscientiousness	
Openness to experience	
Need for cognition	
Growth vs. fixed mindset	
<i>Beliefs</i>	RESOLV, IF-MT, D-ISC
Topic beliefs	
Epistemic beliefs	
<i>Expertise</i>	DMF, MD-TRACE, RESOLV, IF-MT
Domain/document/disciplinary expertise	
Multiple document proficiency	

*Note.* CSI content-source integration model (Stadtler & Bromme, 2014); CRRCS constructively responsive reading comprehension strategies (Cho & Afflerbach, 2017); D-ISC discrepancy-induced source comprehension model (Braasch & Bråten, 2017); DMF documents model framework (Perfetti et al., 1999); IF-MT integrated framework of multiple texts (List & Alexander, 2019); MD-TRACE multiple documents task-based relevance and content extraction model (Rouet & Britt, 2011); RESOLV reading as problem solving model (Britt et al., 2018)

that varied substantially with respect to domain or topic, we wanted to explore whether the role of individual differences in sourcing might vary across the domains or topics addressed in the reading materials.

Specifically, our systematic review was guided by the following four questions:

1. To what extent have the individual difference factors highlighted within conceptualizations of multiple document representation and use been included in the empirical research on sourcing?
2. To what extent have proposed relationships between individual differences and sourcing been supported by the empirical research?



3. Are there any indications that relationships between individual differences and sourcing may vary with the way sourcing is measured?
4. Are there any indications that relationships between individual differences and sourcing may vary with the domain or topic addressed in the reading materials?

## Method

### Search Strategy

We worked with a research librarian in the systematic review team of the university library at our institution to develop the search protocol and conduct the searches. The main search was conducted in January and February, 2019, with additional searches conducted in October 2019 and the beginning of October 2020. The search included three databases: ERIC (Ovid), PsychINFO (Ovid), and ISI Web of Science. Through several meetings with the systematic review team, an extensive search algorithm was developed and tested. The search algorithm was built from 91 search terms and their combinations. In addition to traditional searches in keywords indexed by authors and journals, the search algorithm made it possible to identify records based on the proximity of particular terms in the articles. For example, if the terms “information” or “text\*” occurred in a proximity of three or fewer words from terms such as “source” or “sources” in an article, that article was included for abstract review even though none of these terms were indexed as keywords. The complete search algorithm that was used is included in the online supplementary material. We also performed a manual search of highly relevant journals for studies examining sourcing and individual differences. In addition, the third author examined the publication lists of researchers considered to be central contributors within this line of research. The database search, the manual journal searches, and the publication list examination were all restricted to articles published between January 1, 1991, and October 1, 2020.

### Study Selection and Data Extraction

To be eligible for the review, the studies had to fulfill the following inclusion criteria:

1. The studies should examine the comprehension or evaluation of written resources, either text alone or text in combination with other representations.
2. The studies should report empirical research (e.g., no reviews or theoretical papers were included).
3. The studies should include actual reading. Thus, studies such as survey studies in which participants were asked to rank the trustworthiness of different types of media outlets in general without reading texts from these outlets were not included.

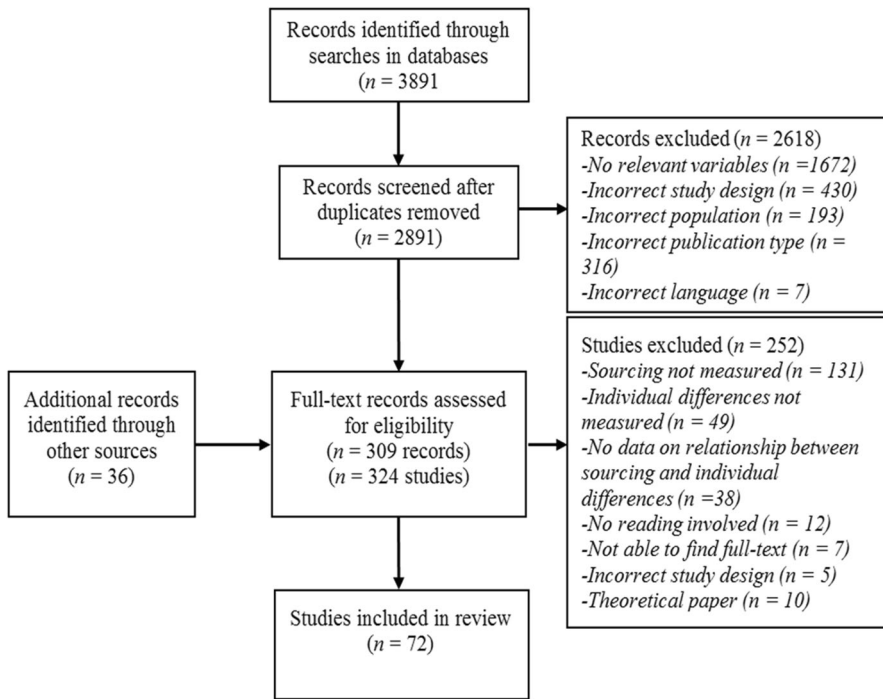
4. The studies should include at least one individual difference variable and one measure of sourcing.
5. The studies should examine typically developed populations (e.g., no participants with learning disabilities).
6. The studies should have a quantitative design and report results through numerical data.
7. The studies should be published in English.

Next, we developed a coding scheme to extract the following information from the selected studies: (a) country and participants (i.e., sample size, gender distribution, age, and grade level); (b) study design (i.e., experimental, quasi-experimental, or correlational); (c) instructions and learning material (e.g., text type, length, and topic; time limit for reading; reading medium; and availability of material during the response); (d) individual difference measures; (e) sourcing measures; (f) relationship(s) between individual difference variable(s) and sourcing variable(s); (g) type of statistics; and (h) field of research.

## Results

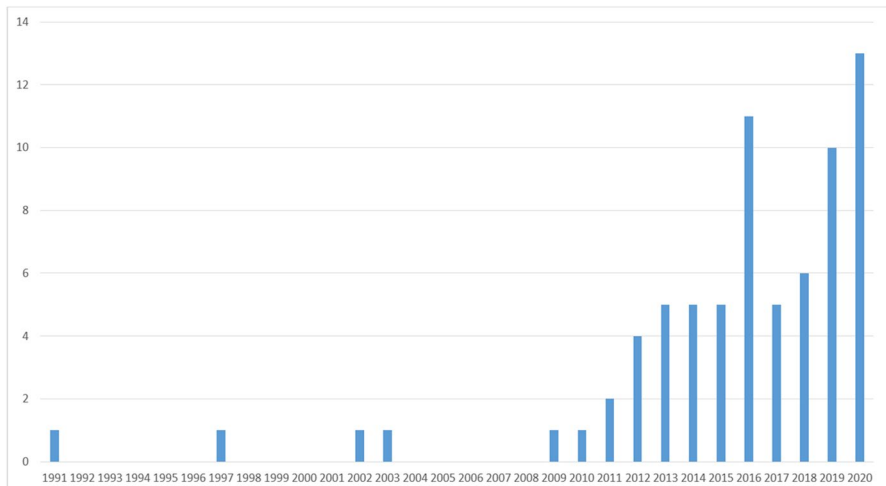
Using the search protocol described in the method section and displayed in the supplemental material, the initial searches in the selected databases yielded 3891 results. After removing duplicates, 2891 records were screened with the abstract screening software Rayyan (Ouzzani et al., 2016). The first, second, and fourth authors collaboratively screened the abstracts of 250 (8.65%) records to establish a common understanding of the inclusion and exclusion criteria. The first and third authors screened the remaining 2641 abstracts. Of the 2891 records that were screened, 2618 were removed from the review for the following reasons: no relevant variables measured (1672 records), incorrect study design (430 records), incorrect population (193 records), incorrect publication type (316 records), or incorrect language (7 records). The remaining 273 records, together with 36 records identified through a hand search in relevant journals and an examination of publication lists, were eligible for a full-text examination. Thus, 309 records, comprising 324 studies, were examined in full text by the first and third authors. Of these 324 studies, 252 were excluded for the following reasons: sourcing was not measured ( $n = 131$ ), individual differences were not measured ( $n = 49$ ), no data was provided on the relationship between sourcing and individual differences ( $n = 38$ ), no reading was involved in the study ( $n = 12$ ), the full text could not be found ( $n = 7$ , all doctoral dissertations), the study design was incorrect ( $n = 5$ ), or the paper was theoretical ( $n = 10$ ). The fourth author randomly drew 54 of these 252 excluded studies (21.4%) and found 100% agreement on whether the study should be included and 96.3% agreement on the reason for exclusion. Seventy-two studies were thus coded by the first and third authors using the detailed coding scheme previously described. Figure 1 displays an overview of the entire coding process.

These 72 studies, which are described in terms of participants, text materials, measures, and main results in Table S1 in the online supplementary material,



**Fig. 1** Overview of the coding process

included 8529 participants with substantial gender skewness. Five studies (Britt & Aglinskas, 2002, study 1; Lucassen et al., 2013; Potocki et al., 2020; Rouet et al., 1997; Wiley et al., 2020), representing 582 participants, did not report gender distribution. Among the remaining 7947 participants, 62.74% ( $n = 4986$ ) were female, and 37.26% ( $n = 2961$ ) were male. Participants represented all educational levels, with seven studies including 797 participants examining elementary school students, five studies including 313 participants examining middle school students, 12 studies including 1325 participants examining high school students, 17 studies including 2538 participants examining undergraduates, two studies including 98 participants examining master/graduate students, 20 studies including 2573 participants examining mixed or unspecified levels of university students, and nine studies including 885 participants examining samples across educational levels or samples outside the educational system. Thus, 5209 (61.07%) participants across the 72 studies were university/college students. The majority of studies ( $n = 52$ ) were conducted in Europe in the following countries: Germany ( $n = 17$ ), Norway ( $n = 15$ ), Italy ( $n = 7$ ), Spain ( $n = 5$ ), France ( $n = 5$ ), and the Netherlands ( $n = 3$ ), with one study conducted in both Germany and Spain. The remaining 20 studies came from the USA ( $n = 14$ ), Israel ( $n = 5$ ), and Japan ( $n = 1$ ). This distribution reflects the country where the data were collected and that several of the studies included in this review have author teams representing several countries. As displayed in Fig. 2, the publication trend indicates an increased number of studies in recent years.



**Fig. 2** Reviewed studies by publication year

With respect to reading medium, participants read digital documents in 46 and printed documents in 22 of the studies. In one study, participants read both digital and printed documents, and in three studies, reading medium was not specified. In the vast majority of the studies (95.83%,  $n=69$ ), participants read a predefined set or library of documents, with the number of documents varying from 1 to 17 ( $M=6.15$ ,  $SD=3.31$ ). In 17 of these studies, participants had to read all the documents, while it was up to the participants to decide which documents they wanted to read in nine studies. In the remaining 43 studies, there was not enough information in the manuscript to determine whether participants had to read all the documents. In the three studies that did not present participants with a predefined document set, they read on the open Internet.

### Inclusion of Individual Differences

Our first research question concerned the extent to which the individual difference factors highlighted within conceptualizations of multiple document representation and use have been included in the empirical research on sourcing. As shown in Table 1, six categories of individual differences (i.e., reading skills and strategies, cognitive factors, motivation and engagement, personality, beliefs, and expertise), covering 19 individual difference constructs, have been proposed as important for multiple document representation and use in contemporary conceptualizations. As displayed in Table 2, 14 of these individual difference constructs were examined in the 72 studies included in this review. Across the 72 studies, 27 different individual difference constructs were examined in relation to sourcing.

