Optional ERIC Coversheet — Only for Use with U.S. Department of Education Grantee Submissions

This coversheet should be completed by grantees and added to the PDF of your submission if the information required in this form **is not included on the PDF to be submitted**.

INSTRUCTIONS

- Before beginning submission process, download this PDF coversheet if you will need to provide information not on the PDF.
- Fill in all fields—information in this form **must match** the information on the submitted PDF and add missing information.
- Attach completed coversheet to the PDF you will upload to ERIC [use Adobe Acrobat or other program to combine PDF files]—do not upload the coversheet as a separate document.
- Begin completing submission form at https://eric.ed.gov/submit/ and upload the full-text PDF with attached coversheet when indicated. Your full-text PDF will display in ERIC after the 12-month embargo period.

GRANTEE SUBMISSION REQUIRED FIELDS

Title of article, paper, or other content

Higher order thinking in comprehension

All author name(s) and affiliations on PDF. If more than 6 names, ERIC will complete the list from the submitted PDF.

Last Name, First Name	Academic/Organizational Affiliation	ORCID ID
McNamara, D. S.	Arizona State University	
Jacovina, M. E.	Arizona State University	
Allen, L. K.	Arizona State University	

Publication/Completion Date—(if *In Press,* enter year accepted or completed) 2015

Check type of content being submitted and complete one of the following in the box below:

- ☐ If article: Name of journal, volume, and issue number if available
- ☐ If paper: Name of conference, date of conference, and place of conference
- ☑ If book chapter: Title of book, page range, publisher name and location
- ☐ If book: Publisher name and location
- ☐ If dissertation: Name of institution, type of degree, and department granting degree

Handbook of individual differences in reading: Reader, text and context, 164-176, Routledge, New York

DOI or URL to published work (if available)

Acknowledgement of Funding— Grantees should check with their grant officer for the preferred wording to acknowledge funding. If the grant officer does not have a preference, grantees can use this suggested wording (adjust wording if multiple grants are to be acknowledged). Fill in Department of Education funding office, grant number, and name of grant recipient institution or organization.

"This work was supported by U.S. Department of Education **[Office name]** National Center for Educat through **[Grant number]** R305A1301 to **Institution]** Arizona State U.The opinions expressed are those of the authors and do not represent views of the **[Office name]** National Center for Educa or the U.S. Department of Education.

Higher Order Thinking in Comprehension

Danielle S. McNamara, Matthew Jacovina, and Laura Varner

Learning Sciences Institute, Arizona State University

Reading is a pervasive activity in the classroom, as well as in everyday activities: comprehending text and discourse is crucial to success and survival in the modern world. Nonetheless, many students struggle to understand text at even a basic level, and even more fail to construct deep level understandings of content. Particularly for complex academic texts, students may understand individual words and sentences; yet, they frequently fail to comprehend the underlying meaning of the material, have little memory for the content, and struggle to learn from the texts. As a result, one goal of educators and reading researchers has been to optimize conditions such that readers engage in the higher level cognitive processes that aid in the construction of coherent, interconnected, and elaborated mental representations of content material. At the heart of this objective is the notion that students should engage in higher order thinking in order to understand material at deep levels (which is conducive to learning).

In this chapter, we discuss the role of higher order thinking in the context of text comprehension. We first describe a sample of frameworks that have been proposed to delineate lower and higher level processes leading to higher order processing. We then discuss some of the challenges faced by educators and researchers in defining and distinguishing lower and higher level processes, processes versus outcomes, as well as issues arising from individual differences among students. Ultimately, we suggest theoretical and educational implications of these concepts, emphasizing the importance of considering the needs of individual readers through sensitivity to the interactions among tasks, processes, and individual differences among students.

Frameworks for Higher Order Thinking

The Bloom Taxonomy. As a starting point, we should define the concept of *higher order thinking*. Consider the results of a Google search (the ultimate higher level answer to all questions), which yielded the following definition[1]:

> *Higher-order thinking, also known as higher order thinking skills, is a concept of Education reform based on learning taxonomies such as Bloom's Taxonomy. The idea is that some types of learning require more cognitive processing than others, but also have more generalized benefits.*

Indeed, one of the most commonly used frameworks to describe the distinction between different levels of processing, and in particular higher order thinking, is the Bloom Taxonomy (e.g., Bloom, 1956; Krathwohl, Bloom, & Masia, 1964; Simpson, 1972). This framework is used across a wide variety of contexts, most often in educational settings (e.g., Granello, 2001; Hanna, 2007). The Bloom Taxonomy, named after Benjamin Bloom, was developed in the 1950s to distinguish between different types of educational objectives. The taxonomy includes three types of processes: cognitive (knowledge, comprehension, application, analysis, synthesis, evaluation), affective (receiving, responding, valuing, organizing, characterizing), and psychomotor (perception, set, guided response, mechanism, complex overt response, adaptation, origination).

