

# Using Reading Times and Eye-Movements to Measure Cognitive Engagement

Brian W. Miller  
*College of Education  
Towson University*

Self-paced reading and eye-tracking can be used to measure microlevel student engagement during science instruction. These methods imply a definition of engagement as the quantity and quality of mental resources directed at an object and the emotions and behaviors entailed. This definition is theoretically supported by models of reading comprehension and visualization comprehension. The use of eye-movement data is based on a number of assumptions, including the assumption that people look longer at words and images because they are thinking about those objects more. Self-paced reading and eye-tracking have strengths such as precision and detail as well as limitations including difficulty of interpretation. Carefully controlled research designs and triangulation of multiple methods are suggested as possible ways to address the limitations and take advantage of the strengths. Two examples drawn from the refutational text literature are described.

In the introduction to this special issue, Sinatra, Heddy, and Lombardi (this issue) suggest that engagement researchers need to consider two key issues when designing experiments. The first is how engagement is defined in the context of your research questions. Researchers can then choose a research method that matches that definition. This is especially important in terms of the components of engagement: behavioral, emotional, cognitive, or agentic (Fredricks, Blumenfeld, & Paris, 2004; Reeve & Tseng, 2011). A second critical task is to choose a method that can measure the grain-size of the phenomena in which you are interested. Sinatra et al. explain that most engagement research can be placed along a continuum ranging from microlevel engagement, in which individuals engagement is studied over brief periods of instruction, up to macrolevel engagement, in which classrooms, schools, and communities are studied over extended periods.

This article addresses two related methods for measuring engagement in science education: self-paced reading and eye-tracking. These two methods imply similar definitions of engagement, and they both measure engagement at a microlevel grain-size. This article begins by describing each of these methods. Then it analyzes the definition of

engagement that is implied by these two research methods including the definitional issues related to measuring microlevel engagement. The next section describes the theoretical assumptions that form the foundation for using these methods to measure engagement. The final section describes an example of research on engagement in science instruction using each of these methods.

## SELF-PACED READING AND EYE-TRACKING METHODS

During self-paced reading, participants are presented with a piece of text on a computer screen. This piece of text can vary in length from a single word to entire paragraphs. The participant is instructed to read the piece of text and then press a key to advance to the next piece. The computer records reading times by recording the time of each key press (see Jegerski, 2014, for a review). Self-paced reading can be used to measure engagement by examining the changes in reading time that take place between individuals or between different pieces of text. Students who are more engaged are expected to slow their rate of reading (see the upcoming section on assumptions), although in cases of increased interest the engaged participant may read faster (Broughton, Sinatra, & Reynolds, 2010).

---

Correspondence should be addressed to Brian W. Miller, College of Education, Towson University, 8000 York Road, Towson, MD 21252. E-mail: bwmiller@towson.edu

Eye-tracking requires a machine that can track the size of the pupil and the location of the eye. There are several categories of eye-tracking machines, but one common type functions by shining an infrared light into the eye and capturing the light that passes through the pupil and is reflected back by the cornea to a video camera. The camera can be located on or near the screen, or it can be mounted on goggles or a headset worn by the participant. The participant sits a known distance from a screen, and a computer coordinates the position of the eyes and what appears on the screen.

Eye-tracking can be very precise. Sampling rates of 250 times per second are routine, as are accuracies of less than  $0.5^\circ$  of visual angle. These accuracies are possible because the density of the most sensitive type of photoreceptor cell, the cone, varies dramatically across the retina (Duchowski, 2007). Only the very small region near the center of the retina called the fovea and to a lesser degree the slightly larger area around it called the parafovea have enough density of cone cells to accurately identify letters and other detailed shapes (Duchowski, 2007; Rayner, 2009). Furthermore, when reading or examining a still image, the eyes do not move smoothly across the page. The eyes pause at one particular spot, called a fixation, and then jump to the next spot. The jump is called a saccade (Duchowski, 2007). The computer uses the layout of the screen and the exact location of the pupil during a fixation to compute what part of that screen falls within the fovea. Eye-tracking can record patterns of eye movements, including when participants skip words or look back and forth in the text. Whereas self-paced reading is restricted to reading of text or looking at whole images, eye-tracking can be used for a larger variety of instructional tasks, because any stimulus that can appear on a screen can be used with eye-tracking including words, diagrams, pictures, animations, movies, and interactive computer programs. New technologies may expand the types of stimuli even further (K. M. Evans, Jacobs, Tarduno, & Pelz, 2012).

Self-paced reading and eye-tracking collect data while the student is reading or looking at images. Using these methods, participants do not have to stop the activity to have their level of engagement recorded. Methods of this type are often called “online measures.” Other online measures include measures that collect data as participants complete activities such as the analysis of logs of what keys students press while working with an interactive software program (Gobert, Baker, & Wixon, this issue; Winne, 2010). Online measures differ from other microlevel methods like self-report that do require the participant to stop at least briefly to answer questions (e.g., self-regulated learning; Cleary & Zimmerman, 2001; Kitsantas & Zimmerman, 2002).

The use of self-paced reading and eye-tracking methodologies are not yet widely utilized as measures of engagement in science education. For example, several prominent reviews of the engagement literature with sections focusing

on measurement including Fredricks et al. (2004) and Jimerson, Campos, and Greif (2003) do not mention self-paced reading or eye-movement data. However, some programs of research related to science instruction such as computer tutoring (D’Mello & Graesser, 2011) and educational game design (Renshaw, Stevens, & Denton, 2009) have embraced these methods.

## DEFINITIONS AND GRAIN-SIZE OF ENGAGEMENT AS MEASURED BY SELF-PACED READING AND EYE-TRACKING

Self-paced reading and eye-tracking are both microlevel measures. As Sinatra et al. (this issue) describe, microlevel measures are defined by short lengths of instruction, the unit of analysis focusing on the individual, and a theoretical framework that characterizes engagement “by the cognitive, emotional, or motivational engagement of the individual learner” (p. 22). Self-paced reading and eye-tracking meet all of these criteria. They allow for frequent and precise sampling over short periods of instruction. The unit of analysis is most likely to be at the person level rather than the communal level, because measuring reading times and eye movements are so sensitive to individual differences. For example, comparing the average reading speed of different classrooms of children on a piece of text is unlikely to be informative, because the variance between individuals will be large and the fine-grained detail will be lost. Finally, the theoretical perspective tends to focus on the cognitive, emotional, and motivational aspects of individuals, because these constructs are what are most accessible to these methods. To accommodate the short instructional time frame, the definition of engagement used in conjunction with these methods should emphasize the situational rather than trait aspects of engagement. Furthermore, the unit of analysis and the theoretical framework suggest a definition of engagement that focuses on the mental processes of individual students.