Of note is that in Table 2, we implemented a more general terminology than the original terms used to describe the measured individual differences in the reviewed studies (see Table S1 in the supplemental material). For example, some of the

**Table 2** Individual differences measured in the reviewed studies

	Individual differences	Number of studies
Measured individual differences included in conceptualizations of multiple document representation and use	<i>Reading skills and strategies</i>	29
	Word-level reading skills	6
	Comprehension	16
	Comprehension strategies	7
	<i>Cognitive constructs</i>	41
	Prior knowledge	35
	Working memory/executive functions	5
	Argumentative reasoning	1
	<i>Motivation and engagement</i>	20
	Achievement goals	-
	Task values	2
	Interest	12
	Attitudes	6
	Self-concept of ability	-
	<i>Personality</i>	3
	Conscientiousness	-
	Openness to experience	-
	Need for cognition	2
	Growth vs. fixed mindset	1
	<i>Beliefs</i>	16
	Topic beliefs	6
	Epistemic beliefs	10
	<i>Expertise</i>	8
Domain/document/disciplinary expertise	8	
Multiple document expertise	-	
Measured individual differences not included in conceptualizations of multiple document representation and use	<i>Cognitive constructs</i>	2
	Vocabulary	1
	Verbal comprehension	1
	<i>Motivation and engagement</i>	2
	Personal relevance	1
	Persistence	1
	<i>Topic familiarity</i>	1
	<i>Demographics</i>	6
	Gender	4
	Age	1
	Parental educational level	1
	<i>Educational level</i>	7
	<i>Emotional/affective constructs</i>	4
	<i>Academic achievement</i>	2
	<i>Internet experience</i>	4
<i>Historical thinking skills</i>	3	

studies used the term “prior knowledge” (e.g., Braasch et al., 2014; List et al., 2017), whereas other studies referred to “prior domain knowledge” (e.g., Kammerer et al., 2021) or “topic knowledge” (e.g., Delgado et al., 2020; Peterson & Alexander, 2020), with “domain” and “topic” suggesting that prior knowledge was measured at different levels of specificity (McCarthy & McNamara, 2021). In Table 2, all these knowledge descriptors are referred to by the more general term “prior knowledge.” Other examples are that “claim agreement” used by Bromme et al. (2015) is considered a “topic belief” in Table 2 or that “word-level reading skills” in Table 2 comprises variables such as “word recognition” (Braasch et al., 2014), “word reading fluency” (Florit et al., 2019), and “reading speed” (Kammerer, Meier, & Stahl, 2016). A final example is that, due to the wording of the items in the measures of “topic involvement” (e.g., “value” and “importance”) used by Kang et al. (2011) and West-erwick (2013, study 1), we coded “topic involvement” in the more general category “task values” in Table 2.

Among the individual differences included in conceptualizations of multiple document representation and use, the most frequently examined category in relation to sourcing was cognitive factors (41 studies), followed by reading skills and strategies (29 studies), motivation and engagement (20 studies), beliefs (16 studies), expertise (8 studies), and personality (3 studies). However, the number of studies varied substantially for the individual difference constructs within these categories. For example, of the 41 studies examining cognitive factors, 35 concerned prior knowledge. Further, of the 29 studies examining reading skills and strategies, 16 concerned reading comprehension. Of note is that all the strategies subsumed under reading skills and strategies referred to strategic processes involved in the comprehension of text. Albeit containing fewer studies, the category motivation and engagement was dominated by studies examining interest ( $n=12$ ) and attitudes ( $n=6$ ). Additionally, two (i.e., achievement goals and self-concept of ability) of the five motivation and engagement constructs highlighted in conceptualizations of multiple document representation and use had not been empirically examined in any of the studies included in this review. Likewise, only two (need for cognition and growth vs. fixed mindset) of the four personality constructs had been examined empirically. Interestingly, personality factors had been examined empirically in relation to sourcing only to a very limited degree ( $n=3$  studies).

Another related issue is whether there are individual differences not highlighted in the conceptualizations of multiple document representation and use that have been examined in relation to sourcing. Table 2 also shows the 13 individual difference constructs we identified that were not included in any of the conceptualizations. Two of these constructs, vocabulary (Macedo-Rouet et al., 2020) and verbal comprehension (Ulyshen et al., 2015), representing the category of cognitive factors, had been examined in one study each. Two additional motivational and engagement variables were identified in the studies we reviewed. Kobayashi (2014) examined “personal relevance,” a composite variable of motivation that concerned importance, experience, and interest of the topic, whereas List, Stephens, and Alexander (2019) included a variable labeled “persistence” that referred to the reading time and the number of texts read. Reading time and texts read have been used as proxies for behavioral engagement

in previous research (e.g., Bråten, Anmarkrud, et al., 2014; Bråten, Brante, & Strømsø, 2018). McCrudden et al. (2016) examined the variable “topic familiarity” based on a combination of measures targeting prior knowledge and interest. Therefore, we did not include this variable in either the cognitive or the motivation and engagement category. Two other studies included in this review referred to measures of “topic familiarity.” Lucassen et al. (2013) based topic familiarity on questions about participants’ interests and disinterests, whereas Van Der Heide and Lim (2016) referred to participants’ experience with an Internet platform in relation to familiarity. Accordingly, in Table 2, these two variables are coded as “interest” and “Internet experience.”

We also identified six studies examining sourcing in relation to a category that we labeled demographic variables, consisting of gender (4 studies), age (1 study), and parental educational level (1 study). Seven studies (e.g., Macedo-Rouet et al., 2013, experiments 1 and 2; Winter & Krämer, 2012) examined participants’ educational level in relation to sourcing. Arguably, educational level can be related to expertise. When, for example, von der Muhlen et al. (2016) compared the sourcing of undergraduates and scientists, the comparison concerned novices and experts and was coded as such, but the comparison was also related to educational level (i.e., undergraduate vs. PhD). However, when Macedo-Rouet et al. (2013, experiment 1) compared the sourcing behavior of fourth and fifth graders and found that fifth graders were better able to identify the most knowledgeable source in texts, the difference was not a matter of expertise since none of the groups could be considered experts on the topics in question (i.e., global warming, nutrition, or public transportation). Thus, we used “expertise” only in the studies in which one group of participants read texts within their area of specialization, such as in Rouet et al.’s (1997) comparison of graduate students in history (experts) and graduate students in psychology (novices) reading history documents. It is important to note, however, that the construct of expertise is multidimensional and can be assumed to involve a configuration of knowledge, strategies, interest, and beliefs in relation to a particular domain or discipline, as conceptualized within the model of domain learning (Alexander, 1997; Alexander et al., 2012). Specifically, within this model, experts in a domain or discipline are characterized by an interplay of a well-integrated body of knowledge, a well-established and efficient repertoire of deeper processing strategies, high individual interest, and adaptive beliefs about the nature of knowledge and the process of knowing.

Finally, we identified four studies examining emotional/affective constructs (Mason et al., 2017; Mason et al., 2020; Mason, Scrimin, Tornatora, et al., 2018; Mason, Scrimin, Zaccoletti, et al., 2018), two studies examining academic achievement (e.g., GPA; Hahnel et al., 2019; Mason et al. 2017), four studies examining Internet experience (Kammerer et al., 2013; Macedo-Rouet et al., 2020; Salmerón et al., 2020; Van Der Heide & Lim, 2016), and three studies examining historical thinking skills (Merkt et al., 2017; Merkt & Huff, 2020, experiments 1 and 2).

In summary, the results concerning research question 1 showed that the majority of individual differences highlighted as important for sourcing in theoretical models and conceptualizations have been examined. However, for the majority of the individual difference constructs, the number of studies was very few. The systematic

**Table 3** Specific studies (represented by study number) supporting and not supporting relationships between individual difference constructs and sourcing

	Proposed relationships supported	Proposed relationships not supported
<i>Reading skills and strategies</i>		
Word-level reading skills	36, 38, 52	7, 19, 27
Comprehension	21, 27, 37, 43, 47, 48, 50, 57	9, 19, 25, 26, 40, 41, 42, 49
Comprehension strategies	1, 32, 33, 59, 64	19, 61
<i>Cognitive constructs</i>		
Prior knowledge	3, 4, 7, 12, 13, 17, 24, 39, 43, 51, 54, 55, 58, 60, 62, 64	2, 5, 9, 10, 11, 19, 25, 26, 27, 30, 31, 34, 40, 41, 42, 59, 63, 65, 67
Working memory/Executive functions	7, 38, 40 <sup>a</sup>	17, 19, 40 <sup>b</sup>
Argumentative reasoning		39
<i>Motivation and engagement</i>		
Task values	28, 69	
Interest	34, 62, 64	2, 5, 9, 35, 54, 57, 58, 59, 61
Attitudes	29, 40, 67	27, 30, 61
<i>Personality</i>		
Need for cognition		16, 72
Growth vs. fixed mindset	7	
<i>Beliefs</i>		
Topic beliefs	11, 57 <sup>c</sup> , 64	16, 25, 26, 57 <sup>d</sup>
Epistemic beliefs	2, 5, 6, 23, 24, 63	15, 16, 65, 70
<i>Expertise</i>		
Domain/document/disciplinary expertise	8, 22, 35, 68, 71	20, 53, 70

*Note.* The numbers in columns two and three refer to the numbers of the respective studies in Table S1 in the supplementary material

<sup>a</sup>Psychophysiological self-regulation

<sup>b</sup>Working memory

<sup>c</sup>Post reading topic beliefs

<sup>d</sup>Prior topic beliefs

review also identified a number of individual difference constructs *not* included in these conceptualizations that have been examined in relation to sourcing.

### Relationships between Individual Differences and Sourcing

Our second research question concerned the extent to which proposed relationships between individual differences and sourcing have been supported by the empirical research. Table 4 shows the substantial variation in how sourcing was measured in the reviewed studies. In addition, different statistical analyses were performed and different types of results were reported across these studies (for



details, see Table S1 in the supplemental material). Thus, it is neither meaningful nor possible to calculate a general and comparable effect size (e.g., Hedges'  $g$ ) to examine the strength of the relationship between the various individual differences and sourcing across the studies. Our main findings regarding the second research question are summarized in Table 3, which is based on the information provided in the supplementary material (Table S1).

Within the category of *reading skills and strategies*, six studies examined the relationship between word-level reading skills and sourcing. Three of these studies (Macedo-Rouet et al., 2013, experiment 1; Macedo-Rouet et al., 2020; Potocki et al., 2020) found a statistically significant relationship between word-level reading and sourcing, whereas three studies could not identify such a relationship (Braasch et al., 2014; Florit et al., 2019; Kammerer, Meier, & Stahl, 2016). These differences in results do not seem to be a matter of the educational level among the participants (i.e., word-level reading skills being more important for younger readers), since students from elementary to high school were represented in both groups of studies (i.e., studies with and without a statistically significant relationship between word-level reading and sourcing). For example, Macedo-Rouet et al. (2013, experiment 1) found a statistically significant relationship between word-level reading and sourcing in their study of fourth and fifth graders' ability to evaluate information sources, whereas Florit et al. (2019) did not in their study of fourth graders. Likewise, Macedo-Rouet et al. (2020) identified a statistically significant relationship between word-level reading and sourcing in a study of ninth graders, whereas Kammerer, Meier, and Stahl (2016) did not.

Sixteen of the studies included in our review examined reading comprehension in relation to sourcing. Eight studies found a statistically significant relationship between reading comprehension and sourcing (e.g., Hahnel et al., 2019; Paul et al., 2019; Salmerón et al., 2020), and eight did not (e.g., Bråten, Brante, & Strømsø, 2018; Florit et al., 2019; Mason et al., 2017). Paul et al. (2019), for example, examined the effect of sourcing prompts and mutually exclusive claims on sourcing in a sample of 89 German fourth graders reading a text set about nutrition (i.e., whether a particular cereal was healthy). The participants stated whether the cereal in question was healthy and provided a justification for their decision. These justifications were coded for the frequency of source citations as well as for expertise and benevolence evaluations. Sourcing was also measured with memory for source-content links. Reading comprehension was measured with a standardized reading comprehension test. The results showed a main effect of reading comprehension on benevolence evaluations and memory for source-content links. Florit et al. (2019) also examined the relationship between reading comprehension and sourcing in a sample of fourth graders reading two text sets about unsettled topics (i.e., nutrition and video games) containing conflicting information. Sourcing was measured through the frequency of source-content links in post-reading essays. Reading comprehension was measured with a standardized test. The authors did not find a statistically significant relationship between reading comprehension and sourcing. In a clear trend, reading comprehension has primarily been measured in relation to sourcing among younger readers. Despite university students being clearly overrepresented across the 72

studies included in this review (i.e., 61.07% of the participants), only three of the 16 studies examining reading comprehension and sourcing included university-level participants, representing 26.74% of the participants in these 16 studies.

Given that students develop reading comprehension over the school years, we explored whether the age of the participants differed systematically in the studies with and without a statistically significant relationship between reading comprehension and sourcing. Both groups of studies included elementary, middle, and high school students as well as university students, and the mean age was similar between the studies with ( $M$  age = 15.36) and without ( $M$  age = 15.11) a statistically significant relationship between reading comprehension and sourcing. There was a trend towards somewhat larger samples in the studies with ( $M$  sample size = 139) than without ( $M$  sample size = 81.90) a statistically significant relationship, a difference that can obviously affect the  $p$  value.