The most common use of Bloom Taxonomy has been to distinguish between cognitive processes. The underlying notion is that students must acquire and master the skills at lower levels in order to advance to higher levels on the spectrum. The taxonomy has been used

---

[1] http://en.wikipedia.org/wiki/Higher-order_thinking

frequently in the development of educational standards to distinguish between types of processes or understanding that would be expected from students at different levels of development. For example, in the context of understanding text, readers would move from being able to describe, and then to evaluating and analyzing concepts within a text. Hence, a student at a lower level within a set of standards might be expected to use knowledge to describe the ideas, whereas students at more advanced levels may be expected to evaluate or analyze the ideas. In turn, rendering judgments based on criteria or evidence (evaluation) goes beyond and partially depends on breaking down ideas into parts (analysis). Hence there is a hierarchical organization wherein processes expected at more advanced levels depend on those at lower levels as the student moves toward higher order thinking. In this sense, higher order thinking comprises cognitive processes such as analysis, synthesis, and ultimately evaluation, in contrast to the use of knowledge, comprehending, and applying information. In turn, activities that call on higher order thinking are assumed to lend to a better, deeper understanding of the material: evaluating an idea is expected to lead to better learning than is describing an idea.

The ICAP Framework. The Bloom Taxonomy has been widely used within the field of education. Another similar approach recently proposed by Michelene Chi (2009), the ICAP (Interactive, Constructive, Active, Passive) framework, focuses on ranking the relative value of learning tasks according to their overt, observable activities and their potential underling cognitive processes. Activities are considered *passive* when students are not visibly doing anything. Thus, watching videos or reading texts would both be considered passive learning activities. Students are defined as *active* when they are overtly doing something during a learning task, such as sliding beads on an abacus or highlighting key themes in a short story. *Constructive* activities require students to produce something that extends beyond the educational material.

This could involve writing down novel uses for various objects or verbally self-explaining sentences in a science text. Finally, students are considered *interactive* when they engage with another person or computer system through dialogue or another joint activity. Some interactive activities may not provide learning benefits above other *constructive* tasks, such as when one student simply provides another student with information, rather than engaging them in conversation. Other interactions, however, that require students to more thoughtfully construct responses and explanations, to formulate their own questions, and to consider someone else's perspective can be more beneficial than basic constructive activities.

This framework primarily relies on information about students' overt behaviors to classify learning activities; however, it also suggests potential cognitive processes that may underlie these activities. For instance, active tasks engage students' attentional processes, (I will cross-reference the chapter on attention here) which can help them to incorporate new information from the instructional materials with their prior knowledge, while simultaneously reinforcing that prior knowledge in memory (I will cross-reference the chapter on prior knowledge here). Constructive tasks require students to engage in inferential processes, which allow them to develop a deeper understanding of the material. According to the framework, interactive processes are similar to constructive processes. However, when students engage in dialogue, they may more easily develop an understanding of complex concepts, thanks to the contributions of an interlocutor.

In contrast to the Bloom Taxonomy, the ICAP framework focuses more on the learning activities rather than on the processes or the outcomes. For example, whereas the Bloom Taxonomy might focus on a student's ability to explain a text (usually after having read it), the ICAP framework would consider the reading process itself to be inherently passive, and focus more on whether the student had explained the text while reading it (which might be *active* or *constructive*). Hence, one focuses more on the assumed processes and outcomes, whereas the other focuses primarily on the overt learning task.

Importantly, both the Bloom Taxonomy and the ICAP framework have strong potential to inspire insights about how individual differences influence higher order thinking for readers. For example, these frameworks may encourage researchers and educators to consider the types of cognitive processes that differ from person to person, or the particular tasks that might allow different readers to thrive. Unfortunately, however, considerations for individual differences are not built into these two frameworks. Therefore, educators and researchers who utilize these frameworks are at risk of making assumptions of homogeneity among learners. Not all readers will progress through Bloom's hierarchy in the same manner, nor will a task lead to the same cognitive processes and ultimate learning benefits for all participants. Thus the frameworks would greatly benefit from the addition of direct guidance that can aide educators in better understanding how individual differences among students may interact with various task demands to influence comprehension and learning.

Comprehension Models. In contrast to frameworks such as the Bloom Taxonomy and the ICAP model, most contemporary models of text and discourse comprehension focus on *the stuff in the middle*: the processes associated with understanding. Most pertinent to higher order thinking, comprehension models differentiate between multiple levels of text understanding (e.g., Kintsch, 1998; McNamara & Magliano, 2009), most notably, the *surface*, *textbase*, and *situation model* levels of understanding (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983).

Accordingly, a reader forms a mental representation of a text. This representation is generally considered in connectionist terms to comprise concepts (nodes) and relations (links) and includes spreading activation between concepts that are explicitly stated within the text as well as to concepts that are unstated in the text, but available in prior knowledge. The surface understanding comprises the explicit words and their relations in the text. The textbase essentially refers to the aspect of the mental representation reflecting understanding and memory for content that occurs within the text. The situation model refers to the aspect of the mental representation that reflects deeper understanding, or the integration and elaboration of ideas in the text and concepts from outside of the text (e.g., prior knowledge; a separate source). Hence, the situation model is most strongly associated with higher order thinking.