### Engagement as the Quantity and Quality of Mental Processes

In self-paced reading and eye-movement research, the speed of reading and the movements of the eye are interpreted as indirect measurements of human cognition. The data that are generated include times and locations. These might seem to be inherently quantitative. However, as is explained throughout this article, self-paced reading and eye-tracking also yield qualitative descriptions of students’ engagement. As such, these methods imply a discontinuous scale of engagement that changes quantitatively and qualitatively across a range. At the lowest end, engagement begins with the simple act of paying attention. Given our understanding of attentional blindness, paying attention is

the baseline below which instruction is simply wasted (Simons & Resnick, 2005). Beyond paying attention, students can allocate varying amounts and types of cognitive resources. Students who are not very engaged would use a small amount of mental resources to conduct shallow processing and heuristic strategies. At high levels of engagement, students would use large amounts of mental resources to conduct deep processing and metacognitive strategies to evaluate and reconstruct their understanding of the material in concert with their prior beliefs.

This view of engagement is similar to the view of the continuum of engagement described in Dole and Sinatra's (1998) cognitive reconstruction of knowledge model (CRKM). This difference is not accidental. The CRKM is influenced by dual-processing theories of cognition (J. B. T. Evans, 2007). These theories suggest that there are two groups of mental processes. One group, often called System 1, is automatic and requires few cognitive resources, whereas another group, often called System 2, is deliberate and requires more cognitive resources. The existence of these two systems helps to explain how measuring the amount of time a student reads a sentence can give clues about the type of thinking in which they are engaged.

### Engagement Components

The definition of engagement as the quantity and quality of mental resources directed at an object of thought strongly implies that self-paced reading and eye-tracking are primarily measurements of the cognitive component of engagement. Although these methods are most closely connected with this component, the emphasis can be overstated. As Sinatra et al. (this issue) warns, the components of engagement (Fredricks et al., 2004) are overlapping and can be difficult to disassociate. I contend that the distinctions between components become increasingly blurred as the grain-size of the measurement decreases. At larger grain-sizes the different components of engagement can be more easily measured, and their individual effects on student outcomes are more easily identified. For example, the emotional engagement involved in feeling competent in mathematics can determine students' future class choices, which will have clear effects on how much science they will eventually learn (Wang, 2013). These emotions can be differentiated from the other components of engagement by carefully worded questionnaires. However, at very short time scales it is difficult to differentiate between components, and it is difficult to see how a particular component has a unique affect on achievement. This is particularly true for eye-tracking for which behavior, emotion, and cognition are all reflected in the same slight movements of the eye. Even for self-paced reading the same slowing of reading speed could indicate an emotional reaction of confusion or an error in interpretation. This increased overlap of components means that the definition should also include the

emotional response (e.g., desire or interest) that might lead to using greater mental resources and the behaviors (e.g., rereading ambiguous words or lingering on key details) that might indicate high metacognitive engagement.

One benefit of using this multifaceted definition is that it helps to differentiate engagement from motivation. Cleary and Zimmerman (2001) defined motivation as an intention while engagement is an action, or as they put it, motivation is the "will" to learn and engagement is the "skill" of actually performing the tasks involved in learning (p. 240). Motivation in this context is closely related to engagement, but it is possible to have the will to perform deep strategies and not actually do so. Engagement in the context of these methods includes the quantity and quality of mental resources as well as the emotional reactions and behaviors that lead to and enable using those resources.

### Definition of Engagement and Theories of Reading and Visual Perception

Self-paced reading and eye-tracking methods are primarily used to measure students' engagement while reading, looking at pictures, or interacting with computer programs. Given the instructional context in which these methods are usually employed, it is important that the definition of engagement that is being used in conjunction with these methods be congruent with existing theories of reading comprehension and visualization comprehension including the landscape model (LM; van den Broek, Young, Tzeng, & Linderholm, 1999), the construction-integration model (van Dijk & Kintsch, 1983), and the model of visualization comprehension (Hegarty, 2005). In this section I briefly outline each theory and how it corresponds with the definition just described.

*The landscape model.* The LM (van den Broek et al., 1999) depicts readers as having the ultimate goal of creating a mental model of the text. Readers differ in their expectations of how coherent their mental model of the text should be. These expectations are called a standard of coherence. Standards of coherence (van den Broek et al., 1999) are similar to the element of desire and interest just described in the definition of engagement.

The LM further explains that due to limitations in attention and working memory, people can attend to only a limited number of concepts at a time and can hold only a limited number of concepts in working memory; therefore, at any point in time a reader must choose which concepts to activate. When reading an easily understood text, this process can occur with great automaticity and relatively low cognitive effort. In contrast, when the text is difficult, as is usually the case with science texts, readers need to continuously monitor and strategically adjust the allocation of their cognitive resources if they are to achieve a highly coherent mental model. In such instances, the LM would predict that

readers who have a high standard will slow their eye movements to give themselves time to accomplish these strategies, but students with low standards will not slow their reading, indicating that they do not have the desire to understand the text and are not using the quantity and quality of cognitive resources that would be necessary. This depiction of high engagement requiring the use of deep cognitive strategies as well as the behavioral and emotional investment that they imply matches very closely with the definition of engagement just given.

*Construction-integration model.* The construction-integration model does not address issues of engagement in the same way as the LM, but it provides a more detailed account of how reading times and eye movements will change in accordance with increased engagement. The construction-integration model of discourse comprehension (Kintsch, 1998; van Dijk & Kintsch, 1983) proposes that there are three levels of discourse representation—surface level, textbase, and situation model. The surface level consists of the meaning and syntax of specific words. The textbase is the meaning derived from the relationship across groups of words. The situation model is an understanding of what the text is about as a whole. Measurements of reading time through self-paced reading or eye-tracking combined with statistical modeling can identify the time that is needed to process the text at each of these levels. Even readers who have similar reading ability will read texts matched on all of these levels at slightly different speeds (Stine-Morrow, Miller, Gagne, & Hertzog, 2008). Some researchers have used this remaining variance as an indication of engagement, because it can represent the effort exerted above and beyond what would normally be required to comprehend the text (Stine-Morrow et al., 2008). For example, Miller et al. (2014) used this approach to show that students who read to prepare for a discussion read more slowly than students without a stated purpose for reading. Defining engagement as dedicating resources above and beyond the level necessary for simple comprehension is also congruent with the cognitive aspects of the definition of engagement just described, although it does not include the emotional and behavior elements that are included in the LM.

