The Dissertation Committee for Derek Joseph Hanson Certifies that this is the approved version of the following Dissertation:

Understanding Metacomprehension: A Multidimensional Examination of Metacognitive Cues and Their Impact

Committee:

Veronica X. Yan, Supervisor

Marlone D. Henderson

Katherine M. Muenks

Diane L. Schallert

Understanding Metacomprehension: A Multidimensional Examination of Metacognitive Cues and Their Impact

by

Derek Joseph Hanson

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2023

Abstract

Understanding Metacomprehension: A Multidimensional Examination of Metacognitive Cues and Their Impact

Derek Joseph Hanson, Ph.D.
The University of Texas at Austin, 2023

Supervisor: Veronica X. Yan

Metacomprehension is a cornerstone in self-regulated learning, crucial for calibrating study strategies and resource allocation. Understanding the mechanisms through which learners form their metacomprehension judgments is therefore crucial for both academic success and lifelong learning. Contemporary metacomprehension cue research often relies on a single open-ended response, but this approach assumes that learners possess accurate insights. My research reveals that this assumption is problematic: learners consistently struggle to explicitly report using these cues without additional prompts. The primary objective of this dissertation is to develop and implement innovative methods for measuring the cues learners employ in making metacomprehension judgments. Synthesizing prior research, I created a set of cues for learners not only to report but also to rate their experienced valence as they studied complex learning materials (e.g., text-based essays and instructional videos).

My research offers four key contributions. First, it emphasizes the importance of measuring multiple cues over the traditional approach of focusing on a single cue. Second, it reveals that learners are often unaware of the cues affecting their judgments: cue valence

ratings were more predictive of their judgments than reported cue use. Third, it demonstrates that the cues that relate to comprehension judgments are not always the same as those that relate to learning. Specifically, cues that reflect the use of deeper strategies are more likely to be uniquely related to comprehension. What this implies is that there may be value in directing learners to focus specifically on the deeper cues that support comprehension. Fourth, it shows how individual differences in reading strategies and topic interest impact the experience of metacognitive cues.

Collectively, this series of studies not only introduces a new paradigm for exploring metacomprehension and the cues that inform these judgments but also investigates how patterns of cue use vary across individuals and are differentially associated with comprehension. The final chapter elaborates on these contributions, discussing their implications for future research and potential interventions.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

"if there is ever going to be a genuine breakthrough in the psychological study of memory . . . it will, among other things, relate the knowledge stored in an individual's memory to his knowledge of that knowledge."

Tulving & Madigan (1970, p. 477)

To navigate our complex world, learning is essential. We learn from many different types of sources: reading the news online, watching a video on our phone, participating in a conversation. In both formal schooling and beyond, the most common sources are visual, and text based. But what does learning involve? It's not just about ingesting and memorizing information so that it can be regurgitated. Rather, learning is based on understanding; we must construct a representation of the new knowledge, connecting the new information with previous understanding and inferences into mental models which we can use to understand the world (Kintsch & Rawson, 2005). In other words, learning involves not just taking in information passively, but requires active meaning-making throughout the learning process. Or, put another way, comprehension is often the target of instruction, rather than pure memory.

Moreover, learners often have to make sense of what they are learning in a self-regulated way. Even in a directed classroom, the learner is the person constructing their own mental representations; they themselves have to know when they are confused so that they can do something about it: focus attention, switch strategies, ask a teacher, and so on. This active construction of comprehension means that being able to accurately monitor one's own current state of knowledge (e.g., checking to see if learning goals have been met) is essential. Many intervention programs are designed to elicit monitoring and reflection in learners (Lipko & Dunlosky, 2007) but accurate monitoring is another challenge altogether (Wiley et al., 2016). In the present series of studies, I will explore the types of cues that people rely on to make metacomprehension judgments and the individual and contextual factors that impact cue use.

WHAT IS METACOMPREHENSION? HOW DOES IT DIFFER FROM COMPREHENSION?

Comprehension refers to the process of understanding something (Kintsch & Rawson, 2005), while metacomprehension involves monitoring and controlling one's own comprehension (Lipko & Dunlosky, 2007; Pitts, 1983). To comprehend something, an individual must construct knowledge by integrating and representing incoming information as a mental model. Concurrently, they engage in a secondary process that monitors the progress of their comprehension, assessing whether the information makes sense, identifying gaps in their mental model, and deciding on the next steps.

Effective monitoring is crucial because control decisions often rely on metacognitive judgments. Inaccurate judgments may lead learners to prematurely terminate their studies due to overconfidence, waste time studying familiar material because of underconfidence, or focus on superficial aspects rather than pursuing a deeper understanding. Metacognitive monitoring heavily depends on meaning-making processes and how learners integrate and represent incoming information in their mental models. In this series of studies, I specifically concentrate on metacomprehension monitoring.

Comprehension: Mental Models of Discourse

Comprehension, or the actual level of understanding, results from a complex series of processes undertaken during learning. While these processes are often highly automated in experienced readers, certain aspects remain directly controlled by the learner, such as the depth of material processing (Tapiero, 2007). The situation model (van Dijk & Kintsch, 1983; Kintsch, 1988, 1998) characterizes comprehension as a set of processes activated by readers as they engage with incoming information. As learners interact with the material, they begin to establish a mental model of the concepts and their relationships. To accomplish this, they access their background information to integrate new information with their prior knowledge (Myers & O'Brien, 1998). They may also need to refer to their knowledge of previous sentences to connect new information with older material (McNamara & Magliano, 2009). As they process more information, their memory undergoes additional activations. As layers accumulate and relationships become more complex, learners must continually update their mental model, resolving discrepancies, drawing analogies, and organizing new knowledge. When reading a passage or watching a lecture, the volume of information to be modeled is substantial, requiring a multi-layered mapping to develop global coherence. This necessitates a more intricate model than when parsing only a sentence or two (Tapiero, 2007).

Discourse psychologists view the processes involved in constructing a mental model as hierarchical and distinct (McNamara et al., 1996; Thiede et al., 2009). At the most basic level of representation, words are paired with their syntactic relations to form a surface representation, a necessary component for understanding but not believed to significantly affect the quality of comprehension (McNamara & Magliano, 2009). Further up, the textbase level contains a mapping of the propositions within the material, a connection of predicates and arguments. At the highest level of analysis, the situation model, the learner draws inferences to connect the propositions within the material to prior knowledge. Comprehension, at the end of processing the material, should map to the depth at which the learner processed the information; more active processing should create more connections outside of the direct text and within the text (situation), less active should limit comprehension to knowledge more limited to propositions within the text itself (textbase)

(McNamara & Magliano, 2009). A high-quality mental representation will integrate ideas from the materials with inferences derived from both the text and prior knowledge (Graesser et al., 1997).

Metacomprehension: Metacognitive Monitoring and Control

Having outlined the steps involved in comprehension, next is the question of how metacomprehension processes take place. What sets humans apart as more effective learners is our ability to be aware of our knowledge and its limitations, or as Tulving & Madigan (1970) described it, "one of the truly unique characteristics of human memory: its knowledge of its own knowledge." This awareness and understanding of our own thought processes is operationalized as metacognition. Any thoughts about one's thoughts, from self-awareness to monitoring memory or regulating behavior mentally, fall under the broad category of metacognitive. This general awareness of our cognition is the broadest category of metacognition depicted in Figure 1.

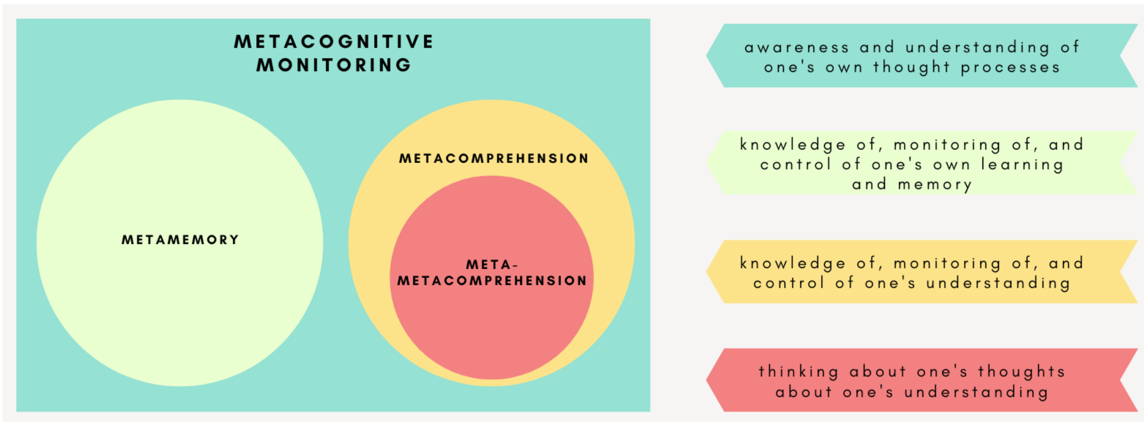
Within the realm of metacognition fall two forms of monitoring: metamemory and metacomprehension. Metamemory research examines the metacognitive judgments and decisions made when memorizing specific stimuli, at the smallest level of memory (Dunlosky & Bjork, 2014; Flavell & Wellman, 1975). When the goal of learning is to remember exact matches, or specific facts, a learner will rely on metamemory strategies to achieve this goal, such as looking for connections to the new stimuli within prior knowledge or generating easy-to-remember pairs (Soderstrom et al., 2015). However, when the goal of learning is to remember the meaning or gist, exact words from the passage or video are quickly lost or never even attended to directly, as the learner instead relies upon metacomprehension strategies to achieve the goal, such as developing a model of the text and combining the new information with previous understanding and inferences

(Kintsch & Rawson, 2005). When the goal of learning moves beyond pure memorization and into understanding broader ideas and concepts, we've moved into metacomprehension (Rawson et al., 2000). Importantly, metacomprehension is applicable to all forms of learning, from reading to watching and listening, despite the term's typical association with reading comprehension.