Deep understanding of a text is conceptualized in terms of the coherence of the reader's mental representation. The degree to which this representation has strong, appropriate connections between concepts within the text and connections to unstated knowledge is the degree to which the representation is coherent and stable. In turn, the coherence and stability of the representation predicts outcomes such as comprehension and memory for the text. Accordingly, the mental representation of the text comprises various aspects, including a textbase and a situation model. Notably, these are not assumed to be hierarchically organized or dissectible from the representation, but rather qualities, much like colors in a painting.

In contrast to educationally oriented approaches to higher order thinking, comprehension models focus on the quality of the mental representation constructed by the reader or comprehender. In turn, different types of assessments are assumed to pick up on these qualities of a reader's mental representation. Just as standing at a different angle or distance affords different perceptions of a painting, the objective in comprehension assessment is to collect various views and angles in order to infer the reader's mental representation. The textbase level of understanding might be inferred from multiple-choice questions, cloze tasks, questions about individual sentences, or paraphrasing. The quality of a reader's situation model is inferred from tasks that rely on the reader having made connections between concepts in the text (e.g., bridging inference questions) or to prior knowledge (e.g., elaboration questions). For example, if a participant were to perform well on questions focused on textbase level information and poorly on situation model questions, then it would likely be deduced that the reader had formed a coherent textbase but had not generated the kinds of inferences necessary to understand the text at a deep level (e.g., McNamara & Kintsch, 1996).

Most models of comprehension assume that prior knowledge is a critical factor in influencing the extent to which individuals construct deep understandings of text. When readers attempt to comprehend a text, they can either be limited or supported by what they already know. Thus, individual differences in prior knowledge for a particular text topic can play a crucial role in a student's comprehension. Outside of prior knowledge, however, most comprehension models tend to place little emphasis on the role of other individual differences among readers, such as motivation or strategy knowledge. Moreover, they rarely focus on the interactions between these individual differences and the various properties of texts that students are attempting to comprehend (e.g., cohesion, topic knowledge, audience, etc.) – in other words, how different readers might engage in different processes for different types of texts. Given the acknowledgement of these interactions, a challenge for researchers is to work towards the development of more comprehensive models of the comprehension process, more specifically incorporating components of both the text and the reader (McNamara & Magliano, 2009).

In relation to the concept of higher order thinking, models of comprehension make a number of predictions. First, assessments that require higher order thinking are more likely to reveal the quality of a reader's situation model (i.e., the extent that the reader incorporated prior knowledge and constructed a coherent mental model of the text). Second, having a reader engage in activities associated with higher order thinking (e.g., explaining, evaluating) while learning from text is more likely to enhance comprehension. These predictions have been well supported in the literature (e.g., Coté, Goldman, & Saul, 1998; Eason et al., 2012; Kintsch, 1998; van der Schoot, Horsley, & van Lieshout, 2010). Nonetheless, there are multiple caveats to these predictions, as we describe in the following sections.

Important Caveats

Lower vs. Higher Level Processing. Higher order thinking implies that it involves higher level processing. One important consideration is how to delineate the line between what is considered *lower* and what is considered *higher*. Lower level processes are generally assumed to be computationally easier than higher level processes, requiring less explicit attention and effort, and relying less on unstated information or knowledge. For example, reading a word is considered to be relatively automatic for a developed reader, and familiar words are read more automatically (e.g., the, a, of, cat, dog), without conscious attention, than are less familiar words (e.g., terricolous, idyll, soporific). Such assumptions lead to definitions of lower and higher level processes as being more or less *automatic* versus *controlled* or *bottom-up* versus *top-down*. As such, across a number of fields, labeling processes as lower or higher level affords relatively quick and easy distinctions about the cognitive resources or knowledge required to complete a task. Such distinctions are appealing and are often quite useful to understanding cognition and learning.

One problem that emerges, however, regards the lack of specificity concerning what processes are considered higher level and the diversity of definitions for lower and higher level processing across disciplines and domains. The intended meaning of lower and higher level cognitive processes depends heavily on the particular domain, the research topic, and the context. For example, a vision researcher might consider the perception of color to be a lower level process whereas perceiving a word would be considered a higher level process. By contrast, some cognitive scientists classify all vision processes as lower level and all language processes as higher level (König, Kühnberger, & Kietzmann, 2013). Similarly, for some reading researchers, lexical decoding might be classified as a lower level process and sentence understanding as a higher level process, whereas for other reading researchers, sentence understanding exemplifies lower level processes and inferencing and elaboration involve higher level processes. Within their respective contexts, these distinctions often make sense. However, across domains, disciplines, and contexts, these terms lose definitional precision. Most importantly, where the line is drawn between lower and higher levels of processing depends on the range of processes under consideration. Hence, the task of delineating a universal definition of higher order thinking seems daunting at best. Moreover, given the utility of these distinctions within domains, regardless of the variance across domains, many researchers would not *want* to adopt a global, all-encompassing definition. They like their own definition.