*Visual comprehension models.* In addition to words, science texts, especially textbooks, are heavily infused with visual representations often having more than one image on every page (Canham & Hegarty, 2010; Florax & Ploetzner, 2010; Seufert, Schütze, & Brünken, 2009). The leading models of visualization comprehension such as Hegarty's theory of visualization comprehension (Hegarty, 2005; Kriz & Hegarty, 2007) have many similar features to the reading comprehension models just described. Due to these similar features, visual comprehension models also support the use of eye-tracking to measure engagement (van Gog &

Scheiter, 2010). Like in other models, the leading models of visualization comprehension assume people have limited mental resources. Creating a useable mental model based on either words or images requires a coordination of top-down and bottom-up processes. In fact, evidence suggests that students require even more cognitively demanding strategies when interpreting diagrams than when reading a science text (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010). Due to the limitations of people and the complexity of the task, comprehension above a certain threshold of difficulty requires the use of strategic allocation of cognitive effort. This requires students who are highly engaged to slow their eye movements to achieve comprehension with images in the same way that it is needed to comprehend text (Hyöna, 2010).

## FOUNDATIONAL ASSUMPTIONS

The use of self-paced reading and eye-tracking to measure engagement as defined in this article hinges on the critical assertion that the speed of reading and the movements of the eye are indications of the quantity and quality of mental resources being used. This assertion is based on making one or more of the four assumptions described in this section. By using the word *assumption*, I am not suggesting that these ideas are taken on faith. There is evidence. Rather, I am emphasizing that the truth of these ideas is foundational to the use of these methods and that their veracity is rarely questioned in research papers. These assumptions are described in the next section, which leads to a discussion of some of the strengths and limitations of these methods. Finally, the article describes two examples of how self-paced reading and eye-tracking can be used to measure engagement.

### Assumption 1—Minimum Reading Times and Foveal View Measure Attention

In the context of self-paced reading and eye-tracking methodologies, the baseline of engagement is attention. Students cannot be even minimally engaged in a science lesson if they are not focusing attention on the material being presented. Both self-paced reading and eye-tracking can be used to establish this baseline.

In the case of self-paced reading, participants are instructed to press a key when they have finished reading a piece of text. Participants do not always follow these instructions. Sometimes participants are so disengaged that they blindly advance through the text without paying attention to the words. Researchers can identify when this occurs by looking for reading times that previous research has shown are too short to allow the reader to comprehend the text. For example, for a single word, it has been established that college-age participants take an average of

approximately 400 ms between key presses (Rayner & Clifton, 2009). If elementary-age participants spend much less than 400 ms on average per word, then they are probably not even minimally engaged in reading the passage. Similar timing minimums can be estimated for other sizes of text.

An eye-tracking machine can measure if participants are paying attention to words or pictures by measuring if they have chosen to place an object in their foveal view for a minimum amount of time. The fovea is very narrow, and only information falling within the fovea can be seen with enough detail to identify words and object features. If a student fixes his or her foveal vision on an object, that alone signals that the student is likely to be devoting at least some attention onto that object (Duchowski, 2007). Furthermore, we have estimates for the average time most people fixate on certain objects. However, although eye movements can identify with reasonable certainty if regions of an image or stretches of words are being ignored, the same conclusion cannot be made about single words or visual features, because single words and features can be skipped entirely and yet their meaning can be inferred by context or identified using limited information from the parafoveal region (Rayner, 2009).

The first assumption is supported by the correlation between reading time and comprehension (Miller et al., 2014), but it is undermined by the phenomena of mind wandering. People sometimes look at objects while their minds wander. Sometimes people can read an entire page before realizing they have no memory of the text they just read. Recently it has been found that participants who are paying attention and students whose minds are wandering have distinctive patterns of eye movements (Reichle, Reineberg, & Schooler, 2010); therefore, there is potential that in the near future assumption one will be strengthened because students will be able to be screened for mind wandering.

### Assumption 2—The Eye-Mind-Engagement Assumption

Assumption 1 describes how reading times and eye movements can be used to infer a baseline level of engagement. Assumption 2 is the theoretical foundation for using these methods to measure higher levels of engagement. Assumption 2 is that the length of reading times or the length of time an eye rests on a word or object can reflect the quantity and quality of cognitive effort a student is using to understand that word or object.

Assumption 2 is derived from what Just and Carpenter (1976) called the eye–mind assumption. According to Just and Carpenter (1980), “the eye–mind assumption, is that the eye remains fixated on a word as long as the word is being processed. So the time it takes to process a newly fixated word is directly indicated by the gaze duration” (p. 330). As evidence, Just and Carpenter (1980) cited the fact that

readers take longer to read words that are thematically important or infrequently used in the English language. These observations are evidence of the eye–mind hypothesis, because these words are more difficult to comprehend but they are not more difficult to decode. The eye–mind assumption can explain why participants look at these words longer, whereas theories of reading based on decoding cannot.

The last four decades of research using a wide variety of participants and visual stimuli have provided many new examples of situations in which people tend to take longer to read words and to examine objects that require more thought to comprehend. For example, participants look at pronouns longer if the antecedent is difficult to identify (Ehrlich & Rayner, 1983). Readers take longer to read words that have ambiguous meanings (Duffy, Morris, & Rayner, 1988) and words that require inferences (O’Brien, Shank, Myers, & Rayner, 1988). Readers also spend more time on interesting sentences than on less interesting sentences (Wade, Schraw, Buston, & Hayes, 1993). In addition, readers spend more time on relevant or important sentences (Kaakinen & Hyönä, 2005; Wade et al., 1993) and sentences related to the purpose for reading (Reynolds & Anderson, 1982; Rothkopf & Billington, 1979).