A subtype of metacomprehension is meta-metacomprehension. Having reflected upon our comprehension during metacomprehension, the next step is when we reflect upon our reflections of our learning. The questions asked of participants within this dissertation are a form of meta-metacomprehension. I am asking learners to reflect upon the cues they believe they utilized to evaluate their learning. So, in a sense, this research is about both metacomprehension cues (at the level of metacomprehension) and our ability to think about how we utilize those cues (meta-metacomprehension).

Figure 1.

Components of Metacognitive Monitoring



In this series of studies, I specifically examine metacomprehension monitoring and not metacomprehension control or metacomprehension strategies. The reason for this focus is that monitoring serves as the initial step in the process. The metacognitive cues that one experiences during the monitoring phase have an impact on the strategies and control mechanisms that are implemented in response. Rather than immediately diving into the metacognitive tools that directly affect outcomes, which may overlook the earlier stages that influenced them, I take a closer look at the monitoring stage. By delving deeper into the monitoring stage and analyzing the impact of cues on comprehension judgments, we can gain more insight into the earlier stages of metacognition that have an impact on the more commonly studied aspects further down the chain.

How Are Metacognitive Judgments Made? Cue-utilization Theory

As with many conceptions of memory and cognition, scholars debate the origin of metacognitive judgments, although two main schools of thought predominate. One view proposes that metacognitive judgments are derived from directly accessing the memory at issue. In the direct access theory, learners simply find the actual memory in storage, then estimate its strength (Dunlosky & Metcalfe, 2008; Koriat, 2012). This estimate of the strength of the memory is the judgment. This theory may explain judgments of specific items or singular facts, but judgments of comprehension, which require multiple memory traces and interrelationships, are likely too complex to be explained through direct access (Kornell, 2014; Schwartz & Metcalfe, 2017; Undorf et al., 2018).

The alternative perspective, which forms the basis of this paper, contends that metacognitive judgments are rooted in inferences drawn from metacognitive cues. According to the cue-utilization theory, learners form judgments based on their preexisting metacognitive knowledge, current metacognitive experiences, or a combination of both cue sources (Flavell, 1979; Koriat, 2012). Learners draw inferences from available cues such as familiarity, fluency, and ease of processing (Kornell, 2014). Intriguingly, some cues, like feelings, may be directly accessible, unlike memory (Kahneman, 2003; Kahneman & Frederick, 2005), yet are still considered inferential cues. Regardless of how cues are accessed (directly or through inference), it is evident that metacognitive judgments often incorporate more than just memory strength. For example, when memory strength is held constant while manipulating the visual (Kornell et al., 2011) or auditory (Rhodes & Castel, 2009) presentation of materials, metacognitive judgments change, indicating that these cues are integrated into metacognitive judgments.

When forming a metacognitive judgment, learners have access to both experience-driven factors and knowledge-based factors (sometimes called "theory-driven factors"), both of which have been shown to influence metacognitive judgments (Blake & Castel, 2018; Koriat, 1997; Yan et al., 2016). In some learning situations, metacognitive judgments seem to rely solely on direct experiences with the materials (e.g., feelings of fluency during reading), independent of implicit theories and beliefs about the learning task (Undorf & Zimdahl, 2019). On the other hand, learners assign higher metacognitive judgments of learning to certain tasks or items that cannot be explained by fluency or experience alone. For instance, learners believe they have learned animated words better than static words (Li et al., 2016), more concrete words better than abstract words (Witherby & Tauber, 2016), and more frequently occurring words better than less frequent ones (Jia et al., 2016).

Despite this dichotomy, it is likely that learners integrate both types of cues into their metacognitive judgments (Frank & Kuhlmann, 2017; Jia et al., 2016; Mendes et al., 2019). Moreover, when learning larger amounts of information (e.g., full sentences rather than single words or word pairs), metacognitive judgments become less reliant on the fluency of the stimuli, such as the disappearing effect of font size as the amount of information available increases (Luna et al., 2018). When more than one cue type (experience-based vs. knowledge-based) seems to affect a single judgment, it suggests that multiple cues may underlie metacognitive judgments. This is not surprising, given the variety of cues available when reading a passage, listening to a lecture, or watching an informative video.

Process Model for Metacomprehension Judgments

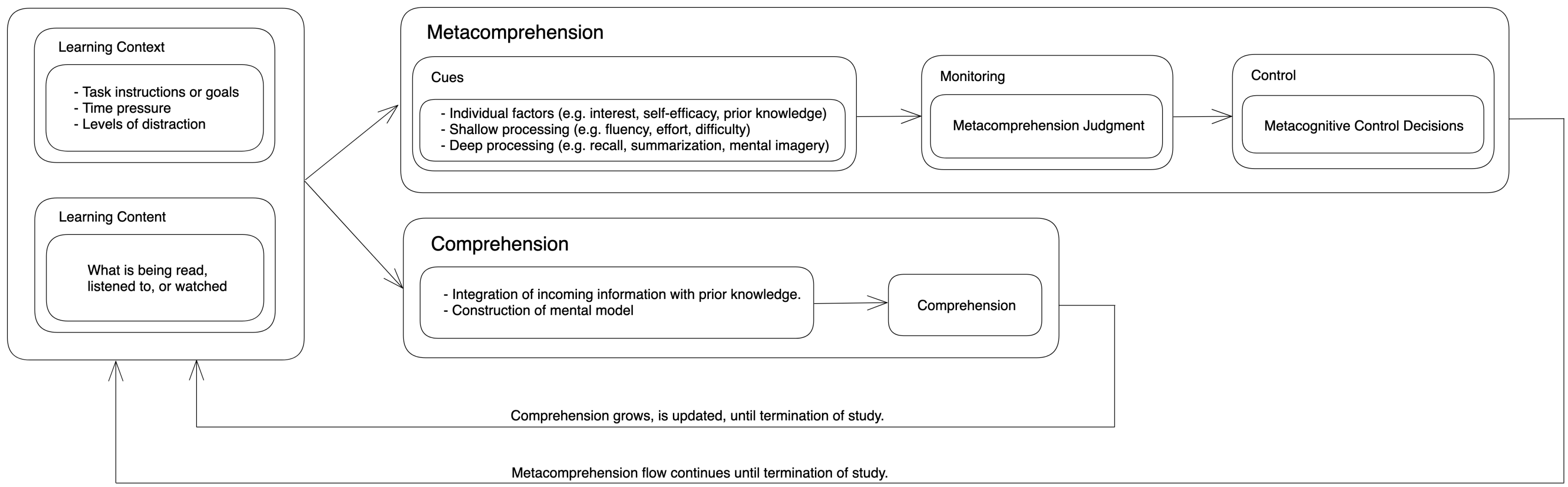
The theoretical process model presented in Figure 2 below illustrates how the comprehension and metacomprehension processes occur in parallel. Comprehension is created through integration of incoming information with prior knowledge and the construction of a mental model. Metacomprehension is influenced both by cues that arise from the comprehension process as well as cues originating from individual factors. Metacomprehension judgments in turn, directly affect metacognitive control decisions.

In this figure, I also depict the role of context; how incoming information—the learning content—that is processed is affected by contextual factors, including task instructions, task demands, time pressure, and levels of distraction. Context can play an integral role in the quality and quantity of learning, from the environment in which learning takes place (Tessmer & Richey, 1997) to the motivational context (Boekaerts, 2010). In this model, context adds an extra layer of influence to the perception of information. Although context is not directly connected to judgments and learning outcomes, it affects those components through metacognitive cues.