Unfortunately, inconsistent distinctions made between lower and higher level processes are not always innocuous. For example, conclusions about higher level processing in visual research are unlikely to map onto conclusions about higher level processing in reading research. Education researchers and educators often have different perceptions of what higher level implies. “Higher order thinking” in the education domain often refers to specific objectives, such as the ability to evaluate texts using particular criteria (Anderson et al., 2001). By design, higher order thinking skills are meant to align with higher level cognitive processes, but because there is no standard definition of what those are, the intended alignment can be misleading.

A related concern is that attempts to import findings on higher level processing from one field to another might lead to costly incompatibilities. Consider a fledgling educator who reads an interesting finding that rereading a text allows readers to focus on higher level reading processes (e.g., Millis, Simon, & tenBroek, 1998). The educator, knowing that an objective of his English class is to encourage the higher order task of analyzing conceptual information from texts, adds a rereading task to his curriculum. The rereading task *might* lead to learning gains, but such a result would not be predicted from the findings in the Millis et al. rereading study, which operationalized the higher level reading process in terms of text-level integration. That is, the conceptual overlap is tenuous between the higher level reading processes described by Millis and colleagues and the higher order reading objectives of the (albeit fictional) educator. While this is a clearly cartoon example, similar misinterpretations are plausible and have likely been committed on much larger scales.

Processes vs. Outcomes. Another caveat to a discussion on higher order thinking, and in turn higher level processes, is the importance of clearly distinguishing between processes and outcomes. For decades, researchers have acknowledged the importance of distinguishing between learning processes and outcome measures (e.g., Kolers & Roediger, 1984). Within the context of reading comprehension research, this distinction between higher level processes and deep comprehension is crucial. The outcome of comprehension or learning processes (through tasks) is the primary variable that is observed through assessments. Typically, a participant is presented with a text (ranging from a sentence to a multi-paragraph passage), and during or after the exposure to the text, various dependent measures might be collected. In the classroom and other educational settings, these assessments often take the form of scores on classroom and standardized tests, and ultimately course grades. In a laboratory setting, these assessments may include reaction time measurements on various reading-related tasks (e.g., reading time, lexical decision-making, word naming) or questions explicitly aimed at assessing a student's comprehension of a passage (e.g., cloze tasks, true-false sentence recognition, sentence verification, multiple choice, open-ended; Cain & Oakhill, 2006; Pearson & Hamm, 2005).

These assessments are used in various ways to infer the state or quality of a reader's understanding. Online measures of comprehension, such as reaction time, cloze tests, or think aloud provide perhaps the closest assessment of the cognitive processes of a particular reader. Reaction time measures can be used to infer, for example, whether the reader makes inferences while reading, is challenged by the difficulty of the text, reads using more automatic processes, or at the least, is paying attention to the task. In particular, word and sentence reaction times are often used for fine-grained deductions regarding the effects of various manipulations such as word ambiguity, cohesion, sentence difficulty, and so on. Indeed, at fine-grained levels, reading times can be quite useful in assessing the effects of manipulations in text and discourse.

At more coarse-grained levels however, their utility is sometimes lost in the limitless processes that readers might have engaged during extended reading. When the reader's task is to process and comprehend a 1000-word text, with as many differences between each sentence as between a Manet and a Monet painting, the inferences that can be deduced from a long or short reading time can be lost in a sea of variance. Moreover, given reading times alone, it can be impossible to deduce what they mean: long reading times may indicate that the reader generated inferences, stumbled through the text, or daydreamed. For example, readers with less knowledge about a topic take might take more time to read a challenging text (compared to a less challenging text) in one situation and less time in another (McNamara & Kintsch, 1996). Each result is equally facile to explain: The readers attempted to generate inferences to resolve conceptual challenges in the first situation, and the readers did not do so (and ultimately *gave up*) in the second.

Online measures also typically assess lower level cognitive processes. For example, cloze tasks omit certain words in a text and ask the reader to fill in or choose the appropriate word. Ideally, the omitted words and the foils are chosen such that the task requires the reader to have a full understanding of the sentences and their relations, rather than a superficial understanding of the sentence. Nonetheless, the task by its very nature is correlated most highly with readers' knowledge of words and their ability to understand individual sentences. Hence, readability measures such as the Flesch Kincaid correlate very highly with performance on cloze tasks (see McNamara, Graesser, McCarthy, & Cai, 2014, for a discussion).

By contrast, think aloud measures come the closest to assessing multiple aspects of a reader's comprehension and in particular, readers' deep level comprehension. However, few researchers and even fewer educators turn to think aloud as a measure of comprehension. An exception is the work by Keith Millis and Joe Magliano (Gilliam et al., 2007; Magliano et al., 2011) who have leveraged think aloud and question asking during reading within an automated assessment called RSAT (Reading Strategy Assessment Tool). This tool asks students both indirect questions (e.g., what are you thinking now?) and direct questions (e.g., why does a tumor develop?) while reading. The reader's answers are scored automatically by comparing the words used in the answers to benchmark sets of words. The scores have been shown to correlate highly with standardized measures of reading as well as measures of post-reading comprehension.