One might argue that the eye–mind assumption as proposed by Just and Carpenter (1980) applies to amounts of processing rather than engagement. However, the definition of engagement previously described connects the two constructs. People do not always exert more effort to read more difficult texts. People only apply the amount of effort that is required to reach the level of comprehension that they desire. Furthermore, theories of reading comprehension, visual comprehension, and dual-processing theories suggest that as people exert more quantity of effort they must switch to qualitatively different strategies (Reynolds, 1992). The application of the self-paced reading and eye-tracking is only beginning to be used to research engagement in science education (Smith, Mestre, & Ross, 2010), but there are examples of how reading time and eye movements can be well explained by an eye–mind–engagement assumption. One example is Madsen, Larson, Loschky, and Rebello (2012), who found that students who looked at the relevant areas of science diagrams longer were more likely to answer physics questions correctly. Another example is research on refutational texts, discussed later in this article.

Although the eye–mind–engagement assumption has theoretical support, it must be interpreted with caution. The first reason is that changes in reading time and eye movements can be used as a measurement of engagement only when all of the other variables have been accounted for or controlled. The variance in eye fixations and reading times across words in the same individual is primarily determined by the difficulty of recognizing the words (e.g., the number of syllables) and identifying the meaning of the words in a sentence (e.g., the frequency of the word in the English language; Rayner, 2009). Most of the variance in reading times

on the same word between individuals can be accounted for by differences in the reading fluency of each individual. The differences in eye movements and reading time that are caused by engagement can be determined only when the difficulty of the text and the skill and knowledge of the readers are already accounted for or controlled.

Even when these ancillary variables are controlled, interpreting reading times and fixations can still be difficult. This difficulty stems from the fact that self-paced reading and eye-tracking can indicate that a piece of text or image caused the participant to think more, but it does not indicate what the participant thought about. The participant can be thinking about errors in reading, distractions, perceived ambiguities, or even related topics brought to mind by the text or image such as intertextual connections. These participants might be highly engaged in the reading task but not in the way that the researcher imagines.

Furthermore, some researchers have suggested that the effects of increased cognitive effort can be masked by other forces that cause the reading time to decrease. In particular, interest and positive affect might allow people to use their cognitive resources more effectively, thereby reducing the reading time (Broughton et al., 2010; O’Keefe & Linnenbrink-Garcia, 2014; Shirey & Reynolds, 1988).

### Assumption 3—Like More, Look More

A corollary to Assumption 2 is that the length of time the eyes rest on a word or object feature reflects the strength of the emotional reaction to that word or object. Although this assumption has had very limited applications in science engagement research, it has become widely accepted in other fields. In advertising, Maughan, Gutnikov, and Stevens (2007) proposed the principle of “like more, look more. Look more, like more” (p. 335). They found “a robust correlation between the number/duration of fixation a person makes, while looking at an advertisement and whether or not they like the advertisement” (p. 341). This finding has contributed to the wide use of eye-tracking technology in marketing research (Pieters & Wedel, 2004). In research on computer game design, Renshaw et al. (2009) found that certain eye movements were associated with dissatisfaction. Jennett et al. (2008) found that immersion in a game was associated with fewer but longer fixations. The Look More Like More assumption could eventually be the basis of a more multifaceted measure of microlevel engagement.

### Assumption 4—Pupil Size

The final assumption applies only to eye-tracking because it utilizes the size of the pupil. This assumption asserts that the size of the pupil reflects a student’s increased cognitive

effort and/or emotional arousal. The size of the pupils are largely determined by the brightness of objects and their distance from the viewer, but once these changes are accounted for or controlled, small changes in the size of the pupil can be observed (Beatty & Lucero-Wagoner, 2000). The pupils have been shown to increase in size when participants were asked to solve difficult problems, identify sounds within noise, recall large amounts of information, and many other challenging situations (Beatty & Lucero-Wagoner, 2000). This technique has also been used to measure emotional arousal. For example, Bradley, Miccoli, Escrig, and Land (2008) found that pupil size increased when students looked at emotionally charged pictures and that these changes correlated with traditional physiological measures of arousal such as skin conductance. Educational applications of this method are rare. Intelligent tutoring systems have been designed to measure and react to users’ emotions, but they tend to rely on features of the entire face rather than pupil dilation alone (D’Mello & Graesser, 2011; Kolakowska, Landowska, Szwoch, Szwoch, & Wrobel, 2013). Like eye movements, pupil size must be used with caution because the measure indicates an increase in arousal and not its cause. For example, emotional arousal could have been caused by the scientific text being presented to a participant, or it could have been caused by an unrelated sound, smell, or passing thought. However, the basic science connecting pupil size and emotions is beginning to reach a level that may allow for educational applications in the near future (Alghowinem, AlShehri, Goecke, & Wagner, 2014; Soleymani, Lichtenauer, Pun, & Pantic, 2012).

## LIMITATIONS OF SELF-PACED READING AND EYE-TRACKING AS MEASUREMENTS OF ENGAGEMENT

As just described, using self-paced reading and eye-tracking as measurements of engagement is based on a number of assumptions that are reasonable but nonetheless are not always true. This is the case with all forms of measurement. For example, self-report assumes that students have interpreted the questions as you intended, that students have accurate knowledge of themselves, and that students are truthful. However, methods like self-report have very well established procedures and statistical methods including factor analysis and item response theory that can be used to address issues of reliability and validity. Some of the eye-tracking techniques such as measuring mind wandering and emotional arousal have not had time to develop mature procedures. The field will be required to pioneer some methodological techniques to ensure reliability and validity as these methods becomes more widely adopted in engagement research.

### Short Well-Structured Interventions Presented on a Screen

One important limitation of self-paced reading and eye-tracking is that they are best suited to short well-structured instructional interventions that can be presented on a screen. These limitations are created by the technological restraints as well as the methodological weaknesses. The technological restraints are created by the fact that both methods are traditionally presented on a computer. This is required for self-paced reading, and it greatly simplifies eye-tracking. There are head-mounted eye-trackers that can be used as participants move freely even in open environments such as geology field trips (K. M. Evans et al., 2012), but the results are difficult to interpret. However, eye-tracking technology continues to develop rapidly, and it will eventually be able to accommodate different types of lessons. In the meantime, being tied to a screen is an important limitation in studying modern science education with its emphasis on participation in scientific and engineering practices. The ideal science classroom is a highly active place that includes classroom discussions, activities, experiments, group work, and the use of various technologies.