At the heart of this research, and at the heart of Figure 2, are the metacognitive cues. There are multiple types of cues. As we process incoming information, we interpret it with individual beliefs we bring into the situation. Cues such as our prior knowledge, interest, and self-efficacy influence how we perceive our learning, and our actual learning itself. In addition, how we process the information, or feel this processing is going, serve as important cues. Fluency cues (e.g., how easy it feels to construct our mental model) or deeper processing cues (e.g., whether we think we can explain and summarize the passage) are guideposts that we use to monitor our learning as we proceed through the task and reflect on it later.

Figure 2.

Process Model for Metacomprehension Monitoring



Metacomprehension judgments hold significant importance, as they result from a metacognitive monitoring phase and inform subsequent decisions (i.e., metacognitive control decisions). If a metacomprehension judgment indicates insufficient knowledge, it signals the need to continue studying or switch to a potentially more effective strategy. Conversely, if the judgment suggests adequate understanding, it prompts us to shift our attention elsewhere, allocating our limited cognitive and motivational resources to the next task (Dunlosky & Metcalfe, 2008; Nelson & Narens, 1990). Monitoring our learning in this way guides our decisions about what, how, and for how long to study (Metcalfe & Finn, 2008; Soderstrom & Bjork, 2014; Thiede & Dunlosky, 1999), potentially enhancing learning effectiveness (Son & Kornell, 2008; Thiede et al., 2003).

However, inaccurate monitoring can lead to poor control decisions, such as terminating study too early or dwelling on a concept for too long (Flavell et al., 1970; Rhodes, 2015). Mistakenly believing you have already mastered a subject when you have not may result in prematurely ending your study, only to experience an unpleasant surprise when you cannot recall the information when needed (e.g., on an exam). Conversely, mistakenly believing you have not learned something when you have can lead to excessive, unnecessary studying at the expense of allocating time to other important tasks.

As a result, examining how individuals monitor their understanding and how these judgments impact behavior constitutes an important area of research. Gaining insight into how people assess their comprehension and what influences these judgments can help us better understand the learning process and enhance both its efficiency and completeness.

METACOMPREHENSION: MOVING BEYOND METAMEMORY

The learning material may present learners with a limited number of cues or a plethora of cues to consider. Relatively simple tasks (e.g., learning foreign language vocabulary terms) provide fewer cues for learners to process when evaluating their learning. In contrast, more complex tasks like reading text passages, which involve identifying key points, retrieving related knowledge, and creating mental models of these relationships, offer a broader set of cues for evaluating one's learning (Wiley et al., 2005). With more cues available, this also potentially leads to greater variety in the use of cues for self-assessment and larger comprehension and metacomprehension gaps between skilled and unskilled readers (Wiley et al. 2016). The majority of the research on metacognitive cues from cognitive psychology literature, however, has focused on paired associates learning; metacognitive cues arising from more complex educational materials are less well understood. Nonetheless, researchers are beginning to explore questions such as whether and when students integrate multiple cues in metacomprehension judgments (Undorf et al., 2018). This dissertation represents one such inquiry into these questions.

When reading a text passage, learners process information metacognitively, engaging with the material through a continuous exchange of thoughts, processes, and actions (Flavell, 1979; Brown, 1977). The act of reading involves more than merely interpreting words on a page for meaning; learners constantly monitor and make decisions as they read the text (Bjork et al., 2013). To evaluate their learning process metacognitively, learners must integrate cues brought into the reading task (prior cues) with those arising from their experience with the text itself (Koriat, 1997). To perform this integration, learners must first develop a mental model or representation from the text (Kintsch, 1998).

Readers start representing the text mentally by combining individual words to comprehend the meaning conveyed by the text (textbase). This basic memory for the facts and details of the text occurs even during the most cursory review. A reader may choose to stop processing at this level, satisfied with a textbase-level understanding (Kintsch &

Rawson, 2005). However, halting processing at this stage often leaves students with memory for the text (recall) but inadequate comprehension of the key concepts and their interconnections. To fully understand a text, readers must engage in deeper processing, developing a situation model by integrating information from the text with prior knowledge (Kintsch & Rawson, 2005). Scholars suggest that inference-level knowledge, the ability to draw connections across topics and to areas of knowledge outside the text, requires processing a text at the situation model level, the deepest level of comprehension (Thiede et al., 2010).

The depth of text processing chosen by a reader is likely influenced by the metacognitive cues brought into the situation, such as interest in the material and prior knowledge, along with perceptions developed while processing the text itself. Factors such as the ease or difficulty of reading, the familiarity of the information, the text's length, and thoughts about the efficacy or success of interpretation (Jaeger, 2012) all play a role in shaping these perceptions. Regardless of how cues are integrated to form the monitoring aspects of metacomprehension, the cognitive thoughts and processes serve as crucial cues to guide learners' understanding (Nelson & Narens, 1990; Dunlosky & Lipko, 2007).

A Framework for Metacomprehension Cue Typology

As the evidence above suggests, we rely on a variety of cues to make our metacognitive judgments. These cues can be broadly categorized into two types: cues that learners bring into a learning situation with them (e.g., beliefs about oneself as a learner, prior knowledge about the content domain) and cues arising from the learning situation itself (e.g., feelings of ease or difficulty, memory for the text). The latter type of cue—situational—might be further subdivided as I shall further describe in my review of past frameworks.

Griffin et al. (2009) demonstrated that learners use both prior and situational types of cues when making judgments of comprehension. Participants read passages about baseball-related topics, made judgments of comprehension estimating performance on related questions, then answered inference-based questions. Holding prior knowledge of baseball constant, those who read the texts were better at estimating their learning than those who answered the questions without having read the texts. Thus, learners appear to integrate cues from the act of reading (situational) in conjunction with prior knowledge to develop judgments of comprehension (Griffin et al., 2009).

The literature supports organizing cues under the categories of prior or situational. Table 1 organizes cue types identified in existing research into this typology. Prior cues, though not always discussed in applied metacomprehension research, include three types: prior knowledge, beliefs about abilities, and interest. Situational cues are a broader category, typically encompassing cues related to relatively superficial cues (fluency and ease of processing, effort, and difficulty) as well as relatively deeper cues (memory for text details, ability to understand or explain the text).

Table 1.

Metacognitive Cue Typology in Research Literature

Article	Prior cues	Situational cues
Griffin, Jee, & Wiley (2009)	**Heuristic** (e.g., interest, prior knowledge, beliefs about abilities)	**Representation-based** (e.g., accessibility of text representation, perceived ease of explanation)
Thiede, Griffin, Wiley, & Anderson (2010)	**Superficial** (e.g., familiarity and interest in the topic)	**Memory-based** (e.g., accessibility of text representation) & **Comprehension-based** (e.g., ability to self-explain)
Jaeger (2012)	**Heuristic** (e.g., interest, prior knowledge)	**Representation-based** (e.g., memory for text details) **Surface** (e.g, length of text) **Comprehension-based** (e.g, ability to understand or explain the text)
Jaeger & Wiley (2014)	-	**Non-comprehension-based** (e.g., memory for the text, interest, or prior knowledge of topic) **Comprehension-based** (e.g., ability to understand, summarize the text)
Wiley, Jaeger, Taylor, & Griffin (2018)	-	**Superficial** (length of text, vocabulary familiarity) **Structure-model** (ability to summarize the text)

Note: Bold words are terms used by the respective authors.

This dissertation aims to unify this typology by focusing on the source of the cues, rather than their perceived utility. As shown in Table 2, this classification method is more straightforward and consistent since the cue source remains constant.

Table 2.

Proposed Metacognitive Cue Typology

Category	Subcategory	Cue	Item
Prior	-	Self-efficacy	My overall confidence in learning new material from reading.
		Prior Knowledge	My knowledge of the topic from previous coursework or life experience.
		Interest	My interest in the topic of the text.
Situational	Shallow	Difficulty	How difficult or easy the text was to understand.
		Effort	The amount of effort I put into trying to learn the material.
	Deep	Recall	My ability to recall information from the text, such as facts or details.
		Mental Imagery	The amount/quality of mental images or examples I came up with to understand the text.
		Summarization	My ability to remember concepts or summarize/explain what the text was about.

Prior Cues: Prior Knowledge, Interest, and Self-efficacy

Metacognitive cues stemming from prior experiences significantly impact learners' comprehension and judgment processes. These heuristic cues, established before the start of the specific learning experience, consistently shape each subsequent learning event by directing learners' expectations, strategies, and evaluations of their understanding.

Researchers have identified three such ‘prior’ cues, so named because they were developed before engaging in a specific learning experience. These cues were identified through open response reports in which learners described the factors they believed influenced their metacognitive judgments. The three cues include: prior knowledge (a learner's pre-existing knowledge of a topic), self-efficacy (a reader's beliefs about their ability to construct understanding from various mediums, such as text or video), and interest (specifically, interest in the topic addressed in the learning task).