While online measures may come closest to assessing both surface and deep comprehension processes, their use is also quite rare. Offline, or post-reading comprehension questions are more commonly used to infer how well a reader has understood a text. Typically, a set of questions is constructed about the text and the average correct is assumed to reflect how well a student understood the text, and in turn the degree to which the reader had deeply processed the text. These questions are usually multiple-choice or true/false questions because these types of questions can be most easily scored. In some cases, open-ended questions are used, requiring the reader to construct the answer without cues. These are less commonly used, both in research studies and educational settings, primarily because of the time to score the answers. Nonetheless, it is generally assumed that open-ended questions are more likely to reveal readers' deep comprehension because they call upon recall for the information in the text rather than recognition, which can be based on a reader's textbase representation (McNamara & Kintsch, 1996).

Whether the assessment is online or offline, the performance is used as a reflection of the processes engaged while reading. For instance, researchers can vary the difficulty of a particular text passage or they can change the instructions to reflect more or less difficult processes. If a student performs well on the assessment for a particularly challenging task, it may be assumed that this student engaged in higher level processes during the task. On the other hand, if the comprehension task is less difficult, performance on the assessment may be assumed to provide information about and relate to lower level processing. A notable issue however emerges when the results from outcome measures are conflated with the processes engaged while reading. Importantly, comprehension assessments do not provide *direct* information about the processes engaged by the learner during the given task. Rather, these processes are inferred based on assessment performance. Admittedly, researchers have made attempts to distinguish between the comprehension processes and outcomes from their research studies. More often than not, however, this distinction is dropped and the results become conflated over time.

Many researchers (ourselves included) describe performance based on the outcome of a particular assessment as if it were a direct reflection of the processes engaged by the learner during comprehension. Obviously, there are examples of processes and outcomes corresponding with each other. For example, McNamara et al. (2006) examined the benefits of engaging in various cognitive processes while self-explaining complex science text (i.e., after training and practice in using reading strategies while self-explaining). They then examined how those processes corresponded to outcomes on a post-training comprehension test. As expected, there was correspondence between paraphrasing and performance on textbase questions and between generating bridging inferences and performance on bridging inference questions. The field is full of examples similar to this one. Thus, we do not want to suggest that processes and outcomes cannot be related – that is certainly not the case. Our principal argument is that researchers and educators should be aware of the distinctions between these two concepts and remain sensitive to the differential effects of certain processes on learning outcomes (i.e., O'Reilly & McNamara, 2007).

By more carefully considering the distinction between processes and outcomes, and the relations between the two, researchers will be better able to detect and understand individual differences among students, and in turn, individual students' particular strengths and weaknesses. Knowing that one student scored 20% lower than another on a comprehension test tells us very little about *why* this particular student struggled or what individual differences may have contributed to comprehension difficulties. Considering both processes and outcomes, as well as nuanced differences in outcomes will move us toward more informed and useful assessment. For example, if we were able to analyze the performance *outcomes* from this struggling student's assessment score more closely and discern that this 20% difference was specifically driven by performance on, for example, deep comprehension questions (and these types of questions comprised some portion of the assessment), a more informed plan for individualized instruction might emerge. Similarly, assessing students' *processes* – perhaps through reading and questing answering times or think aloud protocols – would inform the areas in the text or specific time points where students engage in different cognitive processes, and where and how they may gain from instruction or scaffolding. While it is clearly useful and necessary to identify a student's overall levels of comprehension ability, summative assessments do little to inform instruction. If the ultimate goal is formative, to individualize and guide student instruction, greater attention must be turned to understanding the relations between processes and outcomes, and individual differences.

Processes vs. Tasks. Another distinction to be highlighted is between processes and the tasks in which students engage. Let's consider a few examples. First consider a task such as evaluation, considered to essentially be the epitome of higher order thinking within the Bloom Taxonomy. The process of evaluation can and should involve a great deal of higher order thinking. However, if the material to be evaluated is relatively simple, or the student is highly familiar with the material, the processes may merely comprise the use of knowledge (Kunen, Cohen, & Solman, 1966). For example, the student may be asked to evaluate the quality of an argument within an extended text. If the student has already been exposed to the text, as well as information about its quality, this turns into a memory task (i.e., recalling a prior evaluation). Such cases are unavoidable perhaps. Nonetheless, the familiarity of the materials plays an important role in what cognitive processes will be involved. Along the same line, the assessment used will also plan an important role. Many assessments may rely on multiple choice, where the student chooses the best evaluation (e.g., standardized tests such as the SAT). Such measures might tap into higher order processing (e.g., VanderVeen et al., 2007), but notably less so than open ended questions or think aloud (Magliano, Millis, Ozuru, & McNamara, 2007). On the surface, the Bloom Taxonomy makes good sense; clearly *evaluation* is better than *analysis*. Yet, in practice using the taxonomy to distinguish between the value of various educational activities or outcomes can be more challenging.