The limitations on possible experimental designs are also created by difficulties of interpretation. The effects of engagement need to be isolated from all of the ancillary variables and the nature of the thought needs to be probed. One common approach to solving this problem is to employ carefully designed interventions. Statistical methods can help isolate engagement from other factors (Baayen, Davidson, & Bates, 2008; Richter, 2006), but no statistical procedure can substitute for carefully crafted research design that counterbalances treatments and limits unnecessary variation between groups. In addition to careful designs, researchers are encouraged to combine multiple measures. Sinatra et al. (this issue) call for this type of triangulation. Self-paced reading and eye-tracking data can be combined with self-report measures, think-aloud protocols, discourse analysis, and other methods. By triangulating, researchers are more likely to maintain ecological validity and to answer research questions that are pertinent to the modern science classroom.

### Age Limitations

Another limitation stems from the fact that the eye–mind connection presumes that participants have control of the amount of cognition they are exerting. This is also a presumption of major reading and visual comprehension theories such as the landscape theory. Although this presumption is reasonable, researchers in the field of self-regulation suggest that this presumption might become less plausible for children at very young ages. Some children seem to have trouble setting learning goals, monitoring their own progress, and adjusting their learning strategies accordingly (Zimmerman, 1990). As Butler and Winne (1995) explained, “Several

lines of research suggest that students typically monitor sub-optimally” (p. 261). Without the skills of regulating their strategies and cognitive effort, Assumption 2 no longer holds. The degree of this limitation depends on how challenging the task is and the abilities of the particular students. When reading very difficult texts or using sophisticated study strategies like outlining, this might set a relatively high age and ability limit (Pressley, Ghatala, Woloshyn, & Pirie, 1990; Zimmerman, 1990). Fortunately, in the case of age-appropriate reading, research would suggest that ability and age limits are low. A review by Reichle et al. (2013) found that although children read more slowly than adults in general, their pattern of reading speed is very similar to adults in response to word length, word frequency, and thematic role plausibility (the likelihood of a noun argument being plausible given the theme of the text), suggesting that they monitor their allocation of mental resources in a very similar manner to adults.

### Addressing the Limitations of Eye-Tracking Through Using Multiple Indices

One solution to some of the limitations of eye-tracking in particular deserves special mention. One way that eye-tracking can be made more sensitive is by using a variety of indices in addition to the total amount of time a student fixates on a word or image feature. The most informative additional index in terms of measuring engagement is the regression time (sometimes called the second pass time; see upcoming section on Ariasi & Mason, 2011, for an example of its use). Usually students go from one word to the next or from one image feature to the next as they read or examine images; however, sometimes they will also go backward in order to review a word or feature. This is called a *regression*, and it constitutes about 10%–15% of saccades in skilled readers (Rayner, Chase, Slattery, & Ashby, 2006). Regressions are of particular interest to engagement researchers because readers tend to regress more when they are having difficulty understanding the text (Mikkilä-Erdmann, Penttinen, Anto, & Okinuora, 2008; Sinatra & Broughton, 2011), and increased regressions are associated with difficult conceptual leaps such as conceptual change on a science topic (Ariasi & Mason, 2011; Hyönä, Lorch, & Kaakinen, 2002; Mikkilä-Erdmann et al., 2008). Regressions by definition constitute “extra effort” in that they interrupt the flow of reading and require the reader to exert extra effort to repair errors or to reconsider information.

## STRENGTHS OF SELF-PACED READING AND EYE-TRACKING AS MEASUREMENTS OF ENGAGEMENT

As Sinatra et al. (this issue) suggest, researchers need to match their questions and definitions of engagement

with the methods that they choose. As the limitations underscore, self-paced reading and eye-tracking are not suitable for some research questions concerning engagement in science education, including questions that are exploratory or centered on highly active instructional activities. It is also not appropriate for studying context-oriented macrolevel engagement. However, these microlevel methods of measuring engagement are capable of providing powerful insights on a range of important engagement questions pertaining to science instruction.

One such issue is the idea that engagement is a mediator between classroom interventions and student science achievement (Appleton, Christenson, Kim, & Reschly, 2006; Greene, Miller, Crowson, Duke, & Akey, 2004; Skinner, Furrer, Marchand, & Kindermann, 2008). The view of engagement as a mediator resonates well with theories of science education that include emotional, affective, and motivational factors. The CRKM (Dole & Sinatra, 1998) is an example of such a theory in the area of conceptual change. In the CRKM, engagement serves as a mediator between motivational factors, learner characteristics, instructional features, and conceptual change. Through iterative cycles of instruction, the context of learning influences the degree that students are engaged in the learning task. For example, the match between the text and the readers interest, background knowledge, and prior beliefs will influence their willingness to engage with the text. In turn, the level of engagement may influence the process of conceptual change by determining how deeply students will metacognitively reflect on inconsistencies between their beliefs and the concepts expressed in the text. In this model, engagement does not guarantee conceptual change, but the types of deep metacognitive strategies associated with high engagement ensure that if the student does change his or her conception, that change will be more complete, persistent, and influential.

Self-paced reading, eye-tracking, and other similar microlevel measures of engagement are particularly suited to answer warm conceptual change questions because social, motivational, and affective states can be fleeting and subtle (Graesser & D'Mello, 2011). Shifts in attention also take place over very brief periods, and students are not always aware of where their attention is focused. The ability of eye-tracking to identify changes in engagement during specific tasks allows the researcher to tie specific contexts to specific outcomes and to show that engagement was the mediating factor. Self-report and traditional observations might be useful under these circumstances, but they do not have the precision of eye-tracking, and they require self-awareness in the participants.

#### Example of Using Eye-Tracking to Measure Engagement—Refutational Texts

To better illustrate the strengths and limitations of self-paced reading and eye-tracking as measures of engagement

in science instruction, I now describe an example of research using each method. Although these methods can be used in many areas of science learning, I have chosen to present examples that investigate the same topic to facilitate comparison. The topic I have chosen is refutational texts. This topic is a particularly appropriate subject to investigate with these methods, because instruction through refutational texts can take place over a brief period and because it can be easily built into well-controlled experimental designs that are presented on a computer screen.