Prior knowledge offers a vast array of potential cues that expand as our learning experiences grow. As reading, listening, and watching are nearly unavoidable activities for students, they are likely to develop various beliefs about what aids their learning and to what extent (Schraw & Bruning, 1996). Using reading as an example, learners bring several forms of prior beliefs to a reading experience, including beliefs about their reading ability (Shell et al., 1995) and how the text's structure will impact their learning (Zwaan, 1994). For instance, readers adjust their processing style based on text expectations, focusing more on surface details in literary texts and more on situational details in news texts (Zwaan, 1994). While numerous potential prior beliefs may exist within a given reader, not all of them are useful for determining learning from a specific text, nor are all relevant cues likely activated simultaneously (Flavell, 1979).

In addition to beliefs about general materials or practices, learners bring prior beliefs about themselves into a learning situation (Schraw & Bruning, 1996). Beliefs about the self, such as one’s self-efficacy as a learner, are interpreted through the context of the materials. Am I good at science? How much do I know about electricity? Learners ask themselves questions such as these when confronted by a learning task, and the answers to these general questions bring forth prior beliefs about oneself as a learner and are relevant to the development of estimates of learning, or how much learning one is capable of. Some level of predetermined beliefs about the learner's confidence in learning material in a certain content area (science), by a certain method (text), in a certain situation (classroom), tested in a certain format (multiple choice), combine to influence the resulting metacognitive judgments. These prior beliefs, including self-theories of memory and retention (Kornell & Bjork, 2009), self-efficacy (Bandura et al., 1996; Zimmerman, 2000), and the level of prior knowledge in the domain (Maki & Serra, 1992) have been shown to influence learning estimates, though the relation of these beliefs to metacomprehension judgments remains relatively understudied.

Situational Cues: Metacognitive Experiences of Learning

Many of the cues that learners may choose to attend to relate to the direct experience of engaging with a particular learning task, not general beliefs. Metacognitive experiences are current conscious cognitive or affective experiences that provide cues about how the learning process is going (Flavell, 1979). These experiences activate strategies to achieve learning goals/drive behavior and can impact metacognitive knowledge by adding, deleting, or revising one's beliefs (Flavell, 1979). Feelings of fluency while reading a text, or sensations of lacking understanding during a lecture, are examples of explicit metacognitive experiences (Koriat, 2000). Feelings are the key driver of metacognitive experiences, such as the ease with which information is encoded or retrieved during learning (Koriat & Ma'ayan, 2005) or how easily answers come to mind when searching memory (Koriat, 1993, 1995). Typical metacognitive experiences can include feeling lost in a lecture, feeling a general level of understanding after reviewing notes, or recognizing that a problem feels easy and familiar as you work through it. Flavell (1979) described them as such:

> "Many metacognitive experiences have to do with where you are in an enterprise and what sort of progress you are making or are likely to make: You believe/feel that you have almost memorized those instructions, are not adequately communicating how you feel to your friend, are suddenly stymied in your attempt to understand something you are reading, have just begun to solve what you sense will be an easy problem, and so forth."

To make the distinction clear, prior metacognitive cues originate from an individual's past experiences and knowledge, whereas situational cues pertain to the specific context or environment surrounding a task or learning activity. Like prior cues, research involving open-ended reports from learners has identified five key situational cues. These five cues are: effort (the sensation of exertion invested in the learning task), difficulty (the perception of ease or challenge in comprehension), recall (the recollection of facts and details from the lesson), mental imagery (the quantity or quality of mental images/examples formulated to assist understanding), and summarization (the perceived ability to explain or summarize the lesson).

Situational cues, unlike prior cues, are typically classified into two subcategories: shallow and deep. However, the boundaries between cues defined as deep or shallow are not always clean. For an individual learner in a specific environment, a situational cue may reflect shallow processing, such as using recall of facts and details as a cue when a text was only lightly skimmed. However, had that same learner spent more time processing and building a more thorough mental model, then the same situational cue, recall, may have instead reflected deep processing.

Mirroring the hierarchy of the situation model, a model of reading comprehension in this area of research (Kintsch & Rawson, 2005), has researchers rate cues as more or less predictive of learning quality based upon the implied level of cognitive processing they reflect (Griffin et al., 2009; Thiede et al., 2009; Wiley et al., 2005). If deeper processing of information leads to better comprehension, then metacognitive cues which reflect greater cognitive effort should be more predictive of learning than others. Cues reflective of a situation model level of understanding should have the strongest positive association with learning, such as feeling as if one could easily summarize or remember inferences from the material. Cues reflective of more simplistic mental representations of the text (a textbase model), such as feeling capable of recalling facts or details, should also be positively associated with learning, but only when questions are tied to facts or details, not inferences.

Cues that relate to the materials, but not to the mental model, such as feelings of interest and prior knowledge, may not relate to comprehension and should not have a strong association with learning in either direction (Sarmento, 2018). One may be interested in a topic but not be able to understand the material, or one may have a high level of prior knowledge filled with misconceptions; both would likely increase confidence in one's understanding but associate negatively with learning. Similarly, heuristic feelings, such as perceiving difficulty understanding the material, may arise because it was difficult and little was comprehended, but could also be a metacognitive trap (Jaeger & Wiley, 2015).

In the literature the demarcation is often clear; mental imagery is a cue tied to deep processing; perception of difficulty is tied to shallow processing. But the reality is not as black and white. For the purposes of this paper, cues will be organized by the standard literature hierarchy, but you will notice deviations from these clean organizational structures in the results of several experiments. Keep in mind the flexible nature of these cues to a specific situation and, for now, focus on whether a cue generates from prior experience or the situation itself.

Shallow Processing Cues

When processing text passages, readers often rely on cues such as perceived text length, difficulty, fluency, or ease of processing the information to form judgments of comprehension (Jaeger & Wiley, 2015; Rawson & Dunlosky, 2002; Thiede et al., 2010). Although ease of processing provides a sense of understanding, it may not accurately predict successful learning (Rhodes & Castel, 2008). Shallow cues don't reflect deep engagement with the text, and depending on them can create a "metacognitive illusion" of understanding (Jaeger & Wiley, 2015). However, it is not surprising that learners often use ease of processing heuristics to estimate learning and make decisions (Rawson & Dunlosky, 2002), as research shows a preference for fluent information and its association with enhanced judgments of truth, likeability, and trust (Alter & Oppenheimer, 2009).

Learners may rely on simpler comparisons or heuristics when forming judgments of comprehension (Goldstein & Gigerenzer, 2002; Hertwig et al., 2008). Feelings of familiarity with the information or fluency in retrieving simple details can influence judgment (Glenberg et al., 1987; Glenberg & Epstein, 1987; Benjamin et al., 1998; Morris, 1990). However, these cues are not reliable indicators of comprehension. Remembering a passage's gist or being vaguely familiar with a subject may help answer basic questions but does not reflect a comprehensive understanding of the topic's nuances and connections.

Additionally, learners may consider the perceived ease or difficulty of processing information (Rawson & Dunlosky, 2002). The idea of "easily learned, easily remembered" often guides our judgments. We tend to believe that easily processed information is well-learned (Begg et al., 1989; Kornell et al., 2011), and this belief persists even when the ease of processing doesn't relate to actual learning (Rawson & Dunlosky, 2002). Factors like faster encoding in memory (Hertzog et al., 2003), quicker recall (Benjamin et al., 1998), familiarity (Koriat & Levy-Sadot, 2001), or presentation in larger, brighter, louder, or more dynamic formats (Rhodes & Castel, 2008; Busey et al., 2000; Rhodes & Castel, 2009; Li et al., 2016) can influence this belief. Despite having minimal impact on actual memory, relying on processing cues can lead to overconfidence in learning for easily processed information (Rhodes & Castel, 2009), which may result in stopping study efforts prematurely and lowered exam performance or delayed acquisition of complex knowledge (Dunlosky & Rawson, 2012).

Deep Processing Cues

When students engage deeply with a text, they interact with the material in a more effortful manner, generating different thoughts compared to shallow processing (Thomas & McDaniel, 2007). Metacognitive cues arising from deep engagement include self-explanation of content, visualization of concepts or examples, evaluation of the strengths and weaknesses of the arguments, application of information to new situations, or assessing one's ability to explain the material to someone else. These cues stem from constructing mental models of the information and its implications, rather than general feelings about the reading experience (Wiley et al., 2018).

Cues related to deeper mental model construction, such as summarization, are associated with higher levels of learning and increased metacognitive accuracy (Maki et al., 1990). When experimentally prompted to develop deeper models of discourse through elaborative processes, these actions reduce illusions of knowing. Accuracy improves when explaining the meaning and relevance of each section (Griffin et al., 2008), generating keywords that capture the discourse's essence (de Bruin et al., 2011), mapping concepts within the discourse (Redford et al., 2012), or creating a physical summary (Thiede et al., 2005; Thiede & Anderson, 2003). However, while learners in elaborative conditions likely relied on more informative cues, these experiments cannot conclusively prove that deeper cues were referenced to make more accurate judgments.