Seven of the studies included in our review examined strategic competence in relation to sourcing. Five of these studies (e.g., Anmarkrud et al., 2014; List, Du, et al., 2019, studies 1 and 2) found statistically significant relationships between the use of comprehension strategies and sourcing, whereas two did not (Florit et al., 2019; Strømsø et al., 2020). For example, Anmarkrud et al. (2014) found a statistically significant relationship between the use of comprehension strategies and sourcing in a think-aloud study in which undergraduates read multiple and conflicting documents about a health issue. Participants' verbal utterances during reading were coded for instances of strategic processing, and essays written after the reading of the texts were coded for explicit references to sources and number of source-content links. In addition, participants rank-ordered the documents according to trustworthiness. The results showed that strategy use was positively correlated with the numbers of both explicit references to sources in essays and source-content links. In addition, strategy use was associated with low trustworthiness ranking of a low credibility document. Four of the five studies identifying statistically significant relationships between the use of comprehension strategies and sourcing measured participants' use of rather complex multiple text strategies focusing on the integration of content across documents in samples of university students. The two studies that could not identify a statistically significant relationship between the uses of comprehension strategies included younger students (fourth-grade and high school students).

Within the second category of individual differences, *cognitive factors*, 35 studies examined the relationship between prior knowledge and sourcing. Less than half of these studies ( $n = 16$ ) identified a statistically significant relationship. Bråten, Strømsø, and Salmerón (2011) examined the relationship between prior knowledge and sourcing in a sample of 128 undergraduates reading seven partially conflicting texts with varying degrees of credibility on a science topic (i.e., global warming). Prior knowledge was measured by a researcher made 17-item multiple-choice test, and sourcing was measured by having participants rate the trustworthiness of the different texts and then indicate the extent to which they considered source features when rating the trustworthiness of the texts. The results showed that students with lower levels of prior knowledge placed more trust in biased sources than did students with higher levels of prior knowledge

and that the latter put more trust in objective and balanced texts than in biased texts. These findings were corroborated by Mason et al. (2014), who had 134 ninth-graders read two text sets on conflicting topics (i.e., the health effects of cell phone use and genetically modified food). Prior knowledge was measured with a combination of open-ended and multiple-choice questions, and sourcing was measured with different rank-order tasks (i.e., ranking documents with respect to reliability) and log data (i.e., number of visits and reading time for most and least reliable documents). Prior knowledge was not related to the number of visits or the time spent on the most and least reliable documents in any of the document sets but was related to the reliability ranking of documents in both document sets.

The studies that did and did not find a relationship between prior knowledge and sourcing exhibited one interesting difference. Eleven of the 16 studies finding a statistically significant relationship between prior knowledge and sourcing measured prior knowledge using multiple-choice measures, whereas this type of measure was used in only six of the 19 studies that did not find a statistically significant relationship. Among the latter studies, prior knowledge was to a larger extent measured with open-ended questions ( $n=4$ ; e.g., Bråten, Ferguson, et al., 2014; Mason et al., 2017), true/false questions ( $n=2$ ; e.g., Kammerer, Meier, & Stahl, 2016; Ulyshen et al., 2015), or self-reported perceived knowledge ( $n=4$ ; e.g., Kammerer, Kalbfell, & Gerjets, 2016, experiments 1 and 2). Thus, how prior knowledge is measured might influence the relationship between prior knowledge and sourcing in hitherto unknown ways.

Five studies examined working memory/executive functions in relation to sourcing. Four of these studies (Braasch et al., 2014; Delgado et al., 2020; Florit et al., 2019; Macedo-Rouet et al., 2020) measured only working memory, whereas Mason, Scrimin, Tornatora, et al. (2018) examined both working memory and self-regulation. Two (Braasch et al., 2014; Macedo-Rouet et al., 2020) of the five studies found a statistically significant relationship between working memory and sourcing. For example, Braasch et al. (2014) examined sourcing in a sample of 59 upper secondary school students reading conflicting documents about a scientific topic (i.e., weather patterns). Working memory was measured with a working memory span task and sourcing was measured with a rank-order task in which participants ranked the documents from the most to the least reliable. The results showed that working memory capacity predicted the ability to discriminate between texts that varied in terms of reliability. Mason, Scrimin, Tornatora, et al. (2018) examined both working memory and self-regulation in a study of 7th graders reading conflicting information about a health topic. Working memory was measured with a complex reading span task, (psychophysiological) self-regulation was measured with heart rate variability, and sourcing was measured with a reliability ranking of documents and justifications for those rankings. Working memory was not related to sourcing but students with higher heart rate variability, that is, greater self-regulation, were more accurate in rank-ordering the documents.

In the only study examining argumentative reasoning in relation to sourcing, Mason et al. (2014) had ninth graders read multiple conflicting texts on two health issues. Argumentative reasoning skills were measured with a test on which participants identified informal reasoning fallacies in debates, and sourcing was measured

with four tasks: (a) identify the two most and least reliable texts, (b) rank the texts from the most to the least reliable, (c) record the number of visits to the most and least reliable texts, and (d) note the time spent on them. However, argumentative reasoning was not statistically significantly related to any of these sourcing measures.

In the third category, motivation and engagement, two studies (Kang et al., 2011; Westerwick, 2013) examined task values in relation to sourcing, both finding a statistically significant relationship. Both Kang et al. (2011) and Westerwick (2013, study 1) referred to the individual difference they measured as “topic involvement,” but, given the wording of the items used in their measures (e.g., “value” and “importance”), we considered them measures of task values. Westerwick (2013, study 1) had 574 undergraduates read two online articles on a health topic for which the credibility of the source was manipulated (i.e., high, medium, or low credibility). Task values were measured with four items on a Likert-type scale, and sourcing was measured with a scale on which participants rated the believability, accuracy, trustworthiness, bias, and completeness of the sources. Regression analysis showed that task values predicted the ability to discriminate between high-, medium-, and low-credibility sources.

Of the 12 studies addressing the relationship between interest and sourcing, only three (List, Stephens, & Alexander, 2019; Strømsø et al., 2010; Tarchi, 2019) identified a statistically significant relationship between interest and sourcing. Strømsø et al. (2010) had 126 undergraduates read seven conflicting texts on a science topic (i.e., global warming). Topic interest was measured with a 12-item questionnaire, and sourcing was measured through memory for source-content links. List, Stephens, and Alexander (2019) is the only study included in the review that distinguished between situational and individual interest (Hidi, 2001). In their study, 197 undergraduates read six conflicting texts on a social science topic (i.e., the Arab spring in Egypt) that varied in regard to document type (e.g., essay, blog entry, or newspaper article) and credibility. Individual interest was measured with five items asking participants about their general interest in this topic, whereas situational interest was measured by having them rate the interestingness of each text as they accessed it. Sourcing was measured by the number of citations in the text participants wrote after reading the documents. Situational interest was positively correlated with sourcing, but individual interest was a negative predictor of sourcing in a hierarchical regression analysis. A mediation analysis did not find a mediated effect of situational interest on the number of citations in participants’ written responses via the time participants devoted to text access.

Of the six studies examining the relationship between attitudes and sourcing, four of the studies identified a statistically significant relation. For example, Kobayashi (2014) had 154 undergraduates read two conflicting texts about a health topic (i.e., the relationship between blood type and personality). Attitudes regarding the topic were measured by asking participants to rate four attitude statements on a Likert-type scale, and sourcing was measured with a composite called source acceptability, which was based on the rating of the credibility and persuasiveness of the two texts. In a multivariate analysis of variance, prior attitude was found to be a statistically significant predictor of source acceptability. In another study, van Strien et al. (2016) had 79 university majors read eight conflicting websites on a health topic

(i.e., organic food). First, attitudes regarding the topic were measured with a 15-item questionnaire before the strength of these attitudes was measured with three items. Sourcing was measured in two ways: using eye-tracking data, attention to source information was calculated based on total fixation time on website logos and “about us” information, and credibility judgments were measured by having participants rate the trustworthiness, expertise, and convincingness of each website. Attitude strength correlated negatively with fixation time on source information and credibility rating for pages with attitude-inconsistent information. Moreover, attitude strength correlated positively with credibility rating for pages including attitude-consistent information. Additionally, there was an interaction of attitude strength with consistency/inconsistency on fixation time on source information.

Only three studies included in our review examined the category of personality in relation to sourcing. Neither of the two studies (Bromme et al., 2015, experiment 2; Winter & Krämer, 2012) that examined the need for cognition in relation to sourcing found a statistically significant relationship. Braasch et al. (2014) investigated whether a growth versus a fixed mindset (i.e., viewing intelligence malleable versus stable) was related to sourcing. Participants were 59 upper secondary students reading six conflicting texts with varying credibility on a science topic (weather patterns). Mindset was measured with an eight-item instrument, and sourcing was measured by having participants rank-order the texts from the most to the least reliable. The results showed that a growth mindset was positively and a fixed mindset negatively correlated with the ability to distinguish between texts of varying reliability.

Sixteen studies examined the relationship between the category of beliefs and sourcing. Among the six studies addressing topic beliefs, three (Bråten et al., 2016; Salmerón et al., 2020; Tarchi, 2019) found a statistically significant relationship with sourcing. Salmerón et al. (2020) had 207 4th–6th graders read three conflicting web pages about a health issue (i.e., tap or bottled water). Topic beliefs were measured both before and after the reading of the three web pages with five items asking about preferences and beliefs about tap and bottled water. Higher scores on this measure indicated favoring bottled water. Sourcing was measured with the number of explicit references to source features in texts participants wrote after reading the web pages, memory for source features, and source-content links. Topic beliefs measured before reading were not related to any of the sourcing measures. Post-reading topic beliefs correlated negatively with source-content links but not with any of the other sourcing measures. Bråten et al. (2016) had 71 undergraduates read one of two versions of a text about a health topic (i.e., potential health effects of cell phone use). The two text versions were identical, except for the concluding paragraph, in which one text concluded that no health risks were related to the use of cell phones, and the other version concluded the opposite. Topic beliefs were measured with two items asking participants to rate their agreement on whether cell phones should be considered a health risk. To measure sourcing, participants were asked to describe the text they had read and were given points for each source feature they mentioned. The results showed no relation between topic beliefs and sourcing; yet an interactive effect of topic beliefs with text version was found on sourcing.

Concerning epistemic beliefs, six of 10 studies examining the relationship between epistemic beliefs and sourcing found a statistically significant relationship between these variables. For example, Barzilai and Eseth-Alkalai (2015) had 170 undergraduate and graduate students read four blog posts about a scientific topic (i.e., desalination of sea water) that were written by authors varying in professional backgrounds (i.e., economists or hydrologists). The study included two conditions. In the conflicting condition, two of the blog posts were in favor of desalination, whereas the two other opposed it. In the convergent condition, all blog posts were in favor of desalination. Topic-specific epistemic beliefs were measured with a measure capturing three epistemic perspectives: an absolutist perspective that views knowledge as objective and certain; a multiplist perspective that views knowledge as subjective and justified by personal preferences and judgments; and an evaluativist perspective viewing knowledge as something constructed by people within a particular perspective, grounding this construction on evidence and shared standards. Sourcing was measured with three tasks. First, the participants were asked to connect sentences from the texts to one of the authors. Thus, this author-viewpoint-identification task measured participants' ability to establish source-content links. Second, participants were asked to describe the purpose and viewpoint of each of the four blogs after being given the name of the author and the title of the blog post. This author-viewpoint-description task assessed participants' ability to accurately recall and describe the viewpoint of the individual author. Third, participants were given an author-viewpoint-evaluation task consisting of two steps. In the first, participants rated whether each of the blog posts were believable, accurate, professional, balanced, reliable, correct, true, and trustworthy on a six-point scale, with a blog reliability score calculated based on the mean score on these eight items. In the second step, participants answered the open-ended question, "Why is the blog reliable or not reliable, in your opinion?" for each of the four blogs. Participants had the texts available when providing their answer, and the answers were coded for mentions of the author's viewpoint as justification. The results showed that neither an absolutist nor an evaluativist perspective correlated with any of the sourcing measures. However, a multiplist perspective correlated negatively and statistically significantly with the author-viewpoint-identification and author-viewpoint-evaluation tasks. Based on the scores on the three sourcing measures, a latent variable called author-viewpoint comprehension was created. In a regression analysis, both absolutist and multiplist perspectives were statistically significant negative predictors of author-viewpoint comprehension, whereas an evaluativist perspective was a positive predictor.