Refutational texts in the context of science learning are texts that include not only the scientifically accepted account of a natural phenomenon but also a statement of a common misconception and why it is false. Such texts differ from traditional expository science texts that present only the scientifically accepted explanation. Many studies have shown that refutational texts are more likely to produce conceptual change than traditional expository texts (Guzzetti, Snyder, Glass, & Gamas, 1993; Sinatra & Broughton, 2011; Tippett, 2010). One possible explanation for this result is that these texts allow students to directly compare their prior conceptions with the scientific conception (van den Broek & Kendeou, 2008). Although this explanation has experimental support, there might be other mechanisms that work in conjunction, such as increased engagement. Broughton et al. (2010) used self-paced reading and Ariasi and Mason (2011) used eye-tracking to explore this possibility. These experiments are presented to show how these methods contributed to this area of research and how the researchers addressed the strengths and limitations of the methodologies through their experimental designs.

#### An Example of Self-Paced Reading as a Measure of Engagement in Science—Broughton, Sinatra, and Reynolds (2010)

Broughton et al. (2010) illustrated several of the themes of this article, including the need for caution when interpreting reading times and eye movements and the importance of triangulating reading times and eye movements with other measures. Broughton et al. began by describing a series of self-paced reading studies that produced inconsistent results. Kendeou and van den Broek (2005) reported no differences in the reading times of participants with misconceptions and those without misconceptions when reading a refutational text, but Kendeou and van den Broek (2007) did find that students with misconceptions read slower if they were reading a refutational text. Furthermore, van den Broek and Kendeou (2008) combined self-paced reading with think-aloud protocols to show that students with misconceptions reading a refutational text read slower than students without misconceptions and that they used conceptual change strategies more than other students.

Broughton et al. (2010) hypothesized that these inconsistencies existed because readers who found the text

salient slowed their reading, whereas readers who found the text interesting sped up their reading. Broughton et al. used self-paced reading to test this hypothesis. Students read refutational texts more quickly than expository texts. In addition, Broughton et al. found that it was the refutational sentences that accounted for the difference in time. Students who read the refutational text had fewer misconceptions in an immediate posttest, and they deemed the refutational text both interesting and important. Broughton et al. hypothesized that this result can be explained by the fact that the readers found the refutational text interesting. Interest in a text might make processing easier leading to faster reading times (O'Keefe & Linnenbrink-Garcia, 2014; Shirey & Reynolds, 1988).

This experiment illustrates both the strengths and limitations of the self-paced reading methodology. This study was motivated by the conflicting results of van den Broek and Kendeou (2008) using self-paced reading. This alone illustrates the limitation of the method that it is difficult to interpret until the critical variables have been identified and experimentally manipulated. On the other hand, the fact that the van den Broek and Kendeou experiments motivated the Broughton et al. (2010) study illustrates the strength of reading time methodologies in being able to generate interesting research questions, because the microlevel analysis provides detailed findings that can be further probed.

The methodological choices of Broughton et al. (2010) were examples of some of the ways that interpretation of reading times can be simplified. First, they based their experiment on previous studies, allowing them to identify variables that might be important. They then attempted to isolate the target variables from ancillary variables. They controlled for text characteristics by making the texts vary only in very specific and limited ways between conditions. They also controlled for individual differences in reading ability by including a reading test as a covariate. Finally they combined reading time measurements of overall as well as specific sentences with interviews and pre- and posttest assessments. This triangulation of methods allowed for the development of an interpretation of their findings. Broughton et al. (2010) does not settle the question of how refutational texts work, but the study does advance the field by producing a testable hypothesis that complicates and deepens our understanding of science learning through refutational texts.

#### Ariasi and Mason (2011)

Ariasi and Mason (2011) had students with misconceptions about the tides read either a refutational or traditional expository text about the subject. The texts were controlled for length, number of sentences, and sentence length. Participants completed a pre- and posttest of conceptual

understanding. The eye movements of the participants were recorded while the students read the texts.

The results of the eye movement data illustrate some the limitations and strengths of this method as compared to self-paced reading. Self-paced reading generates one measurement of time per piece of text, because participants read one piece at a time and are not allowed to go backward in the text. In contrast, eye-tracking can generate a large number of different indices for each piece of text. Ariasi and Mason (2011) chose to use the amount of time the participant spent when they first looked at either the refutational or the control text, the total amount of time spent on the text describing the scientific explanation, and the amount of time participants spent the second time they viewed the scientific explanation (if any). This greater detail allows for even more precise measurement that can potentially generate interesting comparisons, but it also makes interpretation even more difficult, because these different indices can differ in their results. This also makes it challenging to compare results from self-paced reading and eye-tracking studies. For example, Ariasi and Mason (2011) found that participants reading the refutational text spent less time on the refutational portion of the text. This is the same result found by Broughton et al. (2010). However, participants spent more time on the scientific portion especially the second time that they looked, which would seem to diverge from Broughton et al. and to replicate van den Broek and Kendeou (2008).

One method that Ariasi and Mason (2011) used to help interpret their results was to triangulate the eye-tracking data with measures of conceptual understanding. This analysis showed that students who read the refutational text learned more if they read the scientific explanation more slowly and the refutational text more quickly, whereas eye-movements did not predict how much students learned from reading the expository text. Once again, this pattern of results generates a fascinating series of hypotheses that advances the field even if it might not be conclusive.

## CONCLUSION

As suggested by Sinatra et al. (this issue) researchers need to carefully choose the proper methodology to match their research questions. This choice is critical, because different methodologies imply different definitions of engagement. Furthermore, different methodologies are better suited to properly answer questions at specific grain-sizes. The research on refutational texts is a good example of the type of microlevel engagement research that has the potential to make important contributions to issues like science textbook design; use of diagrams, animations, and other visuals; the design and use of educational games; and many other questions. Although significant limitations exist, researchers using these methods can take advantage of great

precision and detail, in addition to a well-developed theoretical framework.

## ACKNOWLEDGMENTS

Special thanks to Gale Sinatra, Richard C. Anderson, Elizabeth Stine-Morrow, and Kiel Christianson for their help and guidance.