Using a deeper processing cue, such as the perceived ability to summarize material, can be informative even if the material was not learned well (Thiede & Anderson, 2003).

Attempting to create a summary or simulating one in our minds forces us to retrieve relevant facts, concepts, inferences, and activate prior knowledge. If the material was not well-learned, this act should reveal the flaws in our learning, leading to a lower comprehension estimate. Conversely, if drawing connections and paraphrasing feels easy, our confidence in understanding should increase accordingly (Thiede & Anderson, 2003). Unfortunately, not all cues are as informative as those associated with deeper processing of the discourse.

TYPICAL RESEARCH PARADIGMS

Cue Manipulation of Non-Text Materials

In the laboratory, metacognitive research overwhelmingly utilizes simple methods and materials, such as word pairs or images. Historically, judgments of learning have been solicited in the lab from word lists in which features such as font size are manipulated (Rhodes & Castel, 2008) or paired associates that differ in terms of their relationships, such as varying concrete versus abstract descriptions (Arbuckle & Cuddy, 1969, Hertzog et al., 2003). Thus, most of the research on learners' ability to predict their future learning incorporates learning materials less complex than a simple sentence. The benefit of this simplicity is the ability to control for features of the material or environment, allowing for identification of causal links. Rhodes and Castel (2008), for example, randomly presented words in either large or small font size as a way of manipulating ease of processing. However, these paradigms (a) tend to manipulate just one or two cues and (b) do not generalize to many forms of learning, including learning from texts and videos.

Undorf et al. (2018), provides a rare example in which more than two cues are simultaneously manipulated and hence their influence on metacognitive judgments is measured. They present a series of four experiments, manipulating the form of available cues, their prominence, and the number available to learners. The results showed that the majority of the participants integrated two or more cues into their metacognitive judgments. Interestingly, these cues often encompassed multiple categories, such as intrinsic (e.g., concreteness, emotionality) and situational (e.g., font size) cues, indicating the capacity to integrate disparate cues into a singular judgment.

Nevertheless, the current body of literature has not thoroughly investigated the relationships among the most prevalent metacomprehension cues. When reading the same passage, without changes to font or the availability of imagery, how do learners integrate multiple cues into a single judgment? Are there common patterns in reporting these cues?

Measuring Metacomprehension Cues Via Self-Report

Most of the metacognitive cue use literature has not focused on texts or more complex stimuli. Word lists and paired associates are not complex enough to require tracking multiple concepts or processes, naturally limiting the cues potentially available to learners. Without enough information to be integrated, learners cannot develop a textbase model, leaving only shallow processing cues and prior knowledge upon which to estimate their learning. Such materials limit metacognitive inferences to local judgments on single topics or items. Global judgments, such as estimates of comprehension, require materials more complex than simple word pairs (Thiede et al., 2009).

There have been efforts to examine metacomprehension with more complex text passages, including entire book chapters (Vössing & Stamov-Roßnagel, 2016). Focusing mainly on expository texts, with the goal of having readers learn material they do not already understand, these longer texts allow learners the opportunity to develop detailed situation models to aid their comprehension (Eakin & Moss, 2018). Using expository texts, however, the typical paradigm has tended to focus on asking learners to report the cues they use. However, these paradigms analyze data by focusing either on just a single cue (Jaeger, 2012) or simplistic classification (Wiley et al., 2018). There is an interesting paradox here: Although scholars acknowledge that individuals likely rely on multiple cues and sometimes prompt them to list these cues, metacomprehension analyses have mainly focused on a single cue. For instance, Jaeger (2012) requested participants to list the cues utilized but proceeded to analyze only the “least sophisticated” cue. In other words, their analyses treated metacomprehension judgments as if they were based only on one cue.

Furthermore, studies have only measured cue usage through a single open-ended question. With such a measure it is not possible to quantify the use of multiple cues accurately, if more than one is endorsed by a learner. Did a student use each cue equally, or were some of the cues reported weighted higher in their metacognitive judgment? To understand which cues learners think they use, and estimate the relative importance of each, better measurement techniques are needed.

Investigating the integration of multiple cues and patterns in cue reporting holds significant importance for various reasons. Gaining a deeper understanding of the relationships among different cues can unveil how learners synthesize and prioritize diverse sources of information while making metacomprehension judgments. This knowledge can contribute to the development of models that more accurately represent the cognitive processes involved in metacomprehension. Furthermore, exploring patterns in cue reporting, particularly any commonalities, can illuminate individual differences in cue usage among learners, thus revealing the preferred cues or combinations of cues that individuals depend on during metacomprehension assessments. Among other potential benefits, recognizing these individual differences can foster the creation of personalized learning strategies, specifically tailored to each learner's unique metacognitive profile.

Relationship of Cues to Metacognitive Judgments of Comprehension

The form of cues learners attend to can directly influence their metacognitive judgment, raising or lowering how much learning a student believes took place from reading a text passage. Begg, et al., (1989) found that learners believe some contexts are more likely to aid their learning than others, boosting their metacognitive judgments when such cues are present. When presented with static imagery or reading basic instructions, learners estimated lower judgments of learning than those presented with cues that feel more beneficial to learning, such as interactive imagery (Begg, et al., 1989). Similarly, learners believe that the addition of charts or analogies will enhance their ability to recall information at a later time (Jaeger & Wiley, 2015), giving higher judgments of learning than readers of plain texts.

Thinking of these contexts and beliefs as cues for evaluating comprehension, it seems plausible that reported metacomprehension cues may be associated with higher or lower judgments of comprehension. However, when metacomprehension research has tackled this issue, results have focused on a single cue in a controlled environment (Jaeger & Wiley, 2015). When multiple cues are available, do all cues significantly influence a judgment of comprehension? Do some cues tend to influence judgments more in one direction than the other (boosting or lowering estimated comprehension)? In this multi-cue environment, does controlling for the other available cues change these results?

Relationship of Cue Use to Learning

Irrespective of the cues utilized to formulate an estimate, individuals are generally poor at accurately estimating their comprehension (Dunlosky & Lipko, 2007), particularly when assessing learning from more than just a few sentences (Maki, 1998). Metacognitive accuracy, the ability to correctly judge one's learning, whether you have learned much, or very little, is necessary to wisely direct study actions to appropriately deepen one's learning. Some of this miscalibration can be explained by inherent biases in learners or the actual task/item difficulty level (Schraw & Roedel, 1994). For example, when making an estimate near-in-time to learning the information, current memories are presumed to be available later (Kornell & Bjork, 2009). Operating under the heuristic of "if I know it now, I'll probably know it later", is dangerous because of the forgetting curve; human memory is subject to decay over time (Soderstrom & Bjork, 2014). Even if a learner understands the material immediately after reading it, they might forget important details or concepts over time, particularly if they don't revisit or reinforce the knowledge. Short-term memory feels like a useful cue, creeps into judgments of comprehension and increases estimates when, in reality, forgetting will certainly take place before a later exam occurs (Thiede et al., 2003). Similarly, when learning something new, people tend to overestimate their ability, to the detriment of metacognitive accuracy (Kruger & Dunning, 1999).

A considerable portion of the discrepancy between actual comprehension and estimated comprehension can be attributed to the metacognitive cues on which learners rely (Thiede et al., 2010). Learners rely on metacognitive cues because, through experience, these cues have been associated with correct answers in the past (Koriat, 2012). Nonetheless, not all metacognitive cues are equally informative across all situations. If a discrepancy exists between the demands of a task and the cue used to estimate comprehension, miscalibration is likely (Thiede et al., 2010).

For instance, some cues may be more effective for simpler texts but less informative for complex or dense material. Relying on reading fluency might suffice for straightforward texts but not for materials necessitating deep processing and critical thinking. Additionally, certain cues may be informative in specific contexts but not others. For example, a cue related to prior knowledge could be beneficial when learning new information in a familiar domain but less effective when encountering an unfamiliar subject. Thus, the accuracy of a judgment of comprehension relies heavily on the cue or cues that were used to inform it.

Overview of the Dissertation

Identifying the cues learners rely on to make estimates of their understanding from more complex materials, as well as how these cues are distributed across different learners and contexts, can provide key insight into complex metacognitive processes. If we are to develop methods to improve metacognitive ability in students to promote effective learning, we must first understand the thoughts, insights, and cues that learners use to make metacognitive judgments.

Contemporary research into metacomprehension cues utilizes a single open-ended response, asking students to report the metacognitive cues utilized to develop a judgment of comprehension. Such a method implies that learners are capable of accurate meta-metacomprehension; the ability to monitor their monitoring. Relying upon open response as the sole measure may fail to capture cues learners would otherwise report. Further, as a text-based measure, without numeric responses along a continuum, this method does not capture the level at which the cue was experienced (high/low) and the direction (difficult/easy, positive/negative).