The level of specificity in the measurement of epistemic beliefs seemed to be associated with whether a statistically significant relationship was found with sourcing. Most studies that did not find any relationship between epistemic beliefs and sourcing used either domain-general measures of epistemic beliefs (Ulyshen et al., 2015) or more domain-oriented measures, such as epistemic beliefs about medicine (e.g., Bromme et al., 2015, experiment 1). On the other hand, the six studies reporting a relationship, to a larger degree, used topic- or task-specific measures (e.g., Barzilai & Eseth-Alkalai, 2015; Barzilai & Zohar, 2012; Strømsø et al., 2011).

Finally, in the last category, expertise, we could not identify any studies examining multiple document proficiency in relation to sourcing. Hence, all eight studies addressing expertise concerned domain/document/disciplinary expertise. Five of these studies (Brand-Gruwel et al., 2017; Herrero-Diz et al., 2019; Lucassen et al., 2013; von der Muhlen et al., 2016; Wineburg, 1991) found a statistically significant relationship between domain/document/disciplinary expertise and sourcing. Brand-Gruwel et al. (2017), for example, examined sourcing in a sample of 19 novices (first semester psychology students) and 16 experts (university teachers with a PhD in psychology) reading web pages about two psychological topics (i.e., human memory and altruism). Based on two search engine results pages (SERPs) with 17 and 18 links, participants were asked to select five web pages on each of the topics and rank (prioritize) them. Sourcing was measured by the percentage of web pages on which source information was scanned (eye-tracking) and to what degree participants selected and ranked the most trustworthy web pages. The results showed domain expertise had no effect on the number of web pages on which source information was scanned, but a main effect of domain expertise was found on the ability to identify and select the most trustworthy web pages.

In summary, for research question 2, the results of the systematic review indicated that the empirical backing is rather ambiguous with respect to the relationships between individual differences and sourcing that have been suggested in contemporary conceptualizations of multiple document representation and use.

### **The Potential Role of Sourcing Measurement**

Our third research question concerned the possibility that relationships between individual differences and sourcing might vary with the way sourcing is measured. As shown in Table 4, we identified 21 different measures of sourcing in the 72 studies included in the review, with the most frequently used sourcing measures being various questionnaires asking participants to rate the credibility or trustworthiness of sources (25 studies), citations in written products (e.g., essays, written arguments; 18 studies), and various types of questionnaires examining participants' representation of source-content links (14 studies). Given the low frequency in the use of some of these types of sourcing measures, one should be careful drawing firm conclusions regarding the role of sourcing measurement. However, in an interesting trend, measures of sourcing that examined participants' spontaneous sourcing (e.g., citations in a written product, log data, etc.) seemed to differ from measures of sourcing in which participants were prompted to source (e.g., questionnaires on which the participants rank or rate sources, source selection, memory for source features, etc.), with the latter type apparently yielding more positive findings regarding the relationship between individual differences and sourcing.

### **The Potential Role of Domain or Topic**

Our fourth research question concerned the possibility that relationships between individual differences and sourcing might vary with the domain or topic

**Table 4** Number of reviewed studies supporting and not supporting relationships between individual differences and sourcing by sourcing measure

Sourcing measure	Total number of studies	Support for relationships between individual differences and sourcing	No support for relationships between individual differences and sourcing
Citations in written product	18	5	13
Source-content links in written product	3	1	2
Questionnaire: Source ranking	8	5	3
Questionnaire: Source rating	25	20	5
Questionnaire: Source-content links	14	10	4
Source selection	9	7	2
Source evaluation (open questions)	2	1	1
Think-aloud: Source-content links	1	1	0
Think-aloud: Source-source links	1	1	0
Think-aloud: Source evaluation	8	6	2
Think-aloud: Mention of source features	1	0	1
Think-aloud: Source representation/citation	1	1	0
Identification of source information in text	2	2	0
Eye-tracking: Fixation time on source info	6	2	4
Memory for source features	7	5	2
Justification: Source selection	5	3	5
Justification: Source ranking	5	0	5
Justification: Source rating	5	5	0
Justification: Use of source information to justify position	3	2	1
Log-data (reading time source information/accessing source information)	5	2	3
Composite variable (combination of 2 or more measures of sourcing)	5	4	1

The same study can use more than one measure of sourcing



addressed in the reading materials. Across the 72 studies included in this review, we identified six domains or topics in the textual materials the participants read: health (23 studies; e.g., Barzilai & Zohar, 2012; Bromme et al., 2015, experiment 1; Mason et al., 2014), science (22 studies; e.g., List, Du, et al., 2019, study 1; Strømsø et al., 2010), history (10 studies; e.g., Barzilai et al., 2020, study 1; Merkt et al., 2017; Rouet et al., 1997), social science (6 studies; e.g., Kang et al., 2011; List et al., 2017; Winter & Krämer, 2012), psychology (2 studies; Brand-Gruwel et al., 2017; von der Muhlen et al., 2016), and restaurant reviews (1 study; Van Der Heide & Lim, 2016). In addition, we identified seven studies (e.g., Hahnel et al., 2019; Macedo-Rouet et al., 2020; Salmerón et al., 2016) using reading materials from two or more domains and one study (Herrero-Diz et al., 2019) that did not specify the topic or domain of the reading material. Given the low number of studies within some of these topics, we focused on the studies using reading materials from the three most frequently studied topics: health, science, and history.

As Table 5 shows, no clear general trend indicated that the relationships between individual differences and sourcing varied with the domain or topic addressed in the reading materials. However, some interesting differences can be seen when comparing the studies in which participants read texts about health and science. These two topics are particularly interesting given that the number of studies was comparable: 23 for health and 22 for science. One difference was that reading skills were examined in relation to sourcing to a substantially larger degree in the studies in which participants read about health topics, with this difference driven by reading comprehension, in particular. The majority of the studies (7 of 10) examining reading comprehension and sourcing did not find a statistically significant relationship, and the one study examining reading comprehension and sourcing when participants read about a science topic identified a significant relationship. The relationship between prior knowledge and sourcing also seemed to differ in regard to these two topics. Only 5 of the 16 studies examining prior knowledge and sourcing when participants read documents about health topics found a statistically significant relationship, whereas 9 of the 14 studies found a significant relationship when participants read about science topics.

Finally, the studies in which participants read about historical topics examined very few individual difference factors. Only one study (Wiley et al., 2020) reported results about the relationship between more than one individual difference construct and sourcing.

### **In-Depth Analyses of Relationships between Reading Comprehension, Prior Knowledge, Interest, Epistemic Beliefs, and Sourcing**

Given the mixed findings we obtained regarding relationships between individual differences and sourcing, we further explored relationships between the four most researched individual differences (i.e., reading comprehension, prior knowledge, interest, and epistemic beliefs) and sourcing in a set of more fine-tuned analyses. In

**Table 5** Number of reviewed studies supporting and not supporting relationships between individual differences and sourcing by domain

	Health			Science			History		
	s	ns	ns	s	s	ns	s	s	ns
	<i>Reading skills</i>								
Word-level reading skills	-	2	2	1	1	2	-	-	-
Reading comprehension	3	7	-	1	1	-	-	-	-
Comprehension strategies	3	1	1	2	2	1	-	-	-
<i>Cognitive constructs</i>									
Prior knowledge	5	11	5	9	9	2	-	-	-
Working memory/executive functions	1	2	2	1	1	2	-	-	-
Argumentative reasoning	-	1	-	-	-	-	-	-	-
Topic familiarity	-	-	-	1	1	-	-	-	-
Verbal comprehension	-	1	-	-	-	-	-	-	-
<i>Motivation and engagement</i>									
Interest	1	3	6	1	1	-	-	-	-
Attitudes	3	1	1	-	-	1	-	-	-
Task value	1	-	-	-	-	-	-	-	-
<i>Personality</i>									
Need for cognition	-	-	1	-	-	1	-	-	-
Growth vs. fixed mindset	-	-	-	1	1	-	-	-	-
<i>Beliefs</i>									
Topic beliefs	1	3	2	1	1	2	-	-	-
Epistemic beliefs	2	2	1	4	4	1	-	-	1
<i>Expertise</i>									
Gender	1	1	-	1	1	-	1	-	3
<i>Educational level</i>									
Emotional and affective constructs	1	1	-	2	2	-	1	-	-

**Table 5** (continued)

	Health		Science		History	
	s	ns	s	ns	s	ns
<i>Achievement/grades</i>	-	1	-	-	-	-
<i>Frequency of Internet/social media use</i>	1	-	-	1	-	-
<i>Historical thinking skills</i>	-	-	-	-	3	-

s, supporting; ns, not supporting

these analyses, we carefully attended to how both individual differences and sourcing were assessed in each study, descriptive information about participants' scores on these assessments, psychometric properties of these scores, and participants' level of performance. Detailed information about all these dimensions are included in Tables S2–S5 in the supplemental materials. In the following, we summarize the additional patterns we were able to discern based on these in-depth analyses.

## Reading Comprehension and Sourcing

Our in-depth analysis of the association between reading comprehension and sourcing indicated that the mixed findings regarding reading comprehension, at least in part, were due to how reading comprehension was measured across studies (see Table S2 for details). In the 16 studies examining reading comprehension in relation to sourcing, eight studies used multiple-choice tests (e.g., Mason et al., 2020; Paul et al., 2018, studies 1 and 2), four used cloze tests (e.g., Bråten, Brante, & Strømsø, 2018; Kammerer, Kalbfell, & Gerjets, 2016, experiments 1 and 2), two used open-ended questions (Macedo-Rouet et al., 2013; Salmerón et al., 2020), one used a statement veracity task (Merkt et al., 2017), and one used a multiple source comprehension task (Hahnel et al., 2019). In a clear trend, studies using open-ended questions and requiring extensive inferencing found a relationship between reading comprehension and sourcing (Hahnel et al., 2019; Macedo-Rouet et al., 2013; Salmerón et al., 2020), whereas reading comprehension was not related to sourcing in studies in which the former construct was measured with cloze tests (Bråten, Brante, & Strømsø, 2018; Kammerer, Kalbfell, & Gerjets, 2016, experiments 1 and 2; see, however, Kammerer, Meier, & Stahl, 2016). This trend seemed to be independent of whether sourcing was prompted or unprompted.

The findings were more mixed when reading comprehension was measured with multiple choice tests. However, we noted that most of the studies that did not find any relationships measured only unprompted sourcing in essays, with quite low scores obtained by participants. Also, most of these studies did not inform about the level of inferencing required by the reading comprehension measure.

Taken together, this analysis may suggest that relationships between reading comprehension and sourcing can only be expected to the extent that comprehension measures require the construction of mental representations, preferably at the level of situation(s) model representation (Kintsch, 1988; Perfetti et al., 1999). Further, it may be challenging to demonstrate a relationship when sourcing is unprompted and, consequently, characterized by low score variance.

## Prior Knowledge and Sourcing

Our in-depth analysis of the relationship between prior knowledge and sourcing confirmed that how prior knowledge is measured, indeed, seems to matter (see Table S3 for details). Whereas all the studies measured prior knowledge at a topic-specific level, the formats of the assessments varied, with 17 studies using multiple-choice items (e.g., Barzilai et al., 2015; Kammerer et al., 2021; Stang Lund et al., 2017),

seven studies using open-ended questions (e.g., Braasch et al., 2014; Bråten, Ferguson, et al., 2014; Mason et al., 2017), six studies using perceived knowledge items (Bråten et al., 2016; Kammerer, Kalbfell, & Gerjets 2016, experiment 1; van Strien et al., 2016), three studies using a term identification task (List, 2014; List et al., 2017; List, Stephens, & Alexander, 2019), and two studies using true/false measures (Kammerer, Meier, & Stahl, 2016; Ulyshen et al., 2015).

