



Profiles of warm engagement and cold evaluation in multiple-document comprehension

Helge I. Strømsø¹ · Ivar Bråten¹ · Eva W. Brante^{1,2}

Published online: 11 April 2020
© The Author(s) 2020

Abstract

We explored potential profiles of interest, attitudes, and source evaluation by performing cluster analysis in a sample of Norwegian upper-secondary students. Differences among the profile groups with regard to multiple-document use were examined. The profile groups were partly consistent with the default stances described by the cognitive-affective engagement model of multiple-source use (List & Alexander, 2017), resulting in critical analytic, evaluative, and disengaged profiles. However, the model's assumption that interest and attitude constitute one affective engagement dimension was not confirmed. There were no statistically significant differences between the profile groups in the processing of a set of multiple documents; yet there was a tendency for students who adopted a critical analytic stance to engage in a more thorough text selection process. Those students also included more information units from the selected texts in their written products and integrated information units across the texts more frequently compared to the other profile groups.

Keywords Multiple-text comprehension · Source evaluation · Topic interest · Attitudes · Cluster analysis

Introduction

Since Sam Wineburg's (1991) seminal paper on novices' and experts' reading of multiple documents in history was published three decades ago, students' reading of multiple documents has attracted increasing attention among reading researchers (e.g., Braasch, Bråten, & McCrudden, 2018). Whereas multiple-document literacy originally was studied as discipline-specific heuristics needed to judge and interpret different historical sources (Wineburg, 1991, 1994), such heuristics have later

✉ Helge I. Strømsø
h.i.stromso@iped.uio.no

¹ Department of Education, University of Oslo, P.O. Box 1092, Blindern, 0317 Oslo, Norway

² Present Address: Malmö University, Malmö, Sweden

been described as extending beyond disciplinary boundaries (Wineburg & Reisman, 2015) and representing competencies needed in the information anarchy that characterizes the twenty-first century (Afflerbach & Cho, 2009; Alexander & DRLRL, 2012). Such competencies concern how readers select and navigate among different information sources, evaluate both documents' content and origin (source), and integrate within and across documents (Salmerón, Strømsø, Kammerer, Stadler, & van den Broek, 2018).

The documents model framework (DMF) proposed by Perfetti, Rouet, and Britt (1999) has probably been the most influential framework for research of multiple-document comprehension so far. The framework suggests that when reading multiple documents, good readers construct an integrated mental model of the documents' content, as well as representations of the sources of those documents. A documents model is constructed when readers connect the integrated mental model representing the content and the mental representations of the sources. In brief, readers come to understand "who said what" and relationships that exist among different sources (e.g., whether they agree or disagree, whether they support or oppose each other). Somewhat later, the multiple-document task-based relevance assessment and content extraction model (MD-TRACE) was developed to describe in more detail how readers go about constructing a documents model (Rouet & Britt, 2011), and, recently, Britt, Rouet, and Durik (2018) proposed the RESOLV model as an extension of the MD-TRACE by including readers' interpretations of contextual aspects of the reading task in the framework.

Other frameworks relevant to multiple-document comprehension have also been suggested, such as the information-problem solving on the Internet (IPS-I) model by Brand-Gruwel and colleagues (e.g., Brand-Gruwel & van Strien, 2018) and the new literacies of online research and comprehension framework of Leu and colleagues (e.g., Leu, Kinzer, Coiro, Castek, & Henry, 2013). Building on previous models and frameworks, List and Alexander (2019) suggested the integrated framework of multiple texts (IF-MT) as a metaframework describing what to consider when investigating students' multiple-text use. The IF-MT comprises three different stages of multiple-document comprehension: preparation, execution, and production. An essential aspect of the preparation stage was previously described by the cognitive affective engagement model (CAEM) of multiple-source use (List & Alexander, 2017; 2018). This model was developed in response to the primarily cognitive models typifying earlier phases of research on multiple-document comprehension and, as such, the CAEM stressed the importance of affective components. Thus, although prior models to some extent have acknowledged the importance of individual difference variables for readers' interpretation of multiple documents tasks (e.g., Rouet & Britt, 2011), the role of affective variables such as interest and attitudes have previously not been integrated into these models. In contrast, such affective components are at the heart of the CAEM together with critical reading behavior. Basically, List and Alexander (2017, 2018) argued that the often-challenging nature of multiple-document tasks, such as the need to construct coherence from partly contradictory texts, requires effort and engagement on part of the readers. To our knowledge, the CAEM model has not yet been empirically tested, however, and the goal of the present study was to explore whether the student profiles of interacting cognitive and

affective factors suggested by the CAEM would emerge in a sample of Norwegian upper-secondary students, as well as whether those profiles might be related to processes of selecting, reading, and writing in a multiple-document context. In the following, we briefly describe the overarching IF-MT framework before we discuss the CAEM model and present the questions and expectations guiding the current research.

The integrated framework of multiple texts (IF-MT)

The IF-MT builds on the DMF (Perfetti et al., 1999) and the MD-TRACE (Rouet & Britt, 2011), but List and Alexander (2019) argued that those models had not paid sufficient attention to extant empirical work highlighting the importance of individual factors, such as attitudes, interest, or epistemic beliefs. Additionally, List and Alexander (2019) maintained that prior models in this area had not taken work on learners' strategic processing or argument construction sufficiently into consideration. In this section, we focus on the three main stages of multiple text use described by the IF-MT.

In the first stage, labeled *preparation*, readers form a preparatory stance toward the reading activity by conceptualizing features of the reading task. Perceptions of external features of the task, such as domain or topic, and explicit standards for task completion, are presumably influenced by numerous individual differences, such as prior knowledge, interest, attitudes, and text-processing abilities. List and Alexander (2019) suggested that affective engagement and behavioral dispositions together constitute different default stances in this stage, with those stances representing the CAEM. Thus, the CAEM is an essential part of the preparation stage of the IF-MT. According to the IF-MT, the resulting preparatory stance will influence readers' further processing of multiple texts.

The *execution* stage comprises readers' strategic interaction with the texts. Good readers' processing of multiple texts is assumed to be characterized by the use of behavioral, cognitive, and metacognitive strategies. Behavioral strategies are the observable actions taking place, such as text selection, navigation, and note taking or annotation, while cognitive strategies refer to mental processes related to comprehension. List and Alexander (2019) also differentiated between cognitive processes involved in single-text (intratextual) comprehension and cognitive processes involved in multiple-text (intertextual) comprehension. Finally, metacognitive strategies represent readers' ability to monitor text comprehension, the validity of the content, and potential achievements in light of their reading goal.

The last stage, *production*, concerns cognitive and affective outcomes of readers' use of multiple texts. Readers' knowledge gain from reading represents the cognitive outcome, while changes in topic interest and attitudes may be considered affective outcomes. When reading multiple texts, the optimal cognitive outcome is a documents model in which essential units of information are linked across texts and tagged for their sources (Perfetti et al., 1999; Rouet, 2006).

Whereas the IF-MT is described as a comprehensive framework of multiple-text use, List and Alexander (2019) included the CAEM as an essential part of the

preparation stage in that framework. Our goal is to examine whether the default stances described in the CAEM (List & Alexander, 2017, 2018) appear in a sample of upper-secondary students and the extent to which those stances relate to aspects of the execution and production stages of the IF-MT.