## REFERENCES

- Alghowinem, S., AlShehri, M., Goecke, R., & Wagner, M. (2014). Exploring eye activity as an indication of emotional states using an eye-tracking sensor. In L. Chen, S. Kapoor, & R. Bhatia (Eds.), *Intelligent systems for science and information* (pp. 261–276). Cham, Switzerland: Springer International. [http://dx.doi.org/10.1007/978-3-319-04702-7\\_15](http://dx.doi.org/10.1007/978-3-319-04702-7_15)
- Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology, 44*, 427–445. <http://dx.doi.org/10.1016/j.jsp.2006.04.002>
- Ariasi, N., & Mason, L. (2011). Uncovering the effect of text structure in learning from a science text: An eye-tracking study. *Instructional Science, 39*, 581–601. <http://dx.doi.org/10.1016/j.learninstruc.2010.01.001>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412. <http://dx.doi.org/10.1016/j.jml.2007.12.005>
- Beatty, J., & Lucer-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142–162). Cambridge, MA: Cambridge University Press.
- Bradley, M. M., Miccoli, L., Escrig A., & Land, P. J. (2008) The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45*, 602–607. <http://dx.doi.org/10.1111/j.1469-8986.2011.01223.x>
- Broughton, S. H., Sinatra, G. M., & Reynolds, R. E. (2010). The nature of refutation text effect: An investigation of attention allocation. *The Journal of Educational Research, 103*, 407–423. <http://dx.doi.org/10.1007/s11165-011-9274-x>
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*, 245–281.
- Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction, 20*, 155–166. <http://dx.doi.org/10.1016/j.learninstruc.2009.02.014>
- Cleary, T. J., & Zimmerman, B. J. (2001). Self-regulation differences during athletic practice by experts, non-experts, and novices. *Journal of Applied Sport Psychology, 13*, 185–206. <http://dx.doi.org/10.1080/104132001753149883>
- Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology, 35*, 59–74. <http://dx.doi.org/10.1080/104132001753149883>
- D’Mello, S., & Graesser, A. (2011). The half-life of cognitive-affective states during complex learning. *Cognition & Emotion, 25*, 1299–1308. <http://dx.doi.org/10.1080/02699931.2011.613668>
- Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist, 33*, 109–128. <http://dx.doi.org/10.1080/00461520.1998.9653294>
- Duchowski, A. (2007). *Eye tracking methodology: Theory and practice*. New York, NY: Springer-Verlag.
- Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27*, 429–446.
- Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior, 22*, 75–87. [http://dx.doi.org/10.1016/S0022-5371\(83\)80007-3](http://dx.doi.org/10.1016/S0022-5371(83)80007-3)
- Evans, J. B. T. (2007). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*, 255–278. <http://dx.doi.org/10.1146/annurev.psych.59.103006.093629>
- Evans, K. M., Jacobs, R. A., Tarduno, J. A., & Pelz, J. B. (2012). Collecting and analyzing eye tracking data in outdoor environments. *Journal of Eye Movement Research, 5*, 1–19.
- Florax, M., & Ploetzner, R. (2010). What contributes to the split-attention effect? The role of text segmentation, picture labeling, and spatial proximity. *Learning and Instruction, 20*, 216–224.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*, 59–109. <http://dx.doi.org/10.1016/j.learninstruc.2009.02.021>
- Gobert, J. D., Baker, R. S., & Wixon, M. B. (this issue). Operationalizing and detecting disengagement within online science microworlds. *Educational Psychologist, 50*.
- Graesser, A., & D’Mello, S. K. (2011). Theoretical perspectives on affect and deep learning. In R. A. Calvo & S. K. D’Mello (Eds.), *New perspectives on affect and learning technologies* (pp. 11–21). New York, NY: Springer.
- Greene, B. A., Miller, R. B., Crowson, M., Duke, B., & Akey, K. (2004). Predicting high school students’ cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary Educational Psychology, 29*, 499–517. <http://dx.doi.org/10.1016/j.cedpsych.2004.01.006>
- Guzzetti, B. J., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly, 28*, 117–159.
- Hegarty, M. (2005). Multimedia learning about physical systems. In R. E. Mayer (Ed.), *Handbook of multimedia* (pp. 447–465). New York, NY: Cambridge University Press.
- Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction, 20*, 172–176. <http://dx.doi.org/10.1016/j.learninstruc.2009.02.013>
- Hyönä, J., Lorch, R. F., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94*, 44–55. <http://dx.doi.org/10.1037/0022-0663.94.1.44>
- Jegerski, J. (2014). Self-paced reading. In J. Jegerski & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 20–49). New York, NY: Routledge.
- Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*, 641–661. <http://dx.doi.org/10.1145/1520340.1520494>
- Jimerson, S. R., Campos, E., & Greif, J. L. (2003). Toward an understanding of definitions and measures of school engagement and related terms. *The California School Psychologist, 8*, 7–27.
- Just, M. A., & Carpenter P. A. (1976). The role of eye-fixation research in cognitive psychology. *Behavior Research Methods & Instrumentation, 8*, 139–143.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.
- Kaakinen J. K., & Hyönä, J. (2005). Perspective effects on expository text comprehension: Evidence from think-aloud protocols, eyetracking, and recall. *Discourse Processes, 40*, 239–257. [http://dx.doi.org/10.1207/s15326950dp4003\\_4](http://dx.doi.org/10.1207/s15326950dp4003_4)