Consider further a simple task such as a student repeating a list of words aloud. According to the ICAP framework, a literal interpretation would lead to classifying this behavior as *active* because it involves an overt response. However, the overt observation of an individual repeating words often has little correspondence with the cognitive processes engaged: the learner may be passively and rotely repeating the word or by contrast, the individual might be using complex mnemonics. Hence there is little correlation between word repetition and memory (Craik & Watkins, 1973; McNamara & Scott, 2001). Of course there is some correspondence between the processes inherent to tasks and the underlying processes engaged by individuals. How could we infer cognition otherwise? However, there is a strong tendency to conflate one with the other, with too little consideration of what might be required of the individual for a particular task.

Another consideration regards the tendency to treat frameworks as developmentally hierarchical. For example, the use of the Bloom Taxonomy can lead to assumptions that mastery at lower levels is required prior to advancing to higher levels. An educator (or researcher) may therefore focus on fully developing a learner's lower level skill such as word decoding before tackling a skill that is expected to involve higher level skills, such as comprehension. This poses a problem because developing students can make advances at higher levels before mastering those that came before (Resnick, 1987; Zohar & Dori, 2003). And, some skills that seem hierarchical in nature may best develop in tandem, such as decoding and comprehension (e.g., Kendeou, van den Broek, White, & Lynch, 2009).

Moreover, outcomes are often conflated with processes, as we have already discussed. Thus, poor performance may be assumed to indicate an insufficiency in one skill, but actually be caused by a very different underlying problem (Rapp et al., 2007). A student who seems to have an undeveloped skill may actually be engaging in appropriate cognitive processes, but not showing observable evidence of that skill. Relatedly, students might not engage with tasks and questions as originally conceived by educators. Gierl (1997) examined the relationship between the cognitive processes students were intended to engage while solving math problems (as defined using Bloom's Taxonomy) and the cognitive processes students actually engaged (as measured through think aloud protocols). He found alignment only 54% of the time, with slightly higher agreement for students with higher math ability. This result highlights the concern that educational intentions will frequently mismatch students' cognitive engagement, even when a thoughtfully formulated framework is intelligently implemented.

With regard to the Bloom Taxonomy in particular, some researchers and educators have developed revised versions that aim to make the taxonomy more comprehensive and flexible. For example, the Bloom's Revised Taxonomy (Anderson et al., 2001), spearheaded by Lorin Anderson, a former student of Bloom's, divides the cognitive process level into two dimensions: the cognitive process dimension (remember, understand, apply, analyze, evaluate, create) and the knowledge dimension (factual, conceptual, procedural, metacognitive). By filling in the resulting two-dimensional taxonomy table with lessons and goals, educators can review the number of higher level categories their curriculum manages. Educators might then modify their plans in order to move from lower to higher levels. Likewise, in the area of reading, Afflerbach, Cho, and Kim (2011) proposed a metacognition level to the taxonomy, which they convincingly argue is crucial because students must be able to recognize comprehension errors and make adjustments to their reading strategies, in addition to possessing fundamental reading skills.

These enhanced frameworks can be quite useful as pedagogical tools when used appropriately. Appropriate usage cannot be taken as a given, however. Consider that North Carolina Public Schools' Common Core State and NC Essential Standards' home page invites educators to use the Revised Bloom's Taxonomy by providing the seemingly simple two-dimensional taxonomy table, and a brief explanation of the four types of knowledge and the six cognitive processes (http://www.ncpublicschools.org/acre/standards/). Educators (who are from North Carolina or who find the site through search engines) may not use the framework with nearly the full array of information provided by the authors. This is not to criticize North Carolina's Public Schools; they have consulted with Lorin Anderson in the development of their standards, and provide additional information throughout their website. Still, providing such a simplified version of a complex tool is a large-scale manifestation of the concern that the concept of higher level processing might be misinterpreted. A primary concern is that educators will mistake tasks or assessments designed to promote or measure higher level cognitive processes as definitively doing so.

Interactions among Processes, Tasks, and the Reader. Thus far, we have discussed issues regarding differentiating between lower and higher level processes as well as guarding against conflating processes, tasks, and outcomes. A layer on top of these concerns is the importance of individual differences and interactions between the various factors that influence comprehension.