The cognitive affective engagement model (CAEM)

As noted previously, the description of the preparation stage within the IF-MT was based on the CAEM (List & Alexander, 2017, 2018), with the default stance adopted at this stage expected to influence the reading processes and outcomes in the next two stages (List & Alexander, 2018, 2019). The CAEM takes readers' interpretation of the reading task as its point of departure, with this process drawing on characteristics of the specific task as well as on individual differences. Two features relevant to readers' perception of the task are the topic of reading and the expected standards for the reading outcomes. Based on readers' preliminary representation of the task and individual difference factors, which potentially include prior knowledge, interest, epistemic beliefs, attitudes, and processing abilities, they form a default stance toward the multiple-text task (List & Alexander, 2019). According to List and Alexander (2017), a default stance constitutes a motivational and cognitive orientation toward the multiple-text task. Although default stances are considered to be formed during the preparation stage, they might be modified during the next stages as the reading task unfolds.

Four different default stance profiles are described by the CAEM, with two dimensions defining those stances. One dimension is labeled *affective engagement*, capturing readers' motivational orientation toward the multiple-text task. List and Alexander (2017, 2018) assumed that readers' interest in and attitudes toward the topic are the main components of the affective engagement dimension. The second dimension is labeled *behavioral dispositions* and refers to readers' habituated practices concerning source evaluation and information verification. These two dimensions are considered to form the basis of four default stances that students might adopt when working with multiple information sources. The default stances are supposed to reflect the degree to which readers affectively engage in a particular multiple-text task and their habits regarding source evaluation.

According to the CAEM, readers adopting a *disengaged default stance* will be low in affective engagement as well as source evaluation skills. When facing a multiple-text task, disengaged readers will typically not be interested in the topic or the task; nor are they likely to hold strong attitudes toward the topic. Further, readers categorized as disengaged will not have developed appropriate source evaluation skills. Another group of readers might be highly engaged in the topic and the reading task but lack the habit of critically evaluating the texts that they read. Accordingly, they may be motivated to accumulate extensive information regardless of the quality of the sources. Such readers are said to be adopting an *affectively engaged default stance*. Readers who typically use appropriate source evaluation skills but are not engaged in the current topic or reading task, are said to be adopting an *evaluative default stance*. Even though those readers

are not expected to display much interest in or hold strong attitudes toward the topic or the task, they will routinely perform activities relevant to critical evaluation of the information sources. Finally, the fourth group of readers proposed by the CAEM is characterized by a *critical analytic default stance*. These readers are both affectively engaged in the topic and the reading task and in the habit of performing relevant evaluation activities in a multiple-text context. Comprehension of multiple texts on a specific issue is often a demanding task because readers must construct a mental representation across texts and, thus, take on the task performed by authors of single texts in prioritizing information units and constructing a coherent story from different information sources. To succeed in such a task, both engagement and critical evaluation presumably are needed.

As noted previously, the development of the CAEM was a response to the primarily cognitive emphasis of previous models and sub-models of multiple-document comprehension (e.g., Braasch & Bråten, 2017; Brand-Gruwel & van Strien, 2018; Perfetti et al., 1999; Rouet & Britt, 2011). As such, the CAEM can be considered a valuable and potentially fruitful contribution to the field, consistent with a number of studies indicating that individual differences in the affective domain related to interest, attitudes, and emotions are associated with multiple-document comprehension (e.g., Bråten, Anmarkrud, Brandmo, & Strømsø, 2014; Kobayashi, 2014; Mason, Scrimin, Tornatora, & Zoccoletti, 2017; Richter & Maier, 2017; Strømsø & Bråten, 2009; Trevors, Muis, Pekrun, Sinatra, & Muijselaar, 2017; van Strien, Brand-Gruwel, & Boshuizen, 2014). List and Alexander (2017, 2018) also cited some prior studies that seemed to provide empirical support for the four different CAEM profiles (Kiili, Laurinen, & Marttunen, 2008; Lawless & Kulikowich, 1996). Of note is, however, that the profiles resulting from cluster analysis in those studies were based on behavioral data. Therefore, they do not necessarily reflect the beliefs and dispositions involved in the preparation stage of multiple-text use. There is thus a need for further investigation of the distinct profiles proposed by the CAEM.

The affective engagement dimension described by the CAEM focuses on attitudes and interest, while the behavioral disposition dimension focuses on habits with regard to evaluation of source information. However, List and Alexander (2019) acknowledged that a number of additional individual difference factors may be involved in the formation of readers' default stances during the preparation stage. Specifically, they highlighted the potential importance of epistemic beliefs and prior knowledge. A relationship between readers' epistemic beliefs and multiple-text comprehension has been demonstrated in a number of studies (e.g., Bråten, Britt, Strømsø, & Rouet, 2011; Bråten, Ferguson, Strømsø, & Anmarkrud, 2013; Barzilai and Eshet-Alkalai, 2015; Wiley, Griffin, Steffens, & Britt, 2020). Likewise, prior knowledge conceptualized as academic (Bulger, Mayer, & Metzger, 2014), disciplinary (Rouet, Favart, Britt, & Perfetti, 1997; Wineburg, 1991), or topic specific knowledge (Bråten et al., 2014; Strømsø, Bråten, & Samuelstuen, 2008) has been demonstrated to predict multiple-text comprehension. List and Alexander (2019) speculated somewhat on how these two individual differences, in particular, might relate to the default stances described by the CAEM. In the current study, we also included a measure of participants' topic knowledge to explore relationships

between this individual difference and the default stances, as well as between topic knowledge and processes and products of a multiple-text task.

The present study

Based on the theoretical description of the CAEM, our first goal was to examine whether the four default stances proposed by List and Alexander (2017, 2018, 2019) would emerge in a sample of Norwegian upper-secondary students presented with a multiple-text task on the topic of nuclear power. We chose nuclear power as the topic because it has been a recurrent topic in Norwegian media since the Chernobyl disaster in 1986, with the pasture for Norwegian reindeer and sheep still being contaminated to some degree. The topic is also included in the national curriculum for upper secondary school in Norway. Thus, we expected participants to have some basic knowledge about the topic and considered it likely that they would perceive it as relevant to a Norwegian context.

Given the prominent position of prior knowledge in models of reading and reading research (e.g., McNamara & Magliano, 2009; O'Reilly, Wang, & Sabatini, 2019), we also wanted to explore potential differences in participants' prior topic knowledge across default stances. Considering the theoretical justification for the CAEM, we assumed that the different profiles described by List and Alexander could be identified by performing hierarchical cluster analysis on data collected using measures of interest, attitudes, and source evaluation skills.

Regarding the relationship between prior knowledge and the affective engagement dimension of the CAEM, prior knowledge can be assumed to be associated with topic interest (Hidi & Renninger, 2006), with these variables found to be moderately correlated in some multiple-text studies (Bråten et al., 2014; Strømsø & Bråten, 2009). Although a relationship between prior knowledge and attitudes seems less obvious, a modest correlations have been found across studies (Allum, Sturgis, Tabourazi, & Brunton-Smith, 2008; Lewandowsky & Oberauer, 2016; Strømsø & Bråten, 2017). Further, prior research has demonstrated relationships between topic knowledge (Bråten et al., 2011) and disciplinary expertise (von der Mühlen, Richter, Schmid, Schmidt, & Berthold, 2016; Wineburg, 1991) on the one hand, and source evaluation on the other. In brief, then, prior knowledge could be expected to be positively related to the affective engagement as well as the behavioral dispositions dimension of the CAEM.