When prior knowledge was measured with multiple-choice items and open-ended questions and participants displayed a certain level of prior knowledge (e.g., on average obtained at least 40% of maximum score), the relationship with sourcing was mostly statistically significant, especially when sourcing was prompted and, thus, resulted in higher scores on the sourcing measures. As an example, Stang Lund et al. (2019), who had 140 upper-secondary students read documents on a health topic, found that students, on average, scored between 74 and 79% of maximum score on a multiple-choice prior knowledge measure and between 51 and 65% of maximum score on a prompted memory for source-content links sourcing measure. In that study, prior knowledge correlated with sourcing ( $r=0.31$ ), and path analysis showed a direct effect of prior knowledge on sourcing ( $\beta=0.21$ ) as well as an indirect effect via memory for textual conflicts ( $\beta=0.07$ ). In contrast, in studies in which prior knowledge was very low and studied in relation to unprompted sourcing (e.g., Mason, Scrimin, Zaccoletti, et al., 2018), studies in which only sourcing was very low (e.g., Florit et al., 2019), and studies in which both prior knowledge and sourcing were low (e.g., Bråten, Ferguson, et al., 2014), relationships between prior knowledge and sourcing tended to be statistically non-significant, presumably due to a lack of score variance in prior knowledge, sourcing, or both.

Regarding other types of prior knowledge assessments, none of the studies using true/false measures or term identification tasks, and only two of the six studies measuring perceived prior knowledge (Barzilai et al., 2020, experiments 1 and 2) found statistically significant relationships with sourcing. Although it did not seem to matter whether sourcing was prompted in these studies, participants displayed quite low prior knowledge in the studies using these measures, presumably reducing the functional value of participants' knowledge in relation to sourcing. Taken together, the results of this fine-tuned analysis thus seem to suggest that higher (List & Alexander, 2019) rather than lower (Stadtler & Bromme, 2014) prior knowledge is conducive to sourcing activities.

## Interest and Sourcing

Although the 12 studies that examined interest in relation to sourcing generally found that participants reported substantial individual interest in the topics discussed across documents and measured interest with high reliability, they did not produce much evidence for the importance of individual interest for the subprocess of sourcing (see Table S4 for details). In fact, 10 of the 12 studies found no positive relationships between these variables and one study (List, Stephens, & Alexander, 2019) even found that individual interest negatively predicted sourcing (i.e., the number of citations in written products) when other relevant variables, including prior knowledge, were controlled for. Further, in the two studies that found relationships

between individual interest and sourcing (Strømsø et al., 2010; Tarchi, 2019), correlations were low and statistical significance possibly due to larger samples being included in those studies.

In the only study (List, Stephens, & Alexander, 2019) that measured participants' situational interest during reading in addition to their individual interest (Hidi, 2001), participants who rated the interestingness of the documents higher during reading were also more likely to include citations to those sources in their post-reading essays. The correlation between situational interest and sourcing was quite low, however ( $r=0.15$ ).

Thus, although motivation and engagement recently have been highlighted within models of multiple document literacy (Britt et al., 2018; List & Alexander, 2019), asking participants to self-report their individual interest in the topic of the documents independent of the reading task context does not seem to be a valid indicator of such motivation and engagement. As noted by Bråten, Brante, and Strømsø (2018), possible reasons for this include that students who rate themselves highly on such scales may still not engage much in concrete, challenging multiple document tasks and sometimes even disregard source information because they rely on their own personal opinion about the issue in question. An alternative is therefore to have students rate their task-based or text-based (i.e., situational; Hidi, 2001) interest during reading, as was done by List, Stephens, and Alexander (2019). Another possibility is to focus on their behavioral engagement when working with the documents, that is, on their active, observable involvement in multiple document tasks as typified by time, effort, persistence, and productivity (Bråten, Brante, & Strømsø, 2018; Bråten et al., 2021).

## Epistemic Beliefs and Sourcing

Our in-depth analysis of epistemic beliefs in relation to sourcing (see Table S5 for details) indicated that beliefs in absolute certain knowledge, reliance on personal opinions, and oversimplification of complex issues were negatively related to sourcing activities, whereas acknowledging the tentativeness of knowledge, relying on evidence and expertise, and realizing the need to justify knowledge claims by means of relevant resources both internal and external to the individual were positively related to sourcing (Barzilai & Eshet-Alkalai, 2015; Barzilai et al., 2015; Kammerer et al., 2013, 2021; Strømsø et al., 2011). This trend was most salient when epistemic beliefs were measured at a topic-specific level but was also present in some studies in which epistemic beliefs were measured at a domain-general or domain-specific level (Kammerer et al., 2013, 2021). However, we noted that three of the four studies that did not find any relationships between epistemic beliefs and sourcing did not measure epistemic beliefs at a topic-specific level (Bromme et al., 2015, experiment 1; Ulyshen et al., 2015; Wiley et al., 2020), and with only one of these (Bromme et al., 2015, experiment 1) targeting the justification dimension found to be important in other research.

Taken together, the results of this fine-tuned analysis seem to be consistent with the view that beliefs in the justification of knowledge claims by relying on evidence

and competent sources and using procedures such as cross-checking multiple sources may promote adaptive sourcing, as recently highlighted within the integrated framework of multiple texts (List & Alexander, 2019; see also, Bråten, Britt, et al., 2011). Further, measuring epistemic beliefs in relation to the specific topic discussed across documents seems to be the better option when investigating links between such beliefs and sourcing.

## Discussion

This systematic review sought to advance current knowledge on sourcing by providing answers to four research questions regarding the role of individual differences in attending to, representing, evaluating, and using information about the sources of document content.

### Inclusion of Individual Differences

Our first research question addressed the extent to which the reviewed studies examined the individual differences highlighted in the various conceptualizations of multiple document literacy that included sourcing. The results indicated that the majority of these reader characteristics were considered, with the prevalence of cognitive factors followed by factors related to reading, motivation, beliefs, expertise, and personality.

The most examined individual difference concerned what readers already know about the domain or topic relevant to the document content, which is understandable given its influence on single-text comprehension (e.g., Ozuru et al., 2009). Further, more than half of the studies that investigated reading-related individual differences unsurprisingly examined the role of reading comprehension of single texts in the comprehension and evaluation of multiple texts. Regarding the motivational factors, it should be noted that two heavily researched individual difference constructs within motivation, achievement goals and self-concept of ability, were not included in the reviewed studies. According to the RESOLV model (Britt et al., 2018), they can be assumed to act as resources in constructing mental representations of the reading context and the task, which guide readers' processing of the documents. In research on single-text comprehension and learning from text, mastery achievement goals have been shown to contribute to conceptual learning from science texts (Johnson & Sinatra, 2014). Likewise, self-concept of ability has been linked to reading performance (e.g., Colmar et al., 2019; Sewasew & Koester, 2019). Thus, future studies that include hitherto disregarded yet potentially influential motivation constructs could extend current understanding of sourcing within multiple document literacy. Personality factors have also been very sparsely examined in relation to sourcing, and they are much less frequently assessed in educational psychology research, presumably because they are considered individual differences that are difficult to change through educational interventions. In our review, the only personality factors analyzed at all were need for cognition and growth/fixed mindset, which might

be conceived of as more modifiable, at least to some extent, in appropriate learning environments. These constructs have also been investigated in other research areas, for example, within conceptual change learning in science (Taasobshirazy & Sinatra, 2011) and academic performance in mathematics (Bostwick et al., 2017).

The results showed that the individual differences that were not highlighted in the conceptualizations of multiple-document representation and use, yet investigated in the reviewed studies, spanned the already identified categories of cognitive and motivational factors. Moreover, demographic variables were investigated, such as gender, age, and parental educational level. In addition, some of the studies used educational level as an individual difference variable. Examined individual differences also included experiences with the Internet in terms of usage of the Web, social media, and a specific platform. Few studies concerned the role of emotions in sourcing, which is consistent with the fact that the interplay of cognition and emotion in reading processes is a relatively new area of research (e.g., Trevors et al., 2017; Zaccoletti et al., 2020). Of note is that in all the four studies in this area that were included in our review, affective states were measured physiologically in the school context by recording skin conductance or cardiac activity and using related indices of arousal and self-regulation.

### **Relationships between Individual Differences and Sourcing**

Our second research question asked to what extent the proposed relationships between individual differences and sourcing were supported by empirical research. Overall, the results showed that such relationships emerged in half or slightly more than half of the reviewed studies, indicating that the empirical backing is not strong or, in some cases, clear. The reviewed studies varied in the kinds of analyses performed to investigate the links between individual differences and sourcing. For example, some of the investigations reported bivariate correlations, while others did not, and when individual differences were considered together with control variables in regression analyses, this analytic approach impacted the relationship between an examined individual difference construct and sourcing. However, variety is not only a negative aspect that may make a synthesis of results more difficult and less clear; it can also be considered a positive aspect to some extent, for example because it might indicate whether a range of sourcing tasks share an underlying cognitive mechanism.

Specifically, when considering both word-level reading skills and higher-level reading skills related to comprehension, proposed relationships emerged from approximately half of the studies across age groups and educational levels. These findings show that additional studies are needed to address the role of these reading-related factors in sourcing. Still, the results seem to rule out the possibility that word-level skills and reading comprehension play a role only when sourcing is considered in relatively young readers. With respect to reading comprehension, in particular, our in-depth analysis also suggested that relationships with sourcing may depend on the extent to which reading comprehension measures require mental representations, preferably at the level of situation(s) models (Kintsch, 1988; Perfetti et al., 1999).



Further, the contribution of comprehension strategies to sourcing seemed to be more consistent. However, in general, complex comprehension strategies were considered, and only when readers younger than college and university students were involved did a statistically significant relationship not emerge. This outcome might suggest that comprehension strategies need to be mastered well to play a role in students' sourcing.

Among the cognitive constructs, more than half of the studies documented a contribution of prior knowledge to sourcing. The relationship between these two variables seemed to depend on the measurement of prior knowledge. More specifically, our in-depth analysis of prior knowledge indicated that a certain level of prior knowledge may be needed for this construct to gain any functional value in relation to sourcing (List & Alexander, 2019), with quite a few studies including participants with very low prior knowledge resulting in little score variance, sometimes just to demonstrate that participants were novices with respect to the topic discussed across documents. Not surprisingly, working memory was found to be related to sourcing when readers were tasked with rank-ordering a number of increasingly reliable texts.

Among the motivational constructs, task value and attitudes were found to be most consistently related to sourcing, whereas the proposed role of interest (Britt et al., 2018; List & Alexander, 2019) was generally not confirmed. Interest can be measured at three levels of specificity and stability (Renninger & Hidi, 2011). At the broadest level, individual interest concerns a domain (e.g., science) and can be considered stable and long term. At the narrowest level, situational interest concerns a particular context or task that attracts students' attention and, thus, can be considered unstable and short term. Topic interest is situated between these levels as it concerns individual interest in a specific topic within a domain (e.g., climate change) but, still, can be considered more stable and long term than situational interest. Although most studies measured interest at the topic level, which may seem appropriate for the purpose, the empirical backing of a relationship between this variable and sourcing was weak. This result might be due, at least partially, to low prior knowledge of the topic in question, leading participants to rate their level of interest inaccurately. In addition to an affective component related to emotional engagement in an event or activity, interest has a cognitive component related to the knowledge a person brings to the activity and other aspects of cognitive functions (O'Keefe & Harackiewicz, 2017). However, another possibility suggested by our in-depth analysis of interest in relation to sourcing is that this construct needs to be measured at the situational level or that behavioral indicators need to be used. This is because the challenging, complex tasks that participants typically are presented with in this area of research may require an active involvement that is not captured by self-reports of topic interest independent of the reading task context (Bråten, Brante, & Strømsø, 2018).

For the personality constructs, not much support for the proposed relationships was found. Very few studies examined these constructs in relation to sourcing, and only growth versus fixed mindset was found to be linked to sourcing. Further research is obviously needed to shed light on the role that personality may play in sourcing.

Among the belief constructs, about half of the investigations we reviewed provided empirical support for proposed relationships between sourcing and beliefs about the topic of the texts or about knowledge and knowing (i.e., epistemic beliefs). Regarding epistemic beliefs, the results seemed to suggest that scores on topic- or task-specific measures were more consistently related to sourcing than were scores on measures targeting less specific beliefs. Further, our in-depth analysis indicated that beliefs in justification of knowledge claims by attending to evidence, competence, and consistency across multiple sources were positively related to adaptive sourcing, which is consistent with theoretical accounts (Bråten, Britt, et al., 2011; List & Alexander, 2019). Finally, some support for the relationship between individual differences and sourcing was provided by research on the role of expertise, broadly conceived.