Within the last few decades, a good deal of attention has turned toward the complex dependencies between factors such as the task, the context, the measures, and individual differences. One of the most pervasive problems in education and learning sciences is developing a more thorough understanding of how the same task can do very different things for different learners. There are any number of studies demonstrating such dependencies (*insert current volume reference? NB: I'll do this…).* For example, Voss and Silfies (1996) reported interactions between prior knowledge, reading ability, and text structure. McNamara, Kintsch, Butler, Songer, and Kintsch (1996) demonstrated a three-way interaction between prior knowledge, text cohesion, and the type of outcome measure (or level of understanding). These types of studies point toward the importance of recognizing that how an individual processes a certain text or task depends on the mental processes that may or may not be afforded by abilities, such as knowledge and reading skill (among others). Although frameworks designed to describe stages of processing provide a basic delineation of higher and lower level processes, this distinction often falls apart for individual readers in specific contexts. This is primarily attributable to a conflation of the *task* assigned to the student and the actual cognitive *processes* engaged by that student as well as the success of those processes. It is commonly assumed that all tasks or stimuli have the same effects on all learners, and serve to induce the same internal processes. However, this is clearly not the case.

Take, for example, the task of paraphrasing a text passage. For a young child, this task may be highly complex, requiring the child to engage in the activation of knowledge about words, syntax, as well as the domain of the passage. For an adult, on the other hand, paraphrasing may be much more of a passive task. Of course, this is further complicated by the properties of the text passage itself. Even for adults, paraphrasing can be a challenging *higher* *level* process if the text is difficult enough or if they do not have sufficient prior knowledge of the domain. Current frameworks tend to operate under the assumption that a learner's outward behavior defines their learning outcomes, which is simply not true. Therefore, these frameworks frequently fail to capture the differences that arise from the reader and the nature of the task.

Conclusions

Within the context of comprehension and education, there has been a heavy emphasis placed on an individual's ability to construct a coherent and elaborated mental representation of text content. To this end, research has aimed to establish the theoretical basis behind the comprehension process (e.g., Kintsch, 1998; Pressley & Afflerbach, 1995; Zwaan & Radvansky, 1998), as well as the most effective interventions for improving comprehension skills and strategies (McNamara, 2007; Pressley, 2000). A major problem lies in the task of operationally defining the component processes that contribute to the comprehension of text. A number of frameworks have been proposed to delineate the differences between lower and higher level cognitive processes (Bloom, 1956; Chi, 2009) and to map these various levels of thinking onto the task of reading comprehension (Afflerbach, Cho, & Kim, 2011; Paris, 2005). However, the task of mapping such frameworks on to the reading comprehension process proves difficult, given the influences of various domains as well as the stages of development (Afflerbach et al., 2011).

One particularly salient aspect of implementations of hierarchical frameworks, such as the Bloom Taxonomy or the ICAP framework, is that they are primarily based on overt behaviors, tasks, and performance assessments. This allows educators to directly observe these behaviors and then intervene when students are not engaging in the desired behaviors. It provides tractable goals for encouraging higher level learning, whereas it would likely be futile to monitor and modify students' *actual* cognitive processes during classroom activities. Researchers and educators should be cautious, however, when assuming that certain cognitive processes consistently underlie particular types of learning activities, as overt behaviors do not always reliably indicate the processes in which students are engaged. A student may engage in a task considered to be lower level, but engage in higher level processing, and vice versa, a student may engage in a higher level task, but engage in superficial processing. At the same time, a student who has engaged in higher level processing may manifest the benefits of this processing only at lower levels (which is likely the zone of proximal development for that student). These interactions between reader, task, and outcomes complicate simplistic interpretations of higher level processing.

Given the potential hazards of labeling cognitive processes or learning objectives as lower and higher level, one solution would be to throw out the terminology entirely. However, successful reading comprises a multitude of skills and strategies from the basic (letter recognition) to the complex (self-monitoring of comprehension), such that having hierarchically defined distinctions is clearly useful when building comprehension models (Kintsch & van Dijk, 1978), developing assessments (Magliano et al., 2011), and designing and evaluating classroom activities (Fisher & Hiebert, 1990). Hence, terminology related to higher order thinking is useful on various levels. Nonetheless, one of the primary goals of this chapter is to convince readers to avoid or temper the conflation between learning processes, tasks, and outcomes – particularly if the ultimate research objective is to understand the optimal conditions for enhancing student comprehension and learning. Ultimately, greater consideration must be turned to individual differences among students, and how students' abilities, goals, and dispositions differentially affect comprehension. Clearly the last few decades of research in the area of comprehension have elucidated a good deal in this respect (and much of that progress is described within this volume of work). But we can do better; there remain a multitude of questions to answer, particularly in regard to how to foster and how to scaffold students toward higher order thinking.

References

Afflerbach, P., Cho, B., & Kim, J. (2011). The assessment of higher order thinking skills in reading. In G. Schraw (Ed.) *Current perspectives on cognition, learning, and instruction: Assessment of higher order thinking skills*, (pp. 185-215). Omaha, NE: Information Age Publishing.

Anderson, L. W. (Ed.), Krathwohl, D.R. (Ed.), Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives (Complete edition).* New York: Longman.

Bloom, B. (1956). *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain.* New York: David McKay Co.

Cain, K., & Oakhill, J. (2006). Assessment matters: Issues in the measurement of reading comprehension. *British Journal of Educational Psychology*, *76*, 697-708.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1,* 73-105.