Our next goal was to examine potential relationships between the emerging default stances and processing variables related to multiple-text reading, focusing on processes related to text selection and the time devoted to processing the selected texts. Of note is that while the default stances featured in the CAEM represent the core of the preparation stage within the IF-MT, the processing variables mentioned above are central to the execution stage of the IF-MT (List & Alexander, 2019). Hence, our second goal was to examine whether individual differences integral to the preparation stage would matter in terms of processing taking place within the execution stage. Following List and Alexander (2017, 2018), we expected default stances groups to differ with respect to text selection

in that readers displaying a disengaged default stance would spend less time on text selection and select fewer texts than participants displaying other default stances, and in that participants displaying an affective engagement stance would select the highest number of texts. We also considered it likely that participants displaying evaluative and critical analytic default stances, respectively, would spend more time considering the different texts during text selection than participants in the other groups because they could be assumed to be more selective in terms of the quality of information sources. As part of the text selection process, participants also were asked to justify their text selections, and their responses were coded into content- and source-feature-based justifications, respectively. We expected all default stances groups to refer to content in their justifications at an approximately equal rate, whereas participants displaying evaluative and critical analytic default stances could be expected to justify their text selections by referring to source features at a higher rate than participants in the other groups. That is, participants assumed to habitually engage in source evaluation would probably also refer more often to source features, such as author expertise and publication venue, in justifying why they selected particular texts.

Reading time has been found to be associated with students' multiple-text comprehension (Bråten, Brante, & Strømsø, 2018a; Bråten et al., 2014), and it has also been considered to reflect an important aspect of readers' behavioral engagement in the reading task (Guthrie & Klaua, 2014). In the present study, participants' total reading time for the selected texts was measured, and we expected participants displaying engaged and critical analytic stances to spend more time reading the texts than other participants because the former stances, by definition, are characterized by higher levels of engagement.

Finally, we wanted to examine potential relationships between the emerging default stances and products of multiple-text use, focusing on participants' written products in terms of the number of words, the number of information units from the selected texts, the integration of information units across texts, and the number of source feature references. Again, the variables involved mirror central aspects of stages described within the IF-MT (List & Alexander, 2019), with default stances representing the preparation stage (as elaborated by the CAEM) and written products representing important aspects of the production stage within the IF-MT. Regarding the number of words, we expected that participants displaying stances defined by higher levels of engagement would produce more text than the other participants. Regarding information units, we, in accordance with List and Alexander (2017), expected that participants profiled as engaged or critical analytic would include more content from the texts than the other participants. Regarding integration, we, again following List and Alexander (2017), expected that participants displaying a critical analytic stance would outperform other participants in terms of the integration of information across texts. Lastly, regarding sourcing in the written products, we expected only readers categorized as critical analytic to engage in keeping track of "who said what" when writing up the assignment.

Method

Participants

Participants were 66 students (M age = 16.2, SD = .68; 49.3% female) at an upper-secondary school in southeast Norway who attended college preparatory courses. We recruited students randomly from six different classes with between 9 and 12 students from each class. A majority of the participants (76%) were native-born Norwegians, whereas the others came from families where parents did not speak Norwegian as their first language. The sample was relatively homogenous (i.e., middle class) with regard to socioeconomic status.

Materials

Topic knowledge measure

Knowledge about the topic of nuclear power was assessed by means of a 12-item multiple-choice test. The measure assessed prior knowledge of scientific (e.g., nuclear fission) and political (e.g., the International Atomic Energy Agency) aspects of nuclear power (sample items in Appendix B). The participants' scores were the number of correct responses. Cronbach's α was .68. This measure has been used and validated in prior research, reporting a test–retest reliability of .72 (McCrudden, Stenseth, Bråten, & Strømsø, 2016).

Topic interest measure

A 12-item inventory with a 10-point anchored scale (1 = not at all true of me, 10 = very true of me) was used to measure participants' interest in the topic of nuclear power. Six of the items assessed interest in the topic without targeting any active engagement or involvement, while the other six items focused more on how engaged and involved the participants reportedly were in the topic (sample items in Appendix A). Cronbach's α = .92.

Attitudes

Participants' attitudes toward the topic were measured with an inventory asking them to rate the extent to which they identified with four statements concerning nuclear power (e.g., *I believe nuclear power plants represent environmental risks*) on a 10-point anchored scale (1 = not at all true of me, 10 = very true of me). High scores on this measure indicated that participants held a negative attitude toward nuclear power (i.e., judged nuclear power to be high-risk) and low scores indicated that they considered nuclear power to be a safe form of energy. Cronbach's α = .76.

Source evaluation skills

To assess participants' general knowledge about sources and their ability to use and evaluate source feature information, we administered a Norwegian adaptation of the Source Knowledge Inventory (Rouet, Ros, de Pereyra, Macedo-Rouet, & Salmerón, 2013). This measure consisted of seven tasks. On the first five tasks, participants were presented with brief texts on different natural and social science topics (e.g., nutrition or demography) and asked to rate the sources of each text with respect to expertise and potential bias, using a scale ranging from 1 to 10. Higher scores indicated more general source evaluation skills, with scores, for example, reflecting the extent to which participants considered a pharmacist to have high expertise on medication, and the extent to which they took into account that a pharmacist employed in a big pharmaceutical company might be biased. On the two final tasks, participants were presented with two fictitious search engine results pages (SERPs), each displaying four results on the current topic (biodiversity or freshwater on Earth). The four results on each SERP indicated the source of the website they were representing in several ways, such as URL, title, and key words. For each SERP, the participants were asked to rate each result with respect to whether they wanted to use information from that website in preparing a presentation on the topic, using a scale ranging from 1 to 10 to evaluate the usefulness of each site. Again, higher scores indicated more general source evaluation skills, for example reflecting the extent to which participants considered information about biodiversity on a website provided by a public educational resource to be useful to complete the task, and the extent to which they realized that information from a commercial website promoting a certain product for gardening might be useless in this regard. Cronbach's $\alpha = .70$.

Texts, computer application, and processing and product measures

Participants were presented with a list of 10 texts about the use of nuclear power. In each text, source information (author, credentials, affiliation, text type, venue, and date) was displayed on the first two lines, followed by three sentences of content information. The sources ranged from blog posts written by secondary-school students to textbooks written by high-school teachers and journal articles written by science professors. The three-sentence content information was always relevant and consisted of neutral, factual information as well as information considered controversial.

Participants accessed these 10 texts through a web-based application program, in which they first selected the items they wanted to use in writing a letter to the editor regarding the topic. On a page displaying only the selected texts, they then justified in writing why they had selected each of these texts. Next, they obtained access to a third page containing expanded versions of the selected texts. That is, by clicking on a text, they gained access to an expanded text of approximately 100 words in addition to the source information, and by clicking on another text, that text was expanded and the previous one was again reduced to three-sentence length. Participants could go back and forth between a page where they were writing their letter and the page on which their selected texts were located. After finishing their letters,

participants submitted them to a server. The application program was logging the time participants used for the initial selection task and the total time they used for processing the expanded texts.