### **The Potential Role of Sourcing Measurement**

Our third research question addressed whether relationships between individual differences and sourcing might vary according to the way sourcing is measured. The results showed that a large number of measures were used in the reviewed studies, some more frequently than others. Although substantial variation in this regard may make it more difficult to compare findings across studies, it also highlights the multiple ways an advanced literacy skill such as sourcing can be represented among students across educational levels.

A particularly relevant issue seems to be the distinction between unprompted (i.e., spontaneous) and prompted sourcing. Spontaneous sourcing was targeted, for example, in studies where participants referred to source features when trying to explain conflicting positions on the same issue in an essay task, without receiving any specific instructions to refer to sources (Anmarkrud et al., 2014; Bråten, Brante, & Strømsø, 2018; Kammerer, Kalbfell, & Gerjets, 2016). In contrast, prompted sourcing was targeted, for example, in studies where participants were explicitly asked to rank-order a set of documents according to their credibility (e.g., Mason, Scrimin, Tornatora, et al., 2018) or to indicate from which sources specific content within a document set originated (i.e., memory for source-content links; Strømsø et al., 2010). On the one hand, participants seemed to display rather poor spontaneous sourcing, and the consequential low variance in the adopted measures makes it difficult to identify a statistically significant relationship with individual difference factors. On the other hand, in the studies in which participants were prompted to source, they differed more in sourcing skills, and a significant relationship was thus more likely to emerge.

### **The Potential Role of Domain or Topic**

Our fourth research question addressed whether relationships between individual differences and sourcing might vary according to the domain or topic addressed in the reading materials. Although a clear conclusion could not be drawn, an

interesting variation concerned the contribution of some individual differences, such as prior knowledge and epistemic beliefs, to sourcing when considering the two most investigated domains, health and science. Unlike materials about health issues, those concerning science referred to an academic subject or were closely related to it. It is therefore plausible that what students already know about the subject content makes a difference in sourcing. It is also likely that epistemic beliefs may contribute more consistently to sourcing when participants read science than health documents. Through various activities and tasks involved in learning science as an academic subject, students have many opportunities to reflect on the nature of knowledge and the process of knowing concerning scientific topics, leading to epistemic understanding at various levels (e.g., Kuhn et al., 2000; Yang et al., 2016).

### Limitations and Future Directions

As with every review, this one is not without limitations. We first note that the outcomes are based on samples in which females number almost twice as many as males. This finding is not surprising considering that females are overrepresented among the university populations of psychology and education students, which more frequently participate in the types of studies selected for this review. The inclusion of more gender-balanced samples will allow more solid and generalizable results.

A related limitation is that the majority of the studies involved university students across the USA and Europe, whereas younger students, especially in elementary and middle school, were underrepresented. The underlying reasons are quite obvious, as sourcing is a complex process and can be examined only to a limited extent in younger students. Moreover, it can be practically more difficult to involve them in this type of research, as it is time demanding and often does not fit within school constraints. However, more investigations on younger students who start being confronted with sourcing issues will extend our scientific knowledge about the development of critical thinking and literacy skills and also have implications for educational practice.

Although the 72 studies we reviewed were conducted in nine different countries on three continents, a substantial portion of this research was conducted by a limited number of research groups. We have no reason to believe that this has biased our results in any way. Still, research from an even broader research community would have been desirable. Given the steep increase in interest in this area of research in the last decade, a larger number of research groups can be expected to contribute in the coming years.

With respect to the presentation of documents, findings from studies in which participants search for and select documents themselves are essentially lacking. To what extent certain individual differences (e.g., executive function and self-regulation) become more pertinent in less constrained reading contexts is therefore an issue wide open for further research.

Some limitations more directly concern the variables of interest in this review. In some cases, the labels used to describe the examined variables were unclear. Although this is an issue that does not concern only research on sourcing and

multiple text comprehension, further studies in this area would benefit from terminological clarification, such that variables related to prior knowledge, reading skills, and motivation have clear and unique referents. Moreover, some individual differences need to be investigated more deeply. For example, future research on personal relevance might shed light on the potential role of this variable as a moderator or mediator in sourcing and decision-making processes (Kobayashi, 2014). Some topics appear to be very important from a disciplinary perspective but are hardly perceived as personally relevant by students, especially at younger ages. If the topic of a set of documents has high personal relevance for readers, they may be more likely to invest effort not only in processing document content deeply, which is demanding in itself, but also to pay attention to sources and construct source-content links. Thus, studies comparing the reading of multiple documents that vary with respect to the personal relevance of the topic may extend our understanding of constructs that contribute to sourcing.

Further, in-depth investigations of the role of emotions in sourcing are highly needed. Emotions are not explicitly included in the models described in the background analysis, although both the CAEM (List & Alexander, 2017) and the IF-MT (List & Alexander, 2019) refer to affective engagement, which signals attention to motivational and emotional factors. According to the theory of achievement emotions by Pekrun (2006), motivational aspects such as control and value are antecedents of emotions experienced in relation to specific learning activities and tasks. Readers' emotional reactivity can also be considered an individual tendency to react more or less intensely to emotional stimuli. A set of documents can be more or less emotionally charged depending on content and language style, which means that textual characteristics can interact with individual reactivity to emotional materials and potentially influence the process of sourcing (Bohn-Gettler, 2019).

Finally, we restricted the current review to the role of individual differences in the subprocess of sourcing, leaving aside another important component of multiple document comprehension, that is, the integration of content across documents. Although sourcing can be considered pivotal to multiple document comprehension, and although this selectiveness allowed us to provide an in-depth analysis of the role of several individual difference factors, our approach leaves open the question of whether individual differences might operate differently for content integration than for sourcing. Given theory and extant research within a single text paradigm, there is reason to believe, for example, that prior knowledge may play a more consistent role in content integration than in sourcing (McNamara & Magliano, 2009), and it is difficult to imagine that single text comprehension would not be strongly involved in understanding the content of multiple documents (Florit et al., 2019; Mahlow et al., 2020). To address this question beyond speculation, future systematic reviews of the role of individual differences in content integration within multiple document literacy are needed. In future reviews, it may also be pertinent to focus on one particular individual difference factor at a time, including how the role of that factor in multiple document literacy has been grounded theoretically and how it has been linked empirically to a broader array of multiple document literacy tasks.

## Conclusion

In the current review, we provide a catalogue of individual differences figuring within models of multiple document literacy as a basis for scrutinizing to what extent the assumptions have been empirically supported. At the same time, we did not want to be restricted by the individual differences included in those models, also exploring more exhaustively the individual differences that actually have been researched in relation to sourcing and, thus, providing a catalogue of the individual differences falling outside the scope of the extant models as well. Hopefully, by adopting this approach, our review may encourage theorists in this area to carefully consider the empirical evidence when creating their models. Further, it may suggest individual differences that have not been included in the models but still seem worthy of further investigation.

Despite some limitations, our review provides new insights into the role of individual differences in sourcing. The results have scientific significance because they elucidate to what extent individual differences included in theoretical models of multiple document literacy have empirical support. They also have practical significance in suggesting which individual differences may facilitate or constrain sourcing when students work with multiple documents. Based on our review, we encourage future researchers in multiple document literacy to contribute to further clarification of the role of individual differences in sourcing, as well as to evidence-based knowledge of how the teaching of sourcing skills can be adapted to individual differences among learners. In particular, we call for future research that is specifically designed to investigate the effects of individual differences on sourcing and includes measures and measurements appropriate for discovering such effects.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10648-021-09640-7>.

**Funding** Open access funding provided by University of Oslo (incl Oslo University Hospital).

## Declarations

**Competing Interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

### References marked with an asterisk (\*) are included in the review

- Afflerbach, P., & Cho, B.-Y. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 69–90). Routledge. <https://doi.org/10.4324/9781315759609>
- Alexander, P. A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational, and strategic forces. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 213–250). JAI.
- Alexander, P. A., & The Disciplined Reading and Learning Research Laboratory (2012). Reading into the future: Competence for the 21st century. *Educational Psychologist*, 47(4), 259–280. <https://doi.org/10.1080/00461520.2012.722511>
- \*Anmarkrud, Ø., Bråten, I., & Strømsø, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences*, 30, 64–76. <https://doi.org/10.1016/j.lindif.2013.01.007>
- \*Barzilai, S., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction*, 36, 86–103. <https://doi.org/10.1016/j.learninstruc.2014.12.003>
- \*Barzilai, S., Thomm, E., & Shlomi-Elooz, T. (2020). Dealing with disagreement: The roles of topic familiarity and disagreement explanation in evaluation of conflicting expert claims and sources. *Learning and Instruction*, 69, Article e101367. <https://doi.org/10.1016/j.learninstruc.2020.101367>
- \*Barzilai, S., Tzadok, E., & Eshet-Alkalai, Y. (2015). Sourcing while reading divergent expert accounts: Pathways from views of knowing to written argumentation. *Instructional Science*, 43(6), 737–766. <https://doi.org/10.1007/s11251-015-9359-4>
- \*Barzilai, S., & Zohar, A. (2012). Epistemic thinking in action: Evaluating and integrating online sources. *Cognition and Instruction*, 30(1), 39–85. <https://doi.org/10.1080/07370008.2011.636495>
- Bohn-Gettler, C. M. (2019). Getting a grip: The PET framework for studying how reader emotions influence comprehension. *Discourse Processes*, 56(5–6), 386–401. <https://doi.org/10.1080/0163853X.2019.1611174>
- Bostwick, K. C. P., Collie, R. J., Martin, A. J., & Durksen, T. L. (2017). Students' growth mindsets, goals, and academic outcomes in mathematics. *Zeitschrift für Psychologie*, 225(2), 107–116. <https://doi.org/10.1027/1027151-2604/a000287>
- Braasch, J. L. G., & Bråten, I. (2017). The Discrepancy-Induced Source Comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist*, 52(3), 167–181. <https://doi.org/10.1080/00461520.2017.1323219>
- Braasch, J. L. G., Bråten, I., & McCrudden, M. T. (Eds.) (2018). *Handbook of multiple source use*. Routledge. <https://doi.org/10.4324/9781315627496>
- \*Braasch, J. L. G., Bråten, I., Strømsø, H. I., & Anmarkrud, Ø. (2014). Incremental theories of intelligence predict multiple document comprehension. *Learning and Individual Differences*, 31, 11–20. <https://doi.org/10.1016/j.lindif.2013.12.012>
- Braasch, J. L. G., Rouet, J.-F., Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition*, 40(3), 450–465. <https://doi.org/10.3758/s13421-011-0160-6>
- \*Brand-Gruwel, S., Kammerer, Y., van Meeuwen, L., & van Gog, T. (2017). Source evaluation of domain experts and novices during Web search. *Journal of Computer Assisted Learning*, 33(3), 234–251. <https://doi.org/10.1111/jcal.12162>
- Brante, E. W., & Strømsø, H. I. (2018). Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educational Psychology Review*, 30(3), 773–799. <https://doi.org/10.1007/s10648-017-9421-7>
- Bråten, I., Anmarkrud, Ø., Brandmo, C., & Strømsø, H. I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple text comprehension. *Learning and Instruction*, 30, 9–24. <https://doi.org/10.1016/j.learninstruc.2013.11.002>