The studies that follow were designed to examine the cues that learners use to estimate their comprehension, from obtaining a basic understanding of the nature of these cues and their frequency to the influences of external forces such as individual differences and environmental context. In this exploration, I develop and deploy new methods of measuring reported metacomprehension cues, investigating these cues at greater depths than previously pursued in prior research. The dissertation studies herein address the following research questions:

> **Research Question 1.** What cues do learners report relying on to make their judgements of comprehension, for simple texts (Chapter 2), and complex materials such as essays and videos (Chapter 3)?
>
> **Research Question 2**. Does cue valence predict judgments of comprehension over and above cue use (Chapter 3)?
>
> **Research Question 3.** What impact do individual differences (interest, reading strategies) have on how cue valence is related to judgments of comprehension (Chapter 3)?
>
> **Research Question 4.** Does shifting the goal of reading affect cue use, judgments of comprehension, and learning (Chapter 4)?

Specifically, in Chapter 2, I report on Study 1 in which I investigate the cues people report using to evaluate their understanding and how these cues are distributed across learners. With simple text-based stimuli, I explore the form, frequency, and interrelationships of these cues which learners report using to estimate their comprehension (metacomprehension). This study builds on prior metacomprehension studies by acknowledging and allowing participants to report the use of multiple cues.

The two experiments in Chapter 3 build on the investigation of multiple cues that influence judgments of comprehension. In these studies, participants interact with longer, more complex materials (text-based essays and instructional videos) and new cue measurement methods are introduced. In addition to raw reporting of metacomprehension cues, Studies 2 and 3 ask participants to report the level and direction of each potential metacognitive cue. By asking them the valence of these potential cues, I examine whether cue valences predict participants' metacomprehension judgments over and above their cue use reports. To preview the findings, results show that cue valence is more reliable of a predictor than cue report—the results suggest that even though the cues affect metacomprehension judgments, participants are not always aware of this effect.

Chapter 3 also explores the impact of individual differences on metacomprehension judgments, by measuring individuals' general reading strategies (Study 2) and experimentally manipulating interest level (Study 3).

Finally, in Chapter 4, I report on Studies 4 and 5 which investigate how the context of the learning environment influences the cues reported by the learners. These experiments contain the same materials (text-based essay) and measures as found in Chapter 3, but here the goal for learning has been manipulated between learners. Thus, in Chapter 4, I continue to investigate the interrelationships between cue valence, and judgment of comprehension, but also include analysis of group differences. Further, in Chapter 4, participants completed a comprehension exam; therefore, analyses of the interrelationships between cues, cue valence, condition, and learning outcomes are explored, as well.

CHAPTER 2: CUE UTILIZATION IN METACOMPREHENSION JUDGMENTS

When judging how well we understand something, we draw upon our experiences. Some of these experiences are from the past (e.g., prior knowledge saved in memory), others are from the current learning experience (e.g., feelings of difficulty or ease, memory of the details). In some fashion, we draw upon these cues metacognitively, considering and weighing their utility in helping us judge how well we understand what we just learned. Although a judgment of our understanding is often distilled into a singular estimate, it could be formed from a single, overwhelming cue, or an amalgamation of multiple cues. Perhaps we use only our perceived ability to explain the information in a text to estimate our understanding. Or perhaps this cue is joined with our perception of how interesting the text was to form a weighted estimate where each cue contributes some proportional amount. While researchers have been interested in our ability to estimate our comprehension for decades, surprisingly little is known about the metacognitive cues that go into these estimates.

Often, learners report multiple cues leading to a judgment of their understanding of a text (Jaeger, 2012) and theorists believe learners commonly integrate more than one cue into their judgments (Koriat, 1997). However, much of the research in this area of metacognition has used methods that rely on a single cue as the sole input to an estimate of comprehension. This existing research presents a contradiction: Although researchers acknowledge that participants probably rely on multiple cues and have sometimes asked participants to list the multiple cues they have used, analyses within metacomprehension research have focused on just a single cue.

For example, Jaeger (2012) asked participants to list the cues used but then analyzed only the “most sophisticated” cue. Thiede et al. (2010) asked the same question but analyzed only the "least sophisticated" cue. In other words, their analyses treated metacomprehension judgments as if they were based only on one cue. Furthermore, existing studies have only measured cue usage through a single open-ended question. If more than one cue is endorsed by a learner, it is not possible to quantify the use of multiple cues accurately with such a measure. Did a student use each cue equally, or were some of the cues reported weighted higher in their metacognitive judgment?

Because of these measurement issues, drawing conclusions about the cues that learners think they use from existing research is problematic. Researchers rarely report exact proportions of the cues used and do not analyze the associations of the cues with one another. Thus, it is unknown which cues frequently associate with each other, or whether 'profiles' of cues exist between readers. To understand which cues learners think they use, and estimate the relative importance of each, better measurement techniques are needed.

In the following study, I aim to delve deeper into how individuals employ different cues to make their metacognitive judgments of comprehension. I will examine cue usage in two primary ways: by determining whether a cue was reported (a binary approach) and through qualitative analysis of responses to an open-ended cue prompt. A key aspect of my analysis will be investigating potential associations between different cues to see if distinct 'cue usage profiles' emerge, a concept previously suggested by researchers like Bröder & Undorf (2019). Additionally, I will explore the relationships between cues, judgments of comprehension, and performance on a multiple-choice comprehension test.

RESEARCH QUESTIONS

What cues do learners report relying on to make their judgments of comprehension for simple texts? This overarching research question will be explored in-depth through the following four subquestions:

1. Which cues are most frequently reported by learners, and what relative importance is attributed to each of these cues?

2. Can distinct 'profiles' of cue usage be identified among learners?

3. How does the self-reported utilization of these cues correlate with learners' judgments of comprehension?

4. What is the connection between specific cues and learners' performance on subsequent multiple-choice tests?

Study 1

METHOD

Participants

Recruited from a wide subject pool across a variety of majors (27.87% STEM, 22.54% Social Sciences, Humanities 12.30%) 246 undergraduate participants completed the study in exchange for course credit. Data for two participants were excluded because they acknowledged they did not read the materials, producing a final sample of 244. A majority female (60.25% sample, 54.4% student body), and white (43.03% sample, 38.9% student body), the sample roughly resembled the student body demographics of the large public university at which it took place (Latinx: 19.26% sample, 23.4% student body; Asian: 29.92% sample, 20.2% student body; Black: 5.33% sample, 5.3% student body). The age of participants reflected a typical university environment, as well (20.73 *M*, 1.99 *SD*).

Procedure

The basic task in this experiment involved learning new concepts by reading a series of short text passages (two sentences each) followed by reflecting on how well the information was learned and the metacognitive cues that informed this estimate. To begin, participants were instructed that their task was to learn information in preparation for a comprehension exam at the end of the experiment. The learning materials were then presented for students to read, without a specified time limit. Participants read nine passages describing the nature of outer space phenomena, adapted from previous research (McDaniel & Donnelly, 1996). Exact texts are available in the Appendix. Each passage explained a single concept in one paragraph, averaging 56 words in total length. The complete series of nine passages was read by participants in a total of 2.5 minutes, on average.

When they finished studying to their satisfaction, participants were guided to estimate how well they believed they understood the material with a prompt adapted from (Thiede et al., 2010): *"When you finished reading the text material, how did you decide whether you had understood the passage? That is, when asked to 'grade' your comprehension of that passage, what do you base your grade on so you can say, 'I understood this passage well' or 'I read it, but I didn't understand it?'"*

Learners may judge the quality of their learning from a variety of thoughts, such as familiarity, ease of processing, or what can be recalled when the materials are no longer present. This open response prompt captures this variety without having first primed participants with common responses. Thus, the first measure comprised open responses written in the participants' own words and mimicked the standard method of obtaining cue reports in the literature.

After using their own words to report the metacognitive cues used to gauge their learning, participants were shown a list of the eight most frequently cited metacognitive cues from past experiments in the field (Griffin et al., 2009; Jaeger, 2012; Thiede et al., 2010; Wiley et al., 2018), detailed in Table 3 below. Participants were prompted to select which of these cues matched what they wrote in the previous open response question. Reflecting on the thoughts that went into estimating one's learning can be an unfamiliar process for participants, so a list was designed to help participants ensure that the cues they used were more likely to be reported.

Table 3.

Metacomprehension Cues

Cue	Item
Self-efficacy	My overall confidence in learning new material from reading.
Prior Knowledge	My knowledge of the topic from previous coursework or life experience.
Interest	My interest in the topic of the text.
Difficulty	How difficult or easy the text was to understand.
Effort	The amount of effort I put into trying to learn the material.
Recall	My ability to recall information from the text, such as facts or details.
Mental Imagery	The amount/quality of mental images or examples I came up with to understand the text.
Summarization	My ability to remember concepts or summarize/explain what the text was about.