- Kendeou, P., & van den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology, 97*, 235–245. <http://dx.doi.org/10.1037/0022-0663.97.2.235>
- Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition, 35*, 1567–1577. <http://dx.doi.org/10.3758/BF03193491>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.
- Kitsantas, A., & Zimmerman, B. J. (2002). Comparing self-regulatory processes among novice, non-expert, and expert volleyball players: A microanalytic study. *Journal of Applied Sport Psychology, 14*, 91–105. <http://dx.doi.org/10.1080/10413200252907761>
- Kolakowska, A., Landowska, A., Szwoch, M., Szwoch, W., & Wróbel, M. R. (2013). Emotion recognition and its application in software engineering. In *The 6th International Conference on Human System Interaction (HSI)* (pp. 532–539). Poland: IEEE.
- Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies, 65*, 911–930. <http://dx.doi.org/10.1016/j.ijhcs.2007.06.005>
- Madsen, A. M., Larson, A. M., Loschky, L. C., & Rebello, N. S. (2012). Differences in visual attention between those who correctly and incorrectly answer physics problems. *Physical Review Special Topics-Physics Education Research, 8*, 010122–1–010122–13.
- Maughan, L., Gutnikov, S., & Stevens, R. (2007). Like more, look more. Look more, like more: The evidence from eye-tracking. *Journal of Brand Management, 14*, 335–342. <http://dx.doi.org/10.1057/palgrave.bm.2550074>
- Mikkilä-Erdmann, M., Penttinen, M., Anto, E., & Okinura, E. (2008). Constructing mental models during learning from science text. In P. Ifenthaler, P. Pimay-Dummer, & J. M. Spector (Eds.), *Understanding models for learning and instruction: Essays in honor of Norbert M. Steel* (pp. 62–77). New York, NY: Springer.
- Miller, B. W., Anderson, R. C., Morris, J., Lin, T. J., Jadallah, M., & Sun, J. (2014). The effects of reading to prepare for argumentative discussion on cognitive engagement and conceptual growth. *Learning and Instruction, 33*, 67–80.
- O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 410–420. <http://dx.doi.org/10.1037/0278-7393.14.3.410>
- O'Keefe, P. A., & Linnenbrink-Garcia, L. (2014). The role of interest in optimizing performance and self-regulation. *Journal of Experimental Social Psychology, 53*, 70–78. <http://dx.doi.org/10.1016/j.jesp.2014.02.004>
- Pieters R., & Wedel M. (2004). Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing, 68*, 36–50. <http://dx.doi.org/10.1509/jmkg.68.2.36.27794>
- Pressley, M., Ghatala, E. S., Woloshyn, V., & Pirie, J. (1990). Sometimes adults miss the main ideas and do not realize it: Confidence in responses to short-answer and multiple-choice comprehension questions. *Reading Research Quarterly, 25*, 232–249.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506. <http://dx.doi.org/10.1080/17470210902816461>
- Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10*, 241–255. [http://dx.doi.org/10.1207/s1532799xssr1003\\_3](http://dx.doi.org/10.1207/s1532799xssr1003_3)
- Rayner, K., & Clifton, C. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology, 80*, 4–9.
- Reeve, J., & Tseng, C. M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology, 36*, 257–267. <http://dx.doi.org/http://dx.doi.org/10.1016/j.cedpsych.2011.05.002>
- Reichle, E. D., Livergood, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S., White, S. J., & Rayner, K. (2013). Using EZ Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review, 33*, 110–149.
- Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science, 21*, 1300–1310. <http://dx.doi.org/10.1177/0956797610378686>
- Renshaw, T., Stevens, R., & Denton, P. D. (2009). Towards understanding engagement in games: An eye-tracking study. *On the Horizon, 17*, 408–420. <http://dx.doi.org/10.1108/10748120910998425>
- Reynolds, R. E. (1992). Selective attention and prose learning: Theoretical and empirical research. *Educational Psychology Review, 4*, 1–48. <http://dx.doi.org/10.1007/BF01332144>
- Reynolds, R. E., & Anderson, R. C. (1982). Influence of questions on the allocation of attention during reading. *Journal of Educational Psychology, 74*, 623–632. <http://dx.doi.org/10.1037/0022-0663.74.5.623>
- Richter, T. (2006). What is wrong with ANOVA and multiple regression? Analyzing sentence reading times with hierarchical linear models. *Discourse Processes, 41*, 221–250. [http://dx.doi.org/10.1207/s15326950dp4103\\_1](http://dx.doi.org/10.1207/s15326950dp4103_1)
- Rothkopf, E. Z., & Billington, M. J. (1979). Goal-guided learning from text: Inferring a descriptive processing model from inspection times and eye movements. *Journal of Educational Psychology, 71*, 310–327. <http://dx.doi.org/10.1037/0022-0663.71.3.310>
- Seufert, T., Schütze, M., & Brünken, R. (2009). Memory characteristics and modality in multimedia learning: An aptitude–treatment–interaction study. *Learning and Instruction, 19*, 28–42. <http://dx.doi.org/10.1016/j.learninstruc.2008.01.002>
- Shirey, L. L., & Reynolds, R. E. (1988). The effect of interest on attention and learning. *Journal of Educational Psychology, 80*, 159–166. <http://dx.doi.org/10.1037/0022-0663.80.2.159>
- Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*, 16–20. <http://dx.doi.org/10.1016/j.tics.2004.11.006>
- Sinatra, G. M. & Broughton, S. H. (2011). Bridging reading comprehension and conceptual change in science education: The promise of refutation text. *Reading Research Quarterly, 46*, 374–393. <http://dx.doi.org/10.1002/RRQ.005>
- Sinatra, G. M., Heddy, B. C., & Lombardi, D. (this issue). The challenges of defining and measuring student engagement in science. *Educational Psychologist, 50*.
- Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*, 765–781. <http://dx.doi.org/10.1037/a0012840>
- Smith, A. D., Mestre, J. P., & Ross, B. H. (2010). Eye-gaze patterns as students study worked-out examples in mechanics. *Physical Review Special Topics-Physics Education Research, 6*, 020118. <http://dx.doi.org/10.1103/PhysRevSTPER.6.020118>
- Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2012). A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing, 3*, 42–55. <http://dx.doi.org/10.1109/T-AFCC.2011.25>
- Stine-Morrow, E. A. L., Miller, L. M. S., Gagne, D. D., & Hertzog, C., (2008). Self-regulated reading in adulthood. *Psychology and Aging, 23*, 131–153. <http://dx.doi.org/10.1037/0882-7974.23.1.131>
- Tippett, C. D. (2010). Refutation text in science education: A review of two decades of research. *International Journal of Science and Mathematics Education, 8*, 951–970. <http://dx.doi.org/10.1007/s10763-010-9203-x>
- van den Broek, P., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: The role of co-activation in confronting misconceptions. *Applied Cognitive Psychology, 22*, 335–351. <http://dx.doi.org/10.1002/acp.1418>
- van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of a

- memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71–98). Mahwah, NJ: Erlbaum.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic.
- van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*, 95–99. <http://dx.doi.org/10.1016/j.learninstruc.2009.02.009>
- Wade, S. E., Schraw, G., Buston, W. M., & Hayes, M. T. (1993). Seduction of the strategic reader: Effects of interest on strategies and recall. *Reading Research Quarterly, 28*, 93–114.
- Wang, X. (2013). Why students choose STEM majors motivation, high school learning, and postsecondary context of support. *American Educational Research Journal, 5*, 1081–1121. <http://dx.doi.org/0002831213488622>.
- Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist, 45*, 267–276. <http://dx.doi.org/10.1080/00461520.2010.517150>
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*, 3–17. [http://dx.doi.org/10.1207/s15326985ep2501\\_2](http://dx.doi.org/10.1207/s15326985ep2501_2)

Copyright of Educational Psychologist is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.