- Bråten, I., & Braasch, J. L. G. (2017). Key issues in research on students' critical reading and learning in the 21st century information society. In C. Ng & B. Bartlett (Eds.), *Improving reading and reading engagement in the 21st century: International research and innovations* (pp. 77–98). Springer. <https://doi.org/10.1007/978-981-10-4331-4>
- Bråten, I., & Braasch, J. L. G. (2018). The role of conflict in multiple source use. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 184–201). Routledge. <https://doi.org/10.4324/9781315627496>
- \*Bråten, I., Brante, E. W., & Strømsø, H. I. (2018). What really matters: The role of behavioral engagement in multiple document literacy tasks. *Journal of Research in Reading*, 41(4), 680–699. <https://doi.org/10.1111/1467-9817.12247>
- Bråten, I., Brante, E. W., & Strømsø, H. I. (2019). Teaching sourcing in upper-secondary school: A comprehensive sourcing intervention with follow-up data. *Reading Research Quarterly*, 54(4), 481–505. <https://doi.org/10.1002/rrq.253>
- Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist*, 46(1), 48–70. <https://doi.org/https://doi.org/10.1080/00461520.2011.538647>
- \*Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (2014). Students working with multiple conflicting documents on a scientific issue: Relations between epistemic cognition while reading and sourcing and argumentation in essays. *British Journal of Educational Psychology*, 84(1), 58–85. <https://doi.org/10.1111/bjep.12005>
- Bråten, I., Latini, N., & Haverkamp, Y. E. (2021). Predictors and outcomes of behavioral engagement in the context of text comprehension: When quantity means quality. *Reading and Writing*. Advance online publication. <https://doi.org/10.1007/s11145-021-10205-x>
- \*Bråten, I., Salmerón, L., & Strømsø, H. I. (2016). Who said that? Investigating the Plausibility-Induced Source Focusing assumption with Norwegian undergraduate readers. *Contemporary Educational Psychology*, 46, 253–262. <https://doi.org/10.1016/j.cedpsych.2016.07.004>
- Bråten, I., Stadler, M., & Salmerón, L. (2018). The role of sourcing in discourse comprehension. In M. F. Schober, D. N. Rapp, & M. A. Britt (Eds.), *Handbook of discourse processes* (2nd ed., pp. 141–166). Routledge. <https://doi.org/10.4324/9781315687384>
- \*Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly*, 44(1), 6–28. <https://doi.org/10.1598/RRQ.44.1.1>
- \*Bråten, I., Strømsø, H. I., & Salmerón, L. (2011). Trust and mistrust when students read multiple information sources about climate change. *Learning and Instruction*, 21(2), 180–192. <https://doi.org/10.1016/j.learninstruc.2010.02.002>
- \*Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, 20(4), 485–522. [https://doi.org/10.1207/S1532690XC12004\\_2](https://doi.org/10.1207/S1532690XC12004_2)
- Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J.-F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative, comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 209–233). Erlbaum. <https://doi.org/10.4324/9781410603135>
- Britt, M. A., & Rouet, J.-F. (2012). Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (pp. 276–314). Cambridge University Press.
- Britt, M. A., Rouet, J.-F., & Braasch, J. L. G. (2013). Documents as entities: Extending the situation model theory of comprehension. In M. A. Britt, S. R. Goldman, & J.-F. Rouet (Eds.), *Reading: From words to multiple texts* (pp. 160–179). Routledge. <https://doi.org/10.4324/9780203131268>
- Britt, M. A., Rouet, J.-F., & Durik, A. (2018). *Literacy beyond text comprehension: A theory of purposeful reading*. Routledge. <https://doi.org/10.4324/9781315682860>
- \*Bromme, R., Scharrer, L., Stadler, M., Homberg, J., & Torspecken, R. (2015). Is it believable when it's scientific? How scientific discourse style influences laypeople's resolution of conflicts. *Journal of Research in Science Teaching*, 52(1), 36–57. <https://doi.org/10.1002/tea.21172>
- Chinn, C. A., Barzilai, S., & Duncan, R. G. (2021). Education for a “post-truth” world: New directions for research and practice. *Educational Researcher*, 50(1), 51–60. <https://doi.org/10.3102/0013189X20940683>
- Cho, B.-Y., & Afflerbach, P. (2017). An evolving perspective of constructively responsive reading comprehension strategies in multilayered digital text environments. In S. E. Israel (Ed.), *Handbook of research on reading comprehension* (2nd ed., pp. 109–14). Guilford.

- Cho, B.-Y., Afflerbach, P., & Han, H. (2018). Strategic processing in accessing, comprehending, and using multiple sources online. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use*. (pp. 133–150) Routledge. <https://doi.org/10.4324/9781315627496>
- Colmar, S., Liem, G. A. D., Connor, J., & Martin, A. J. (2019). Exploring the relationships between academic buoyancy, academic self-concept, and academic performance: A study of mathematics and reading among primary school students. *Educational Psychology*, 39(8), 1068–1089. <https://doi.org/10.1080/01443410.2019.1617409>
- \*Delgado, P., Stang-Lund, E., Salmerón, L., & Bråten, I. (2020). To click or not to click: Investigating conflict detection and sourcing in a multiple document hypertext environment. *Reading and Writing*, 33(8), 2049–2072. <https://doi.org/10.1007/s11145-020-10030-8>
- Duncan, R. G., Chinn, C. A., & Barzilai, S. (2018). Grasp of evidence: Problematizing and expanding the next generation science standards' conceptualization of evidence. *Journal of Research in Science Teaching*, 55(7), 907–937. <https://doi.org/10.1002/tea.21468>
- \*Flanagin, J. A., & Metzger, M. J. (2003). The perceived credibility of personal web page information as influenced by the sex of the source. *Computers in Human Behavior*, 19(6), 683–701. [https://doi.org/10.1016/S0747-5632\(03\)00021-9](https://doi.org/10.1016/S0747-5632(03)00021-9)
- \*Florit, E., Cain, K., & Mason, L. (2019). Going beyond children's single-text comprehension: The role of word reading, working memory, comprehension monitoring and source use in 4th graders' multiple-document comprehension. *British Journal of Educational Psychology*, 90(2), 449–472. <https://doi.org/10.1111/bjep.12288>
- Goldman, S. R., & Brand-Gruwel, S. (2018). Learning from multiple sources in a digital society. In F. Fisher, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International handbook of the learning sciences* (pp. 86–95). Routledge. <https://doi.org/10.4324/9781315617572>
- Goldman, S. R., & Scardamalia, M. (2013). Creating knowledge in the information age: Next-generation challenges and opportunities. *Cognition and Instruction*, 31(2), 255–269. <https://doi.org/10.1080/10824669.2013.773217>
- \*Gottlieb, E., & Wineburg, S. (2012). Between veritas and communitas: Epistemic switching in the reading of academic and sacred history. *Journal of the Learning Sciences*, 21(1), 84–129. <https://doi.org/10.1080/10508406.2011.582376>
- \*Hahnel, C., Kroehne, U., Goldhammer, F., Schoor, C., Mahlow, N., & Artelt, C. (2019). Validating process variables of sourcing in an assessment of multiple document comprehension. *British Journal of Educational Psychology*, 89(3), 524–537. <https://doi.org/10.1111/bjep.12278>
- \*Herrero-Diz, P., Conde-Jiménez, J., Tapia-Frade, A., & Varona-Aramburu, D. (2019). The credibility of online news: An evaluation of the information by university students. *Culture and Education*, 31(2), 407–435. <https://doi.org/10.1080/11356405.2019.1601937>
- Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review*, 13(3), 191–209. <https://doi.org/10.1023/A:1016667621114>
- Johnson, M. L., & Sinatra, G. M. (2014). The influence of approach and avoidance goals on conceptual change. *Journal of Educational Research*, 107(4), 312–225. <https://doi.org/10.1080/00220671.2013.807492>
- \*Kammerer, Y., Bråten, I., Gerjets, P., & Strømø, H. I. (2013). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during web search on a medical issue. *Computers in Human Behavior*, 29(3), 1193–1203. <https://doi.org/10.1016/j.chb.2012.10.012>
- \*Kammerer, Y., Gottschling, S., & Bråten, I. (2021). The role of Internet-specific justification beliefs in source evaluation and corroboration during web search on an unsettled socio-scientific issue. *Journal of Educational Computing Research*, 59(2), 342–378. <https://doi.org/10.1177/0735633120952731>
- \*Kammerer, Y., Kalbfell, E., & Gerjets, P. (2016). Is this information source commercially biased? How contradictions between web pages stimulate the consideration of source information. *Discourse Processes*, 53(5–6), 430–456. <https://doi.org/10.1080/0163853X.2016.1169968>
- \*Kammerer, Y., Meier, N., & Stahl, E. (2016). Fostering secondary-school students' intertext model formation when reading a set of websites: The effectiveness of source prompts. *Computers & Education*, 102, 52–64. <https://doi.org/10.1016/j.compedu.2016.07.001>
- \*Kang, H., Bae, K., Zhang, S. K., & Sundar, S. S. (2011). Source cues in online news: Is the proximate source more powerful than distal sources? *Journalism & Mass Communication Quarterly*, 88(4), 719–736. <https://doi.org/10.1177/107769901108800403>
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review*, 95(2), 163–182. <https://doi.org/10.1037/0033-295X.95.2.163>



- \*Kobayashi, K. (2014). Students' consideration of source information during the reading of multiple texts and its effect on intertextual conflict resolution. *Instructional Science*, 42(2), 183–205. <https://doi.org/10.1007/s11251-013-9276-3>
- Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, 15(3), 309–328. [https://doi.org/10.1016/S0885-2014\(00\)00030-7](https://doi.org/10.1016/S0885-2014(00)00030-7)
- \*List, A. (2014). *Modeling multiple source use: Using learner characteristics and source use behaviors to predict response quality* [Doctoral dissertation, University of Maryland]. Digital Repository at the University of Maryland. <http://hdl.handle.net/1903/16254>
- List, A., & Alexander, P. A. (2017). Cognitive Affective Engagement Model of multiple source use. *Educational Psychologist*, 52(3), 182–199. <https://doi.org/10.1080/00461520.2017.1329014>
- List, A., & Alexander, P. A. (2019). Toward an integrated framework of multiple text use. *Educational Psychologist*, 54(1), 20–39. <https://doi.org/10.1080/00461520.2018.1505514>
- \*List, A., Alexander, P. A., & Stephens, L. A. (2017). Trust but verify: Examining the association between students' sourcing behaviors and ratings of text trustworthiness. *Discourse Processes*, 54(2), 83–104. <https://doi.org/10.1080/0163853X.2016.1174654>
- \*List, A., Du, H., Wang, Y., & Lee, H. Y. (2019). Toward a typology of integration: Examining the documents model framework. *Contemporary Educational Psychology*, 58, 228–242. <https://doi.org/10.1016/j.cedpsych.2019.03.003>
- \*List, A., Stephens, L. A., & Alexander, P. A. (2019). Examining interest throughout multiple text use. *Reading and Writing*, 32(2), 307–333. <https://doi.org/10.1007/s11145-018-9863-4>
- \*Lucassen, T., Muilwijk, R., Noordzij, M. L., & Schraagen, J. M. (2013). Topic familiarity and information skills in online credibility evaluation. *Journal of the American Society for Information Science and Technology*, 64(2), 254–264. <https://doi.org/10.1002/asi.22743>
- \*Macedo-Rouet, M., Braasch, J. L. G., Britt, M. A., & Rouet, J.-F. (2013). Teaching fourth and fifth graders to evaluate information sources during text comprehension. *Cognition and Instruction*, 31(2), 204–226. <https://doi.org/10.1080/07370008.2013.769995>
- \*Macedo-Rouet, M., Salmerón, L., Ros, C., Pérez, A., Stadler, M., & Rouet, J.-F. (2020). Are frequent users of social network sites good information evaluators? An investigation of adolescents' sourcing abilities. *Journal for the Study of Education and Development*, 43(1), 101–138. <https://doi.org/10.1080/02103702.2019.1690849>
- Magliano, J. P., McCrudden, M. T., Rouet, J.-F., & Sabatini, J. (2018). The modern reader: Should changes to how we read affect theory and research? In M. F. Schober, D. N. Rapp, & M. A. Britt (Eds.), *Handbook of discourse processes* (2nd ed., pp. 343–361). Routledge. <https://doi.org/10.4324/9781315687384>
- Mahlow, N., Hähnel, C., Kroehne, U., Artelt, C., Goldhammer, F., & Schoor, C. (2020). More that (single) text comprehension? On university students' understanding of multiple documents. *Frontiers of Psychology*, 11, Article 562450. <https://doi.org/10.3389/fpsyg.2020.562450>
- \*Mason, L., Junyent, A. A., & Tornatora, M. C. (2014). Epistemic evaluation and comprehension of web-source information on controversial science-related topics: Effects of a short-term instructional intervention. *Computers & Education*, 76, 143–157. <https://doi.org/10.1016/j.compedu.2014.03.016>
- \*Mason, L., Scrimin, S., Tornatora, M. C., & Zaccoletti, S. (2017). Emotional reactivity and comprehension of multiple online texts. *Learning and Individual Differences*, 58, 10–21. <https://doi.org/10.1016/j.lindif.2017.07.002>
- \*Mason, L., Scrimin, S., Tornatora, M. C., Suitner, C., & Moé, A. (2018). Internet source evaluation: The role of implicit associations and psychophysiological self-regulation. *Computers & Education*, 119, 59–75. <https://doi.org/10.1016/j.compedu.2017.12.009>
- Mason, L., Scrimin, S., Zaccoletti, S., Tornatora, M. C., & Goetz, T. (2018). Webpage reading: Psychophysiological correlates of emotional arousal and regulation predict multiple-text comprehension. *Computers in Human Behavior*, 87, 317–326. <https://doi.org/10.1016/j.chb.2018.05.020>
- \*Mason, L., Zaccoletti, S., Scrimin, S., Tornatora, M. C., Florit, E., & Goetz, T. (2020). Reading with the eyes and under the skin: Comprehending conflicting digital texts. *Journal of Computer Assisted Learning*, 36(1), 89–101. <https://doi.org/10.1111/jcal.12399>
- McCarthy, K. S., & McNamara, D. S. (2021). The multidimensional knowledge in text comprehension framework. *Educational Psychologist*. Advance online publication. <https://doi.org/10.1080/00461520.2021.1872379>
- \*McCrudden, M. T., Stenseth, T., Bråten, I., & Strømsø, H. I. (2016). The effects of topic familiarity, author expertise, and content relevance on Norwegian students' document selection: A mixed