The texts provided to participants both described challenges that nuclear power plants might represent and new developments that might make them safer. For example, potential consequences of nuclear accidents were described in a newspaper article on the devastating incidents in Chernobyl and Fukushima, and a scientific journal article written by a professor described how earthquakes can damage nuclear power plants. The problem of radioactive waste from nuclear power plants was described in several texts authored by both students and experts, whereas other texts, written by a teacher and by several experts, described how new technology and international agreements can make the use of nuclear power as a source of energy much safer. Given the fatal consequences of nuclear accidents, we assumed that students would be interested in new developments concerning safety.

Data on different aspects of text selection, reading, and writing were collected. As indicators of the process of *text selection*, we used the time devoted to the initial selection task, the number of texts selected, and participants' justifications for selecting these particular texts. Following Braasch, Bråten, Strømsø, Anmarkrud, and Ferguson (2013), we coded the justifications into content-based and source-feature-based justifications, respectively. Two independent raters scored a random selection of 20% of the justifications, resulting in 92% agreement on the type of justification provided for text selection. The *total reading* time for the expanded texts was used as a measure of how intensely and thoroughly participants processed the selected texts. Of note is that the time devoted to reading has been considered an important indicator of engagement with the texts within reading motivation (Guthrie & Klauda, 2014, 2016), and as an indicator of effortful processing, it has uniquely predicted performance on multiple-text reading tasks among upper-secondary and undergraduate students when other motivational and cognitive variables, including basic reading skills (i.e., word recognition) have been controlled for (e.g., Bråten et al., 2014, 2018a; List, Stephens, & Alexander, 2019).

We used four different *writing measures* as indicators of the products of multiple-text use. The number of words in participants' written products (letters to the editor) was counted, as well as the numbers of information units from the texts they included, switches between information units from different texts, and references to source features. The number of information units in the written products indicated the degree of content coverage. When a sentence or part of a sentence in the written product contained information that corresponded to information contained in a particular part of one of the selected texts, it was coded as an information unit coming from that text. The number of switches between information units from different texts indicated the degree of content integration. This way of measuring content integration across multiple texts was developed by Britt and Sommer (2004) and has been validated in a number of more recent studies (e.g., Bråten, Brante, & Strømsø, 2019; Bråten et al., 2018a; Gil, Bråten, Vidal-Abarca, & Strømsø, 2010). For example, if a written product contained seven information units altogether, and the first four information units came from one text and the next three information units came from another text, this would count as one switch and indicate poor content

integration. Finally, the number of references to source features (i.e., author, author credentials, author affiliation, text type, venue, and date) in the written products indicated the extent to which accurate source information was linked to information units from the texts. Two raters independently scored a random selection of 20% of the written products, resulting in 92% agreement on the texts from which the information units came. Independent scoring of a random selection of 20% of the written products for the number of source-feature references yielded an interrater reliability coefficient (Pearson's r) of .99.

Procedure

We collected data in two sessions separated by 8 weeks. The first session was a 45-min class period in which all participants completed a demographics survey and the pre-reading individual difference measures on paper.

The second session was a 60-min class period in which participants used the application program to perform the selection, justification, reading, and writing activities described in the previous section. Before logging on with their laptops to access the application, they received a brief introduction providing some factual background information and mentioning a controversy concerning the topic (i.e., the issue of the safety of nuclear power plants). After this introduction, the task instruction read: *You will be writing a letter to the editor where you discuss the safety of nuclear power plants. When you log on, you will see a list referring to 10 web texts. From this list, you are going to select the web texts you want to use when writing the letter to the editor.* On the first page of the application, the 10 texts were listed in random order for each participant.

Following Bråten, McCrudden, Stang Lund, Brante, and Strømsø (2018b), we used “writing a letter to the editor” in this task instruction because it seemed suitable for eliciting argumentative reasoning by the students although no direct argument prompt, which may be difficult to understand for many students (Britt, Richter, & Rouet, 2014), was given. Moreover, according to the teachers, students could be considered familiar with this literary genre.

Results

Cluster analysis

We performed hierarchical cluster analysis using the Ward method (Everitt, Landau, Leese, & Stahl, 2011; Yim & Ramdeen, 2015) to profile participants based on their topic interest, attitudes, and source evaluation skills. Of note is that this method is well suited to the current sample size (Kulikowich & Sedransk, 2012). Inspection of the dendrogram indicated a three-cluster solution. The three clusters resembled three of the four default stances suggested by List and Alexander (2017). Cluster 1 constituted a *critical analytic* group ($n=31$) with high scores on source evaluation skills and attitudes and moderate scores on topic interest. Cluster 2 was labeled the

evaluative group ($n=20$), based on its high scores on source evaluation skills, moderate attitudes, and low topic interest. Finally, participants in Cluster 3, which we labeled the *disengaged* group ($n=15$), had moderate source evaluation skills, strong negative attitudes and low topic interest (see Table 1).

A multivariate analysis of variance (MANOVA) was performed with cluster group as the independent variable and topic interest, attitudes, and source evaluation skills as the dependent variables. Results indicated a statistically significant overall difference between clusters, Wilk's $\lambda = .14$, $F(6, 122) = 33.88$, $p < .001$, $\eta^2 = .63$. Follow-up analyses of variance (ANOVAs) showed statistically significant univariate effects for all the dependent measures, $F_s(2, 63) > 29.17$, $ps < .001$. The effect sizes (partial η^2) were .48 (source evaluation skills), .41 (attitudes), and .43 (topic interest). A series of multiple comparisons with Fisher's least significant differences (LSD) showed that participants in the critical analytic and evaluative groups had statistically significantly higher scores on the source evaluation measure than participants in the disengaged group, and that participants in the critical analytic and disengaged groups had statistically significantly higher scores on the attitude measure than participants in the evaluative group. Regarding the topic interest measure, participants in the critical analytic group scored statistically significantly higher than participants in the two other clusters.

Discriminant function analysis showed that overall group membership was accurately predicted for 92.4% of the cases. Prediction accuracy for the three clusters was 90.3% for the critical analytic, 90.0% for the evaluative, and 100% for the disengaged group.

Prior knowledge

We conducted a one-way between-subjects ANOVA to compare the levels of prior knowledge between the three clusters. Although participants in the critical analytic group ($M=6.94$, $SD=2.85$) and the evaluative group ($M=6.65$, $SD=2.68$) had somewhat higher scores on the prior knowledge measure than participants in the disengaged group ($M=5.60$, $SD=2.20$), the ANOVA showed no statistically significant overall difference between the clusters, $F(2, 63) = 1.29$, ns , $\eta^2 = .04$. Still, estimations of Cohen's d showed medium effect sizes for the difference between the

Table 1 Mean scores on pre-reading measures for the three profiles

Variable	Profile		
	Cluster 1 ($n=31$) Critical analytic M (SD)	Cluster 2 ($n=20$) Evaluative M (SD)	Cluster 3 ($n=15$) Disengaged M (SD)
Interest	4.86 (1.43)	2.67 (1.45)	2.36 (1.11)
Attitude	8.02 (1.31)	5.64 (1.40)	8.33 (1.58)
Source evaluation	8.08 (.82)	8.19 (.50)	6.65 (.44)

disengaged and critical analytic groups ($d=0.52$) and the difference between the disengaged and the evaluative groups ($d=0.44$) with respect to prior knowledge.