Next, the full list of cues disappeared, with participants now seeing only the cues they selected in the previous question. In this phase, participants were asked to weigh the relative importance of each cue to decide how well they learned the material, with the total of all estimates forced to equal one hundred. This creates a proportional weighting for each cue in the estimate.

At this stage, both the presence/absence of each cue and the weight of how important each cue was to the overall learning estimate of the participant were gathered. Asking participants to weigh the relative importance of each cue may provide valuable additional information. Assuming a learner utilizes more than one cue to estimate their learning, it is likely that each cue is not equally represented in the estimate. Some thoughts likely predominate, whereas others are informative, but less so. Capturing this imbalance amongst reported cues is useful to evaluate more precisely the variation in cues and the causes of this variation.

After reporting and rating cues, participants next estimated how well they would do on an upcoming multiple-choice quiz on the learning materials with which they recently engaged with, on a scale from 0-100%. This metacognitive judgment of comprehension provides a general estimate of how well the participant believes they learned the material. After making these estimates, participants completed a 36-question multiple choice comprehension exam, with four questions per topic covered in the texts. Half of the questions tested recall of facts from the passages, the other half required inferencing to apply concepts from the passages to similar, but novel situations. As with the passages themselves, the questions were adapted from previous research (McDaniel & Donnelly, 1996). Each question could be answered by recalling information from the text.

RESULTS

Deep Cues are Reported Most Often in Open-Response Measures

Participants reported the cues used to estimate their comprehension initially through an open response item. Analysis of the text responses revealed that participants signaled their use of cues through keywords and phrases, such as “if I can”, “how well”, “able to”, and “whether”, following these phrases with a description of their metacognitive evaluation. The descriptions that followed overwhelmingly referenced three metacognitive cues: summarization, recall, and mental imagery. Summarization was the most reported cue, with 19.39% of all responses describing an ability to “explain it to someone” or “summarize in my own words” as a way in which participants estimated their understanding. Considering their ability to recall the information (“recall and explain what I read” or “regurgitate the information”) and/or their ability to visualize the information in the texts (“visualize the terms in my head” or “visualize the information”) were also common (9.84% and 8.61%, respectively). These unprompted responses were also commonly carried through to later cue-reporting measures, with 100% of those reporting mental imagery in the open response item also reporting this as a cue when prompted for their cues later, 90% of those reporting summarization doing the same, and 70.83% of those referencing recall as a cue also carrying these responses through their reporting. Participants rarely, if ever, reported using the remaining five cues commonly cited in similar research (prior knowledge: 1.23% of respondents, ease, or difficulty: .04%, interest: .82%, self-efficacy: 0%).

These numbers do not add up to 100 percent since some participants appear to have been confused about how to respond to the prompt, responding with "I understood the passage well" or "I understood a fair amount of information" instead of providing a description of how they came to this estimate. This observation casts doubts about the findings of previous research utilizing such a prompt as the sole measure of metacognitive cues. However, in this study participants were given additional opportunities to report cues in more quantitatively focused measures, which each of the remaining results focus exclusively on.

People Often Use Multiple Cues, Especially Deep and Shallow Processing Cues, to Assess Their Comprehension

Among the 244 participants, only 79 selected a single cue (32.40%), all others selected multiple cues (67.60%, overall sample M = 2.86, SD = 1.74). Table 4 shows a summary of how participants in the study indicated their usage of the different cues. The table summarizes (a) the proportion of participants who selected each cue (i.e., binary), (b) the mean relative weighting given to a cue when participants selected it, and (c) the mean relative weighting given to a cue across all participants (i.e., includes '0' for all those who did not select it).

A majority of participants (54.92%) indicated that they relied on their perceived capacity to summarize the text as a cue and assigned it the highest relative importance. This implies that most learners engaged in and acknowledged the utilization of a deep processing cue. However, many participants also relied heavily on fluency to derive their estimates of comprehension, with nearly half reporting difficulty as a cue (49.18%) or a general ability to recall facts and details (48.36%). Interestingly, although the development of mental images of concepts was rarely reported (31.15%), when participants reported this cue, it was a strong factor in their judgment of comprehension (35.92% relative weight when selected). Several cues were reported at lower rates, mentioned by a quarter of the participants or less. Interest as a cue was the least reported (20.90%), followed by self-efficacy as a reader (24.59%).

Table 4.

Appearance of Metacomprehension Cues in Self-report Data

Cue	Percentage Selecting as a Cue	Relative Weight When Selected *M*(*SD*)	Relative Weight, Overall *M*(*SD*)
Self-efficacy	24.59	30.33 (29.11)	7.46 (19.42)
Prior Knowledge	27.87	38.06 (29.89)	10.61 (23.21)
Interest	20.90	29.40 (25.63)	6.14 (16.69)
Difficulty	49.18	34.48 (28.38)	16.96 (26.32)
Effort	29.51	28.32 (22.06)	8.35 (17.60)
Recall	48.36	37.63 (30.88)	18.20 (28.54)
Mental Imagery	31.15	35.92 (29.23)	11.19 (23.27)
Summarization	54.92	38.41 (27.88)	21.09 (28.15)

Metacomprehension Cues Lack Distinct Factor Structure: Individual Cue Analysis is Warranted Over Factor Groupings

To determine if there are distinct ‘profiles’ of reported cue usage, binary cue use data were examined in two ways. First, item-centered analyses of the reported cues aimed to identify latent groupings within the response data, verifying if the hypothesized groupings from the metacomprehension literature align with real-world data. Second, person-centered analyses were carried out to find potential ‘profiles’ of cues using probabilistic cluster analysis. In both approaches, binary data were employed to analyze cue reporting (coded 0 if not reported, 1 if reported). The rationale for not using relative weightings is their lack of independence, as the proportion of each cue depends on the other cues selected and their quantity. Binary cue data, while not continuous, are independent and can be utilized to model relationships through tetrachoric correlations and binary clustering algorithms.

Correlations among Binary Cue Data Suggest Presence of Underlying Factors

As shown in Table 5, tetrachoric correlations of binary cue data indicate that most relationships exist within predefined cue types (i.e., prior, situational-shallow, situational-deep), as expected. Shallow cues (difficulty and effort) displayed the strongest connection (r = 0.58), although they were also positively related to summarization (difficulty and summarization: r = 0.23) and interest (effort and interest: r = 0.44). Deep cues (mental imagery and summarization) demonstrated a strong positive correlation (r = 0.36) with each other. Additionally, summarization had strong positive associations with recall (r = 0.48) and difficulty (r = 0.23). Prior cues (self-efficacy as a reader, prior knowledge, and interest) were all strongly interconnected and occasionally exhibited significant relationships with shallow cues (effort and difficulty). While some cues had negative associations, none of the negative relationships were significant.

Table 5.

Tetrachoric Correlations Among Binary Cue Data

Cue	Self-efficacy	Prior Knowledge	Interest	Difficulty	Effort	Recall	Mental Imagery
Prior Knowledge	.23**						
Interest	.35**	.34**					
Difficulty	.24**	.32**	.35**				
Effort	.16	.20*	.44**	.58**			
Recall	-.03	-.06	-.02	.02	.12		
Mental Imagery	.08	.13	-.00	.13	.12	-.02	
Summarization	.13	-.01	-.11	.23**	.16	.48**	.36**

Note. *p < .05, ** p < .01.