- methods study. *Journal of Educational Psychology*, 108(2), 147–162. <https://doi.org/10.1037/edu0000057>
- McGrew, S. (2021). Internet or archive? Expertise in searching for digital sources on a contentious historical question. *Cognition and Instruction*. Advance online publication. <https://doi.org/10.1080/0737008.2021.1908288>
- McNamara, D. S., & Magliano, J. P. (2009). Towards a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (vol. 51, pp. 297–384). Elsevier.
- \*Merk, M., & Huff, M. (2020). Does the position of source information for multiple documents matter? Insights from two experiments. *Contemporary Educational Psychology*, 62, Article e101900. <https://doi.org/10.1016/j.cedpsych.2020.101900>
- \*Merk, M., Werner, M., & Wagner, W. (2017). Historical thinking skills and mastery of multiple document tasks. *Learning and Individual Differences*, 54, 135–148. <https://doi.org/10.1016/j.lindif.2017.01.021>
- O’Keefe, P. A., & Harackiewicz, J. M. (Eds.) (2017). *The science of interest*. Springer. <https://doi.org/10.1007/978-3-319-55509-6>.
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan – a web and mobile app for systematic reviews. *Systematic Reviews*, 5, Article e210. <https://doi.org/10.1186/s13643-016-0384-4>
- Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, 19(3), 228–242. <https://doi.org/10.1016/j.learninstruc.2008.04.003>
- \*Paul, J., Cerdan, R., Rouet, J.-F., & Stadler, M. (2018). Exploring fourth graders’ sourcing skills. *Journal for the Study of Education and Development*, 41(3), 536–580. <https://doi.org/10.1080/02103702.2018.1480458>
- \*Paul, J., Stadler, M., & Bromme, R. (2019). Effects of a sourcing prompt and conflicts in reading materials on elementary students’ use of source information. *Discourse Processes*, 54(2), 155–169. <https://doi.org/10.1080/0163853X.2017.1402165>
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315–341. <https://doi.org/10.1007/s10648-006-9029-9>
- Perfetti, C. A., Britt, M. A., & Georgi, M. C. (1995). *Text-based learning and reasoning: Studies in history*. Erlbaum.
- Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Towards a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 99–122). Erlbaum. <https://doi.org/10.4324/9781410603050>
- \*Peterson, E. G., & Alexander, P. A. (2020). Navigating print and digital sources: Students’ selection, use, and integration of multiple sources across mediums. *The Journal of Experimental Education*, 88(1), 27–46. <https://doi.org/10.1080/00220973.2018.1496058>
- \*Potocki, A., de Pereyra, G., Ros, C., Macedo-Rouet, M., Stadler, M., Salmerón, L., & Rouet, J.-F. (2020). The development of source evaluation skills during adolescence: Exploring different levels of source processing and their relationships. *Journal for the Study of Education and Development*, 43(1), 19–59. <https://doi.org/10.1080/02103702.2019.1690848>
- Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist*, 46(3), 168–184. <https://doi.org/10.1080/00461520.2011.587723>
- 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>
- Richter, T., & Maier, J. (2018). The role of validation in multiple-document comprehension. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 151–167). Routledge. <https://doi.org/10.4324/9781315627496>
- Rouet, J.-F. (2006). *The skills of document use: From text comprehension to Web-based learning*. Erlbaum.
- Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19–52). Information Age.

- Rouet, J.-F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, 52(3), 200–215. <https://doi.org/10.1080/00461520.2017.1329015>
- Rouet, J.-F., Britt, M. A., Mason, R. A., & Perfetti, C. A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology*, 88(3), 478–493. <https://doi.org/10.1037/0022-0663.88.3.478>
- \*Rouet, J.-F., Favart, M., Britt, M. A., & Perfetti, C. A. (1997). Studying and using multiple documents in history: Effects of discipline expertise. *Cognition and Instruction*, 15(1), 85–106. [https://doi.org/10.1207/s1532690xci1501\\_3](https://doi.org/10.1207/s1532690xci1501_3)
- \*Salmerón, L., Gil, L., & Bråten, I. (2018). Effects of reading real versus print-out versions of multiple documents on students' sourcing and integrated understanding. *Contemporary Educational Psychology*, 52, 25–35. <https://doi.org/10.1016/j.cedpsych.2017.12.002>
- \*Salmerón, L., Kammerer, Y., & García-Carrión, P. (2013). Searching the Web for conflicting topics: Page and user factors. *Computers in Human Behavior*, 29(6), 2161–2171. <https://doi.org/10.1016/j.chb.2013.04.034>
- \*Salmerón, L., Macedo-Rouet, M., & Rouet, J.-F. (2016). Multiple viewpoints increase students' attention to source features in social question and answer forum messages. *Journal of the Association for Information Science and Technology*, 67(10), 2404–2419. <https://doi.org/10.1002/asi.23585>
- \*Salmerón, L., Sampietro, A., & Delgado, P. (2020). Using Internet videos to learn about controversies: Evaluation and integration of multiple and multimodal documents by primary school students. *Computers & Education*, 148, Article e103796. <https://doi.org/10.1016/j.compedu.2019.103796>
- Scharrer, L., & Salmerón, L. (2016). Sourcing in the reading process: Introduction to the special issue. *Reading and Writing*, 29(8), 1539–1548. <https://doi.org/10.1007/s11145-016-9676-2>
- Sewasew, D., & Koester, L. S. (2019). The developmental dynamics of students' reading self-concept and reading competence: Examining reciprocal relations and ethnic-background patterns. *Learning and Individual Differences*, 73, 102–111. <https://doi.org/10.1016/j.lindif.2019.05.010>
- Sinatra, G. M., & Lombardi, D. (2020). Evaluating sources of scientific evidence and claims in the post-truth era may require reappraising plausibility judgments. *Educational Psychologist*, 55(3), 120–131. <https://doi.org/10.1080/00461520.2020.1730181>
- Stadtler, M., & Bromme, R. (2014). The content-source integration model: A taxonomic description of how readers comprehend conflicting scientific information. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 379–402). The MIT Press. <https://doi.org/10.7551/mitpress/9737.001.0001>
- \*Stang Lund, E., Bråten, I., Brandmo, C., Brante, E. W., & Strømsø, H. I. (2019). Direct and indirect effects of textual and individual factors on source-content integration when reading about a socio-scientific issue. *Reading and Writing*, 32(2), 335–356. <https://doi.org/10.1007/s11145-018-9868-z>
- \*Stang Lund, E., Bråten, I., Brante, E. W., & Strømsø, H. I. (2017). Memory for textual conflicts predicts sourcing when adolescents read multiple expository texts. *Reading Psychology*, 38(4), 417–437. <https://doi.org/10.1080/02702711.2016.1278417>
- \*Stenseth, T., & Strømsø, H. I. (2019). To read or not to read: A qualitative study of students' justifications for document selection in task-oriented reading. *Scandinavian Journal of Educational Research*, 63(5), 771–788. <https://doi.org/10.1080/00313831.2018.1434828>
- Strømsø, H. I., & Bråten, I. (2013). Multiple documents literacy. In L. Meyer (Ed.), *Oxford bibliographies in education*. Oxford University Press. <https://doi.org/10.1093/OBO/9780199756810-0092>
- \*Strømsø, H. I., Bråten, I., & Brante, E. W. (2020). Profiles of warm engagement and cold evaluation in multiple-document comprehension. *Reading and Writing*, 33(9), 2337–2359. <https://doi.org/10.1007/s11145-020-10041-5>
- \*Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction*, 20(3), 192–204. <https://doi.org/10.1016/j.learninstruc.2009.02.001>
- \*Strømsø, H. I., Bråten, I., & Britt, M. A. (2011). Do students' beliefs about knowledge and knowing predict their judgement of texts' trustworthiness? *Educational Psychology*, 31(2), 177–206. <https://doi.org/10.1080/01443410.2010.538039>
- Taasoobshirazi, G., & Sinatra, G. M. (2011). A structural equation model of conceptual change in physics. *Journal of Research in Science Teaching*, 48(8), 901–918. <https://doi.org/10.1002/tea.20434>
- \*Tarchi, C. (2019). Identifying fake news through trustworthiness judgements of documents. *Culture and Education*, 31(2), 369–406. <https://doi.org/10.1080/11356405.2019.1597442>

- Trevors, G., Muis, K. R., Pekrun, R., Sinatra, G. M., & Muijselaar, M. M. L. (2017). Exploring the relations between epistemic beliefs, emotions, and learning from texts. *Contemporary Educational Psychology, 48*, 116–132. <https://doi.org/10.1016/j.cedpsych.2016.10.001>
- \*Ulyshen, T. Z., Koehler, M. J., & Gao, F. (2015). Understanding the connection between epistemic beliefs and Internet searching. *Journal of Educational Computing Research, 53*(3), 345–383. <https://doi.org/10.1177/0735633115599604>
- \*Van Der Heide, B., & Lim, Y.-S. (2016). On the conditional cueing of credibility heuristics: The case of online influence. *Communication Research, 43*(5), 672–693. <https://doi.org/10.1177/0093650214565915>
- \*van Strien, J. L., Kammerer, Y., Brand-Gruwel, S., & Boshuizen, H. P. (2016). How attitude strength biases information processing and evaluation on the web. *Computers in Human Behavior, 60*, 245–252. <https://doi.org/10.1016/j.chb.2016.02.057>
- \*von der Mühlen, S., Richter, T., Schmid, S., Schmidt, E. M., & Berthold, K. (2016). The use of source-related strategies in evaluating multiple psychology texts: A student-scientist comparison. *Reading and Writing, 29*(8), 1677–1698. <https://doi.org/10.1007/s11145-015-9601-0>
- \*Westerwick, A. (2013). Effects of sponsorship, web site design, and Google ranking on the credibility of online information. *Journal of Computer-Mediated Communication, 18*(2), 194–211. <https://doi.org/10.1111/jcc4.12006>
- \*Wiley, J., Griffin, T. D., Steffens, B., & Britt, M. A. (2020). Epistemic beliefs about the value of integrating information across multiple documents in history. *Learning and Instruction, 65*, Article e101266. <https://doi.org/10.1016/j.learninstruc.2019.101266>
- \*Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83*(1), 73–87. <https://doi.org/10.1037/0022-0663.83.1.73>
- \*Winter, S., & Krämer, N.C. (2012). Selecting science information in web 2.0: How source cues, message sidedness, and need for cognition influence users' exposure to blog posts. *Journal of Computer-Mediated Communication, 18*(1), 80–96. <https://doi.org/10.1111/j.1083-6101.2012.01596.x>
- Yang, F.-Y., Chang, C.-C., Chen, L.-L., & Chen, Y.-C. (2016). Exploring learners' beliefs about science reading and scientific epistemic beliefs, and their relations with science text understanding. *International Journal of Science Education, 38*(10), 1591–1606. <https://doi.org/10.1080/09500693.2016.1200763>
- Zaccoletti, S., Altoè, G., & Mason, L. (2020). Enjoyment, anxiety and boredom, and their control-value antecedents as predictors of reading comprehension. *Learning and Individual Differences, Article e101869*. <https://doi.org/10.1016/j.lindif.2020.101869>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.