Further, zero-order correlations among prior knowledge, topic interest, attitudes, and source evaluation skills were computed. This analysis showed a modest but statistically significant correlation between prior knowledge and source evaluation skills, $r = .24$, $p < .05$. No other statistically significant correlations were observed.

Finally, zero-order correlations were computed to explore whether prior knowledge correlated with processing measures related to text selection and reading or with product measures related to the writing task. No statistically significant correlations were found, with $r_s < .14$, $p_s > .31$. Accordingly, prior knowledge was not included in further analyses of relationships between the cluster groups and the multiple-text processing and product measures.

Text selection, reading, and writing

To compare the cluster groups with regard to the text selection, reading, and writing measures, a one-way between-subjects ANOVA was performed. There were no statistically significant differences on any of the text selection measures or on total reading time for the selected texts (see Table 2). However, we noted that participants in the critical analytic cluster had higher mean scores than participants in the other two clusters on all the text selection measures and on total reading time for the selected texts.

Regarding text selection time, both the critical analytic ($M = 185.39$, $SD = 79.39$) and the disengaged group ($M = 184.13$, $SD = 86.69$) devoted more time to the initial selection task than did the evaluative group ($M = 165.70$, $SD = 83.96$), but the effect sizes were small (respectively $d = .25$ and $d = .22$). The difference in number

Table 2 Mean scores on dependent variables for the three profiles

Variable	Profile		
	Cluster 1 ($n = 31$) Critical analytic M (SD)	Cluster 2 ($n = 20$) Evaluative M (SD)	Cluster 3 ($n = 15$) Disengaged M (SD)
Selecting texts			
Text selection time	185.39 (79.39)	165.70 (83.96)	184.13 (86.69)
Number of texts selected	4.81 (2.01)	4.35 (1.79)	3.93 (1.53)
Content based justifications	3.23 (2.62)	2.15 (2.06)	3.20 (1.74)
Source feature based justifications	4.26 (5.07)	3.40 (4.27)	3.27 (4.10)
Reading			
Total reading time	285.63 (189.25)	229.85 (154.33)	263.33 (193.10)
Writing			
Number of words	133.67 (81.58)	96.37 (59.28)	120.71 (81.24)
Number of information units	8.25 (4.72)	4.95 (4.66)	5.77 (4.21)
Number of switches	2.21 (1.82)	1.32 (1.38)	1.08 (1.12)
Source feature references	.76 (1.83)	2.05 (3.72)	.23 (.83)

of selected texts between the disengaged ($M=3.93$, $SD=1.53$) and critical analytic ($M=4.81$, $SD=2.01$) groups was medium large ($d=0.48$). Regarding the number of content-based justifications, there were also medium effect sizes for the difference between the critical analytic ($M=3.23$, $SD=2.62$) and the evaluative ($M=2.15$, $SD=2.06$) groups ($d=.46$) and the difference between the evaluative and the disengaged ($M=3.20$, $SD=1.74$) groups ($d=.56$). Thus, participants adopting an evaluative stance produced fewer justifications for their text selections that referred to the texts' content than did the two other groups. As for source-feature-based justifications, the critical analytic group had the highest mean number of justifications (see Table 2). There were, however, no statistically significant differences between the groups and the effect sizes were rather small. Still, the actual justifications for text selections may suggest some interesting tendencies across the default stances.

Thus, students adopting a critical analytic stance typically seemed to rely on both content and source features in selecting texts, for example referring to relevance as well as trustworthiness: "I chose this text because it concerns how long radioactive waste will remain dangerous, which is relevant when writing a letter to the editor. The article is written by another professor in natural sciences, and when facts are validated by several professors [referring to another text], the trustworthiness of my text will be stronger." In comparison, students in the evaluative group tended to rely more on source features than on content, as in the following example: "Because it [the selected text] is from a Norwegian journal on nuclear physics, authored by a professor at the Department of natural sciences. That is a trustworthy source." Finally, there were several examples that students in the disengaged group paid little attention to criteria for selecting sources, or that they produced justifications that were superficial. For example, one student in this group referred to the author's name, but not to his credentials or affiliation: "I chose this text because Jan Karlsen describes future nuclear power plants. I may end my letter to the editor with this such that I do not frighten my readers too much." These examples may illustrate potential differences between the groups that are in accordance with the CAEM.

For the writing measures, there were statistically significant differences for the number of information units from the texts included in participants' written products [$F(2, 57)=3.24$, $p=.05$, $\eta^2=.10$] and for the number of switches between information units from different texts [$F(2, 58)=3.13$, $p=.05$, $\eta^2=.10$]. Post hoc comparisons showed that the critical analytic group ($M=8.25$, $SD=4.21$) included more information units than the disengaged group ($M=5.76$, $SD=4.21$), although not statistically significantly more ($p=.11$, $d=0.57$). However, the critical analytic group included statistically significantly more information units ($p=.02$, $d=0.72$) than the evaluative group ($M=4.95$, $SD=4.66$). Regarding the number of switches, the written products of the critical analytic group ($M=2.21$, $SD=1.82$) had statistically significantly ($p=.03$, $d=0.71$) more switches than those of the disengaged group ($M=1.08$, $SD=1.12$), whereas the difference between the critical analytic group and the evaluative group ($M=1.32$, $SD=1.38$) did not quite reach a conventional level of statistical significance ($p=.06$, $d=0.55$). Although no statistically significant differences between the groups occurred for the number of words, we noted a medium effect size ($d=.52$) for the difference between the critical analytic group ($M=133.67$, $SD=81.58$) and the evaluative group ($M=96.37$, $SD=59.28$).

Given the skewed distribution of the number of source feature references in the written products, a Kruskal–Wallis test was performed to assess differences between the groups. The test showed no statistically significant differences among the three groups. One should note, however, that only 12 participants included references to source features in their written products.

Discussion

The results partly supported the theoretical description of the CAEM, as well as some of the suggested relationships between the CAEM and processes involved in multiple text reading and products resulting from such reading (List & Alexander, 2017, 2018). Our first research question concerned whether the four default stances represented in the CAEM would emerge in a sample of upper-secondary students. The cluster analysis resulted in only three clusters, however, with no evidence of an affective engagement cluster. Because the three clusters to a certain extent reflected default stances profiles suggested by the CAEM, we decided to retain the labels used in the original model for those profiles: disengaged, evaluative, and critical analytic.

Participants displaying a disengaged stance scored low on interest, high on attitude, and medium on source evaluation skills. We believe the difference between interest and attitude scores for this profile illustrates a challenge in highlighting interest and attitudes as reflecting affective engagement as a unified dimension. Strong attitudes may represent engagement either in favor of or against a certain issue (Ajzen, 1989), whereas high interest will only represent positive engagement. Thus, the two variables do not necessarily correlate (Stenseth, Bråten, & Strømsø, 2016), which also seems to be the case in our sample. Specifically, in the disengaged group, participants showed low interest in the topic of nuclear power and, at the same time, strong attitudes against the use of nuclear power. This might also be a measurement issue, however. Attitudes have traditionally been described as comprising both an affective and a cognitive component (Ajzen, 1989). The attitude measure that we used focused on beliefs and to a lesser extent on expressions of emotions. Thus, participants might have held strong negative beliefs about the use of nuclear power without necessarily being emotionally engaged in the issue.