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes*, 25, 1-53.

Craik, F., & Watkins, M. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *12*, 599-607.

Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader–text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology*, *104*, 515–528.

Fisher, C. W., & Hiebert, E. H. (1990). Characteristics of tasks in two approaches to literacy instruction. *The Elementary School Journal*, *91*, 3-18.

Gierl, M. J. (1997). Comparing cognitive representations of test developers and students on a mathematics test with Bloom's Taxonomy. *The Journal of Educational Research*, *91*, 26-32.

Gilliam, S., Magliano, J. P., Millis, K. K., Levinstein, I., & Boonthum, C. (2007). Assessing the format of the presentation of text in developing a Reading Strategy Assessment Tool (RSAT). *Behavior Research Methods*, *39*, 199–204

Granello, D. H. (2001). Promoting cognitive complexity in graduate written work: UsingBloom's Taxonomy as a pedagogical tool to improve literature reviews. *Counselor Education and Supervision*, *40*, 292-307.

Hanna, W. (2007). The New Bloom's Taxonomy: Implications for music education. *Arts Education Policy Review*, *108*, 7–16.

Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. (2009). Predicting Reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101,* 765-778.

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

Kintsch, W., & van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological review*, *5*, 363-394.

Kolers, P. A., & Roediger, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, *23*, 425-449.

König P, Kühnberger K.U., & Kietzmann TC (2013). A unifying approach to high- and low-level cognition. In U. V. Gähde, S. Hartmann and J. H. Wolf (Eds.), *Models, simulations, and the reduction of complexity,* (pp. 117-139). Berlin, Germany: De Gruyter.

Krathwohl, D.R., Bloom, B.S., & Masia, B.B. (1964). *Taxonomy of Educational Objectives: Handbook 2: The Affective Domain.* New York: David McKay Co.

Kunen, S., Cohen, R., & Solman, R. (1966). A Levels-of-Processing Analysis of Bloom ' s Taxonomy, *Journal of Educational Psychology, 73,* 202–211.

Magliano, J. P., Millis, K. K., Levinstein, I., & Boonthum, C. (2011). Assessing comprehension during reading with the Reading Strategy Assessment Tool (RSAT). *Metacognition and Learning*, *6*, 131–154.

Magliano, J. P., Millis, K. K., Ozuru, Y., & McNamara, D.S. (2007). A multidimensional framework to evaluate reading assessment tools. In D.S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies,* (pp. 107-136). Mahwah, NJ: Erlbaum.

McNamara, D. S. (Ed.). (2007). *Reading comprehension strategies: Theory, interventions, and technologies*. Mahwah, NJ: Erlbaum.

McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix.* Cambridge: Cambridge University Press.

McNamara, D.S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes*, *22*, 247-288.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*, 1-43.

McNamara, D. S. & Magliano, J. P. (2009). Towards a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation,* (Vol. 51, pp. 297-384). New York, NY: Elsevier Science.

McNamara, D. S., O'Reilly, T. P., Best, R. M., & Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research*, *34*, 147–171.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, *29*, 10-17.

Millis, K. K., Simon, S., & tenBroek, N. S. (1998). Resource allocation during the rereading of scientific texts. *Memory & Cognition*, *26*, 232–246.

O'Reilly, T., & McNamara, D.S. (2007). Reversing the reverse cohesion effect: good texts can be better for strategic, high-knowledge readers. *Discourse Processes, 43,* 121-152.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, *40*, 184-202.

Pearson, P. D., & Hamm, D. N. (2005). The assessment of reading comprehension: A review of practices—Past, present, and future. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment*, (pp. 13–69). Mahwah, NJ: Lawrence Erlbaum Associates.

Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Volume III* (pp. 545-561). Mahwah NJ: Erlbaum.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Earlbaum.

Rapp, D. N., van den Broek, P., McMaster, K. L., Panayiota, K., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading*, *11*, 289–312.

Resnick, L. (1987). *Education and learning to think*. Washington, DC: National Academy Press.

Simpson, E. (1972). *The classification of educational objectives in the psychomotor domain: The psychomotor domain. Vol. 3.* Washington, DC: Gryphon House.

Van der Schoot, M., Horsley, T. M., & van Lieshout, E. C. D. M. (2010). The effects of instruction on situation model construction: An eye fixation study on text comprehension in primary school children. *Educational Psychology*, *30*, 817–835.

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

VanderVeen, A., Huff, K., Gierl, M., McNamara, D.S., Louwerse, M., & Graesser, A.C. (2007). Developing and validating instructionally relevant reading competency profiles measured by the critical reading sections of the SAT. In D.S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 137-172). Mahwah, NJ: Erlbaum.

Voss, J., & Silfies, L. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. *Cognition and Instruction*, *14*, 45-68.

Zohar, A., & Dori, Y. J. (2003). Higher order thinking skills and low-achieving students: Are they mutually exclusive? *Journal of Learning Sciences*, *12*, 145-181.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123,* 162-1855.