Participants in the evaluative group scored low on interest and medium on attitudes and had relatively high scores on the source evaluation measure. This profile seemed to fit the evaluative default stance suggested by the CAEM quite well. The participants in this group did not express high engagement in the topic; yet, they were proficient in dealing with multiple texts in terms of source evaluation. Finally, participants in the critical analytic group scored medium on interest, high on attitude, and high on source evaluation. This profile represented the largest group of students in the sample, and the cluster fit the critical analytic default stance described by the CAEM fairly well.

Theoretical models, such as the CAEM, will seldom be perfectly reflected in smaller samples. Nevertheless, our results were fairly consistent with at least three of the profiles of the CAEM. The fourth profile, an affective engagement default stance, did not show up in our results. That profile supposedly reflects high scores

on both the interest and the attitude measures in combination with low source evaluation skills. One possible explanation for the lack of topic interest among a majority of the participants is the fact that there are no nuclear power plants in Norway and the topic is not heatedly debated. Thus, many participants were against the use of nuclear power plants but did not experience the issue to be very relevant in a Norwegian context and consequently did not invest much interest in the topic. Additionally, with the context of the task being a research project, participants' interest in the task and topic may have been lower than under other circumstances (Bråten et al., 2018b; Britt et al., 2018).

Our second research question concerned potential differences in participants' prior knowledge across the default stances. No statistically significant differences were found, but medium effect sizes suggested that participants displaying a disengaged default stance had a somewhat lower prior knowledge score than participants in the two other clusters. Although the correlation was moderate, the association between prior knowledge and source evaluation skills might have contributed to this result (see also, Bråten et al., 2011). Regarding the affective dimension, a relationship between prior knowledge and interest has been demonstrated in prior studies (Bråten et al., 2014; Strømsø & Bråten, 2009), whereas a relationship between prior knowledge and attitudes seems more uncertain (Allum et al., 2008; Strømsø & Bråten, 2017). The results from the present study showed no relationships between prior knowledge and the components of the affective engagement dimensions. The lack of a prior knowledge—attitudes relationship was not surprising given results from prior studies, and it is consistent with an assumed distinction between the cognitive component of attitudes (beliefs) and knowledge (Wolfe & Griffin, 2018). Regarding the prior knowledge—interest relationship, a number of studies have not identified such a relationship when students are low in interest (Schiefele, 1999; Tobias, 1994), which was also the case in the present study. Finally, prior knowledge did not relate to any of the processing or outcome measures. The participants' knowledge about nuclear power was not particularly high. O'Reilly and colleagues (2019) suggested that a certain level (threshold) of prior knowledge might be necessary for such knowledge to facilitate text comprehension. Thus, the participants in the present study might generally have had too little prior knowledge to be able to profit from what they already knew when reading and using the texts.

In their theoretical model, List and Alexander (2017) also assumed that the default stances would be related to a number of multiple-text processing variables. Therefore, our third research question focused specifically on whether participants categorized in different profile groups would differ in processes related to text selection and reading. Analyses showed no statistically significant differences among the groups on the text selection and reading measures. It is, however, worth noting that participants in the critical analytic group had higher scores than the other groups on all the processing measures and that medium effect sizes appeared for some of the differences. For example, the critical analytic group selected more texts than did the disengaged group, while both the disengaged and critical analytic groups produced more content-based justifications for text selections than did the evaluative group. In general, the results thus showed a trend in the direction of more thorough text selection behavior among students in the critical analytic group than in the two other

groups, with this trend also aligned with the predictions set forth by List and Alexander (2017). Regarding total reading time, the critical analytic group, as expected, used somewhat more time than the other two groups, although there were not statistically significant differences among the groups.

Our last research question concerned differences among the profile groups on measures based on the written products. Again, there was a general tendency for higher scores for the critical analytic group than for the two other groups. Although not statistically significant, there was a medium effect size for the difference in the number of words between the evaluative and the critical analytic groups, indicating that the critical analytic participants invested somewhat more effort in writing from the multiple texts. The critical analytic participants also included statistically significantly more information units and switches between those units in the written products than did participants in the other two groups, indicating better content coverage and integration, respectively. Those results are in line with List and Alexander's (2017) assumptions regarding the performance of the different profile groups on measures of recall and integration, and they are also consistent with the potential role of interest and strategic reading demonstrated in Bråten et al. (2014). Finally, the number of source feature references in the written products did not differ significantly among the three groups. However, only 18% of the students included any source feature references at all.

In summary, our results do, to some extent, support the structure of the cognitive affective engagement model of List and Alexander (2017, 2018) and predicted relationships between the model's default stances and indicators of processing and products in a multiple text context. In that respect, we believe the CAEM could be helpful in developing a better understanding of the role of individual differences in students' multiple text use. Specifically, affective engagement has been lacking in prior models (e.g., Brand-Gruwel & van Strien, 2018; Rouet & Britt, 2011). However, although interest and attitudes are certainly relevant for students' reading of multiple texts (e.g., Richter & Maier, 2017; Strømsø & Bråten, 2009; van Strien et al., 2014), other variables in the affective domain have also been demonstrated to affect reading of single and multiple texts (Mason et al., 2017; Wigfield, Gladstone, & Turci, 2016). Accordingly, Britt et al. (2018) suggested including several additional variables related to achievement goals, task values, and self-beliefs in their recent RESOLV model. Thus, although the CAEM should be empirically examined in further studies, we also suggest that the roles of other variables from the affective domain are studied more thoroughly in multiple text contexts.

Our study has several limitations, of course. For example, to be able to identify all four default stances of the CAEM, more participants may have to be interested in the topic of the reading task. Further, the ecological validity of the task and context may affect the results, with high-stakes tasks potentially mobilizing more engagement and effort than the present researcher-initiated task. Yet another issue related to the affective engagement dimension of the model is that by measuring the affective and cognitive components of attitudes separately (e.g., See, Petty, & Fabrigar, 2013), the affective engagement dimension of the CAEM may be represented in a more complete way. As suggested by the IF-MT framework (List & Alexander, 2019), several other individual difference variables may also affect students' default stances. More

specifically, in addition to the affective and cognitive variables included in the present study, students' epistemic beliefs and reading motivation are hypothesized to influence the default stances represented in the CAEM. Future studies should therefore examine the relationships between those variables and the CAEM profiles.

Regarding processing measures, the present study primarily assessed students' text selection behavior, whereas List and Alexander (2017) also connected the CAEM profiles to other aspects of students' processing of multiple texts, such as strategic verification of texts' content. And, although reading time has been used as an indicator of text processing in several previous multiple-text studies (Bråten et al., 2014, 2018b, 2019; List et al., 2019), we can not exclude the possibility that some students may display longer reading times for reasons other than actively engaging with the texts, for example because they are mind-wandering or lack basic word-level or comprehension skills (Latini, Bråten, Anmarkrud, & Salmerón, 2019). Data from eye-tracking or think-aloud studies may capture different reading processes more fully and should therefore be considered for future studies.

Finally, the model was tested in a relatively small group of Norwegian high-school students. While the sample size obviously affected statistical significance (or the lack of it) in this study, it could be argued that attention to the effect sizes of the differences between the profiles may be more relevant than attention to the level of statistical significance (Kline, 2004; Wasserstein & Lazar, 2016). Accordingly, we focused on the substantial effect sizes of the differences across profiles in the present study, also when these differences did not reach a conventional level of statistical significance. That said, further studies should be conducted including not only larger samples but also other populations.

Despite the limitations of the present study, we believe that the results may have both theoretical and practical implications. First, our results provide preliminary support for List and Alexander's theoretical model concerning students' various default stances when facing multiple-text tasks, as well as for some of the hypothesized relationships between those stances and processes and products of multiple text comprehension. Thus, the CAEM could be considered a fruitful model for future studies on why students relate differently to multiple-text tasks. Regarding educational implications, identifying student profiles within multiple-document comprehension may provide insights into subgroups that exist within a community of learners and help instructors adapt their instructional approach to various subgroups. As an example from reading research, McMaster et al. (2012), who constructed profiles based on the text processing of struggling readers, showed that students in different profiles responded differentially to interventions, thus demonstrating the potential utility of profile analysis for classroom practice. In contrast, a variable-centered approach that treats the sample (rather than the person) as the unit of analysis and focuses on the average (rather than the personal) may be less informative when applied to classroom practice (Chen, 2012; Molden & Dweck, 2006). In particular, our study indicated that a focus on developing students' source evaluation habits is probably not sufficient to improve multiple-document comprehension for many students. The affective engagement dimension of the CAEM needs to get more attention in instruction of multiple-document comprehension. This is consistent with a recent review indicating that the engagement dimension has not

been sufficiently emphasized in the majority of prior intervention studies (Brante & Strømsø, 2018). Presumably, teachers need to create affectively engaging reading tasks for individual students to both mobilize and develop the skills needed to become competent twenty-first-century readers.

Acknowledgement Open Access funding provided by University of Oslo (incl Oslo University Hospital). The research reported in this article was funded by Grant 237981/H20 from the Research Council of Norway to Ivar Bråten and Helge I. Strømsø.

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/>.

Appendix A

Sample items for the topic interest measure:

I'm interested in conditions that influence how safe it is to use nuclear power plants.

The safety of nuclear power is of interest to me.

I like to keep updated on issues concerning nuclear power plants.

Appendix B

Sample items for the topic knowledge measures (correct answers are in italics).

3. Radioactive waste from nuclear power plants is ...

- (a) spent radon
- (b) *spent uranium fuel*
- (c) spent bio-fuel
- (d) spent nuclear power

5. The responsibility for ensuring that nuclear power is used for peaceful purposes is held by ...

- (a) The OECD's Atomic Energy Institute
- (b) *The International Atomic Energy Agency*

- (c) The UN Panel on Climate Change
- (d) The International Renewable Energy Agency

9. A radioactive element...

- (a) *has unstable atomic nuclei*
- (b) has stable atomic nuclei
- (c) is among the lightest elements
- (d) emits visible radiation

References

- 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. 69–114). New York: Routledge.
- Ajzen, I. (1989). Attitude structure and behavior. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude, structure, and function* (pp. 241–274). Hillsdale: Erlbaum.
- Alexander, P. A., & the Disciplined Reading and Learning Research Laboratory. (2012). Reading into the future: Competence for the 21st century. *Educational Psychologist, 47*, 259–280. <https://doi.org/10.1080/00461520.2012.722511>.
- Allum, N., Sturgis, P., Tabourazi, D., & Brunton-Smith, I. (2008). Science knowledge and attitudes across cultures: A meta-analysis. *Public Understanding of Science, 17*, 35–54. <https://doi.org/10.1077/0963662506070159>.
- 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>.
- Braasch, J. L. G., & Bråten, I. (2017). The Discrepancy-Induced Source Comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist, 52*, 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*. New York: Routledge.
- Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology, 38*, 180–195. <https://doi.org/10.1016/j.cedpsych.2013.03.003>.
- Brand-Gruwel, S., & van Strien, J. L. H. (2018). Instruction to promote information problem solving on the Internet in primary and secondary education: A systematic literature review. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 401–422). New York: Routledge.
- Brante, E. W., & Strømsø, H. I. (2018). Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educational Psychology Review, 30*, 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., Brante, E. W., & Strømsø, H. I. (2018a). What really matters: The role of behavioural engagement in multiple document literacy tasks. *Journal of Research in Reading, 41*, 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*, 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*, 48–70. <https://doi.org/10.1080/00461520.2011.538647>.
- Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (2013). Justification beliefs and multiple-documents comprehension. *European Journal of Psychology of Education, 28*, 879–902. <https://doi.org/10.1007/s10212-012-0145-2>.
- Bråten, I., McCrudden, M. T., Stang Lund, E., Brante, E. W., & Strømsø, H. I. (2018b). Task-oriented learning with multiple documents: Effects of topic familiarity, author expertise, and content relevance on document selection, processing, and use. *Reading Research Quarterly, 53*, 345–365. <https://doi.org/10.1002/rrq.197>.
- Britt, M. A., Richter, T., & Rouet, J.-F. (2014). Scientific literacy: The role of goal-directed reading and evaluation in understanding scientific information. *Educational Psychologist, 49*, 104–122. <https://doi.org/10.1080/00461520.2014.916217>.
- Britt, M. A., Rouet, J.-F., & Durik, A. M. (2018). *Literacy beyond text comprehension. A theory of purposeful reading*. New York: Routledge.
- Britt, M. A., & Sommer, J. (2004). Facilitating text integration with macro-structure focusing tasks. *Reading Psychology, 25*, 313–339. <https://doi.org/10.1080/02702710490522658>.
- Bulger, M. E., Mayer, R. E., & Metzger, M. J. (2014). Knowledge and processes that predict proficiency in digital literacy. *Reading and Writing: An Interdisciplinary Journal, 27*, 1567–1583. <https://doi.org/10.1007/s11145-014-9507-2>.
- Chen, J. A. (2012). Implicit theories, epistemic beliefs, and science motivation: A person-centered approach. *Learning and Individual Differences, 22*, 724–736. <https://doi.org/10.1016/j.lindi.2012.07.013>.
- Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). Sussex: Wiley.
- Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010). Summary versus argument tasks when working with multiple documents: Which is better for whom? *Contemporary Educational Psychology, 35*, 157–173. <https://doi.org/10.1016/j.cedpsych.2009.11.002>.
- Guthrie, J. T., & Klauda, S. L. (2014). Effects of classroom practices on reading comprehension, engagement, and motivations for adolescents. *Reading Research Quarterly, 49*, 387–416. <https://doi.org/10.1002/rrq.81>.
- Guthrie, J. T., & Klauda, S. L. (2016). Engagement and motivation processes in reading. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reading, text, and context* (pp. 41–53). New York: Routledge.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111–127. https://doi.org/10.1207/s15326985ep4102_4.
- Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating internet sources: From versatile evaluators to uncritical readers. *Journal of Educational Computing Research, 39*, 75–95. <https://doi.org/10.2190/EC.39.1.e>.
- Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.
- Kobayashi, K. (2014). Students' consideration of source information during the reading of multiple texts and its effect on intertextual conflict resolution. *Instructional Science, 42*, 183–205. <https://doi.org/10.1007/s11251-013-9276-3>.
- Kulikowich, J. M., & Sedransk, N. (2012). Current and emerging design and data analysis approaches. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook* (Vol. 1, pp. 33–60). Theories, constructs, and critical issues Washington, DC: American Psychological Association.
- Latini, N., Bråten, I., Anmarkrud, Ø., & Salmerón, L. (2019). Investigating effects of reading medium and reading purpose on behavioral engagement and textual integration in a multiple text context. *Contemporary Educational Psychology, 59*, 101797. <https://doi.org/10.1016/j.cedpsych.2019.101797>.
- Lawless, K. A., & Kulikowich, J. M. (1996). Understanding hypertext navigation through cluster analysis. *Journal of Educational Computing Research, 14*, 385–399. <https://doi.org/10.2190/DVAP-DE23-3XMV-9MXH>.
- Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., & Henry, L. A. (2013). New literacies: A dual level theory of the changing nature of literacy, instruction, and assessment. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 1150–1181). Newark: International Reading Association.

- Lewandowsky, S., & Oberauer, K. (2016). Motivated rejection of science. *Current Directions in Psychological Science*, 25, 217–222. <https://doi.org/10.1177/0963721416654436>.
- List, A., & Alexander, P. A. (2017). Cognitive affective engagement model of multiple source use. *Educational Psychologist*, 52, 182–199. <https://doi.org/10.1080/00461520.2017.1329014>.
- List, A., & Alexander, P. A. (2018). Cold and warm perspectives on the cognitive affective engagement model of multiple source use. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 34–54). New York: Routledge.
- List, A., & Alexander, P. A. (2019). Toward an integrated framework of multiple text use. *Educational Psychologist*, 54, 20–39. <https://doi.org/10.1080/00461520.2018.1505514>.
- List, A., Stephens, L. A., & Alexander, P. A. (2019). Examining interest throughout multiple text use. *Reading and Writing: An Interdisciplinary Journal*, 32, 307–333. <https://doi.org/10.1007/s11145-018-9863-4>.
- 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>.
- McCrudden, M. T., Stenseth, T., Bråten, I., & Strømsø, H. I. (2016). The effects of author expertise and content relevance on document selection: A mixed methods study. *Journal of Educational Psychology*, 108, 147–162. <https://doi.org/10.1037/edu0000057>.
- McMaster, K. L., van den Broek, P., Espin, C. A., White, M. J., Rapp, D. N., Kendeou, P., et al. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences*, 22, 100–111. <https://doi.org/10.1016/j.lindif.2011.11.017>.
- McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 297–384). New York: Elsevier.
- Molden, D. C., & Dweck, C. S. (2006). Finding “meaning” in psychology. *The American Psychologist*, 61, 192–203. <https://doi.org/10.1037/0003-066X.61.3.192>.
- O’Reilly, T., Wang, Z., & Sabatini, J. (2019). How much knowledge is too little? When a lack of knowledge becomes a barrier to comprehension. *Psychological Science*, 30, 1344–1351. <https://doi.org/10.1177/0956797619862276>.
- Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 88–108). Mahwah: Erlbaum.
- Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist*, 52, 148–166. <https://doi.org/10.1080/00461520.2017.1322968>.
- Rouet, J.-F. (2006). *The skills of document use: From text comprehension to web-based learning*. Mahwah: 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). Charlotte: Information Age.
- 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, 85–106. https://doi.org/10.1207/s1532690xci1501_3.
- Rouet, J.-F., Ros, C., de Pereyra, G., Macedo-Rouet, M., & Salmerón, L. (2013). Teenagers’ developing awareness of source quality. *Paper presented at the annual meeting of the Society for Text and Discourse*, Valenica, Spain.
- Salmerón, L., Strømsø, H. I., Kammerer, Y., Stadler, M., & van den Broek, P. (2018). Comprehension processes in digital reading. In M. Barzillai, J. Thomson, S. Schroeder, & P. van den Broek (Eds.), *Learning to read in a digital world* (pp. 91–120). Amsterdam: John Benjamin.
- Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading*, 3, 257–279. https://doi.org/10.1207/s1532799xssr0303_4.
- See, Y. H. M., Petty, R. E., & Fabrigar, L. R. (2013). Affective-cognitive meta-bases versus structural bases of attitudes predict processing interest versus efficiency. *Personality and Social Psychology Bulletin*, 39, 1111–1123. <https://doi.org/10.1177/0146167213490807>.
- Stenseth, T., Bråten, I., & Strømsø, H. I. (2016). Investigating interest and knowledge as predictors of students’ attitudes towards socio-scientific issues. *Learning and Individual Differences*, 47, 274–280. <https://doi.org/10.1016/j.lindif.2016.02.005>.

- Strømsø, H. I., & Bråten, I. (2009). Beliefs about knowledge and knowing and multiple-text comprehension among upper secondary students. *Educational Psychology, 29*, 425–445. <https://doi.org/10.1080/01443410903046864>.
- Strømsø, H. I., & Bråten, I. (2017). Students' trust in research-based results about potential health risks presented in popular media. *Bulletin of Science, Technology & Society, 37*, 3–14. <https://doi.org/10.1177/0270467617740495>.
- Strømsø, H. I., Bråten, I., & Samuelstuen, M. S. (2008). Dimensions of topic-specific epistemological beliefs as predictors of multiple text understanding. *Learning and Instruction, 18*, 513–527. <https://doi.org/10.1016/j.learninstruc.2007.11.001>.
- Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research, 64*, 37–54. <https://doi.org/10.2307/1170745>.
- Trevors, G. J., 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>.
- van Strien, J. L. H., Brand-Gruwel, S., & Boshuizen, H. P. A. (2014). Dealing with conflicting information from multiple nonlinear texts: Effects of prior attitudes. *Computers in Human Behavior, 32*, 101–111. <https://doi.org/10.1016/j.chb.2013.11.021>.
- 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: An Interdisciplinary Journal, 29*, 1677–1698. <https://doi.org/10.1007/s11455-015-9601-0>.
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician, 70*, 129–133. <https://doi.org/10.1080/00031305.2016.1154108>.
- Wigfield, A., Gladstone, J. R., & Turci, L. (2016). Beyond cognition: Reading motivation and reading comprehension. *Child Development Perspectives, 10*, 190–195. <https://doi.org/10.1111/cdep.12184>.
- 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*, 101266. <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*, 73–87. <https://doi.org/10.1037/0022-0663.83.1.73>.
- Wineburg, S. (1994). The cognitive representation of historical texts. In G. Leinhardt, I. L. Beck, & C. Stainton (Eds.), *Teaching and learning in history* (pp. 85–135). Hillsdale: Erlbaum.
- Wineburg, S., & Reisman, A. (2015). Disciplinary literacy in history: A toolkit for digital citizenship. *Journal of Adolescent & Adult Literacy, 58*, 636–639. <https://doi.org/10.1002/jaal.410>.
- Wolfe, M. B., & Griffin, T. D. (2018). Beliefs and discourse processing. In M. F. Schober, D. N. Rapp, & M. A. Britt (Eds.), *The Routledge handbook of discourse processes* (pp. 295–314). New York: Routledge.
- Yim, O., & Ramdeen, K. T. (2015). Hierarchical cluster analysis: Comparison of three linkage measures and application to psychological data. *The Quantitative Methods for Psychology, 11*, 8–21. <https://doi.org/10.20982/tqmp.11.1.p008>.

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