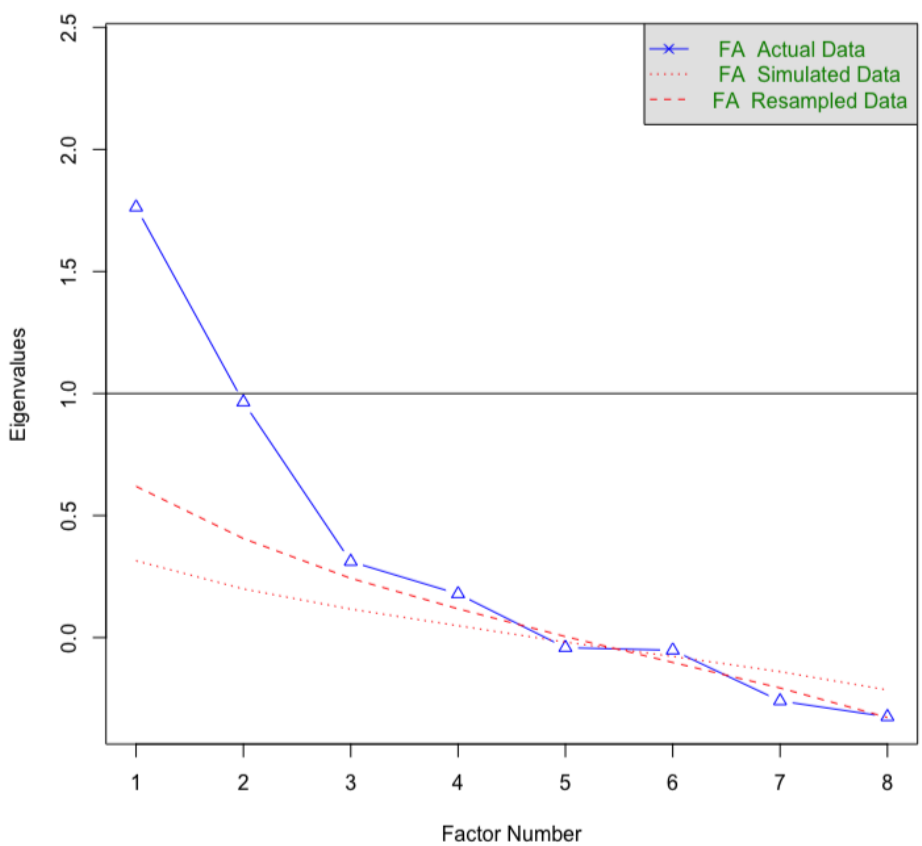
Overall, these patterns suggest that there may be underlying factors. In general, cues tend to associate according to the three main types (prior, situational-shallow, situational-deep), which may be reflected in a factor analysis. Checking the data for suitability for factor analysis, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy score of .57 is greater than .50, suggesting the proportion of variables which may be caused by underlying factors may be adequate (Kaiser, 1970; Kaiser & Rice, 1974). Additionally, Bartlett's Sphericity Test of Sphericity (*K2* = 19.168, df = 7, *p* = .008) indicates the variables may be related closely enough to warrant investigation (Bartlett, 1950). To further explore whether a sufficient underlying structure exists between cues, a factor analysis of the binary cue data was completed.

Factor Analysis Fails to Capture Distinct Underlying Structures for Metacomprehension Cues

For the item-centered analysis, an exploratory factor analysis using maximum likelihood estimation with varimax (orthogonal) rotation was performed for the eight-item collection (N=244) of reported cues using the fa() function of the psych package in R. As shown in Figure 3, parallel analysis (Velicer & Jackson, 1990) using simulated data suggests four components, while a visual inspection of the plot through the Scree test (Cattell, 1978; Cattell & Vogelmann, 1977) suggests a three-factor solution. As such, both a three-factor and a four-factor model were run. The three-factor and four-factor solutions yielded a complete set of factors, each with eigenvalues greater than one. The four-factor model explained a larger total proportion of the variance for the set of cues, 58.9% of the total variance, versus 48.7%, respectively.

Figure 3.

Parallel Analysis Scree Plots



A three-factor model was hypothesized based on the theorized distribution of cues from literature. The factors brought into the learning task (self-efficacy, prior knowledge, interest) were labeled ‘prior’. The factors pertaining to the task experience that typically involve shallow processing (difficulty, effort) were labeled ‘situational–shallow’. The factors pertaining to the task that typically involve deep processing (recall, mental imagery, summarization) were labeled ‘situational–deep’. Using loading cutoffs of |.3|, the three-model solution fits this theorized model well, as shown in Table 6. The factor with the highest eigenvalue extracts the three ‘situational-deep’ cues cleanly, with none loading above |.14| on any other factor. A second factor, consisting mostly of ‘situational-shallow’ fluency cues, is predominated by effort, with both difficulty and interest mapping above the cutoff value. The third factor contains mostly ‘prior’ factors, though there is cross-loading between this factor and the ‘situational–shallow’ fluency factor for both the interest and difficulty cues.

Table 6.

Factor Loadings for Factor Analysis of Binary Cue Data with Principal Axis Factoring after Varimax Rotation, Three-factor Model

Cues	Prior	Situational-Shallow	Situational-Deep	Communalities
Self-efficacy	**.524**	.057	.114	.29
Prior Knowledge	**.531**	.122	-.033	.30
Interest	**.542**	**.381**	-.138	.46
Difficulty	**.407**	**.503**	.206	.46
Effort	.165	**.975**	.132	.99
Recall	-.128	.084	**.482**	.26
Mental Imagery	.146	.048	**.352**	.15
Summarization	.035	.028	**.996**	.99
SS loadings	1.083	1.376	1.442	
Proportion of Variance	.135	.172	.180	
Cumulative Variance	.135	.307	.487	

Note: Loadings over |.30| are in bold.

A four-factor model, shown in Table 7, explains nearly 10% more of the overall variance in cue reporting. This model has similar loadings for two factors, again labeled 'prior' factors and 'situational-shallow' factors but splits out the individual cue 'recall' from the 'situational-deep' factor as its own factor. Removing recall from the factor labeled 'situational-deep' maintains loadings above the cutoff value (|.30|) for both factors and each factor retains an eigenvalue greater than one, suggesting a four-factor model may fit the data best. This four-factor model implies that using recall as a cue to assess comprehension may not be strictly deep, potentially serving as a midpoint between shallow and deep cues. This outcome is reasonable, as recalling facts or details can be linked to shallow processing (when one remembers facts from a brief skim of a text) or deep processing (when one recalls facts through deeper engagement with the text).

Table 7.

Factor Loadings for Factor Analysis of Binary Cue Data with Principal Axis Factoring after Varimax Rotation, Four-factor Model

Cues	Prior	Situational-Shallow	Recall	Situational-Deep	Communalities
Self-efficacy	**.515**	.045	-.023	.167	.30
Prior Knowledge	**.509**	.109	-.037	.012	.27
Interest	**.622**	**.343**	.004	-.125	.52
Difficulty	**.379**	**.505**	-.004	.235	.62
Effort	.194	**.972**	.074	.089	.99
Recall	-.056	.055	**.990**	.093	.99
Mental Imagery	.073	.073	-.061	**.414**	.19
Summarization	-.026	.060	**.393**	**.914**	.99
SS loadings	1.102	1.342	1.146	1.123	
Proportion of Variance	.138	.168	.143	.140	
Cumulative Variance	.138	.306	.449	.589	

Note: Loadings over |.30| are in bold.

While the factor analysis demonstrated that certain clusters of metacomprehension cues are consistent with theoretical predictions, the overall fit indices suggest that these clusters do not form a cohesive factor structure. Specifically, the Root Mean Square Error of Approximation (RMSEA) values for both the three-factor (.111) and four-factor (.114) models fall into the range characterized as 'poor' fit (Browne, 1993). Similarly, the Tucker-Lewis Index (TLI) values for the three-factor (.77) and four-factor (.76) models are substantially lower than the recommended threshold of .90 (Hu & Bentler, 1999; Tucker & Lewis, 1973), indicating that the models fail to adequately capture the underlying structure of the data. These discrepancies suggest that while the metacomprehension cues may be logically grouped, their utility as latent factors for evaluating comprehension is not supported by the current dataset. Consequently, the subsequent analyses will not aggregate these cues into factors but will instead treat each cue as an individual predictor in the examination of their relationship with judgments of comprehension and recall.

Depth of Processing Influences Comprehension Estimation and Exam Performance: Insights from Four Cue Profile Classes

For the person-centered analysis, a latent class analysis (LCA) was used to examine common profiles of cue reporting amongst participants (Hagenaars & McCutcheon, 2002). LCA is a clustering algorithm that can analyze binary data and provides the probability of an individual belonging to a particular latent class. LCA segments individuals into class profiles, but does so flexibly, allowing for individuals to represent aspects of different profiles. In this type of model, a person is given a probability of appearing in each respective class, as opposed to forcibly being placed into a class that may be more of a 'stretch', a drawback of k-means and many other clustering algorithms. For LCA, a model that fits the data well will be confident in its classifications, demonstrating a high entropy R-squared (Granado, 2015).

This study evaluated several latent class models with different numbers of classes to determine which model best explained the variation in cue reporting by individuals. Although the two-class model had the lowest Bayesian Information Criterion (BIC) score (Raferty, 1995) (2427.79 for class 2, 2457.68 for class 4), indicating it was more parsimonious, the low entropy R-squared (.61 for class 2, .80 for class 4) and small class size meant that all eight cues were relatively likely in either class. This resulted in high probabilities of class membership for many forms of cue reporting, making it less effective in explaining variation.

In contrast, the four-class model had a higher entropy R-squared statistic, which more readily explained the variation in cue reporting by producing four distinct profiles. As shown in Table 8, other models with a single class, two classes, or more than four classes either had higher BIC scores or lower entropy R-squared scores, indicating that they were less effective in explaining the variation in cue reporting. Therefore, the four-class model was considered the most appropriate model, based on the data.

Table 8.

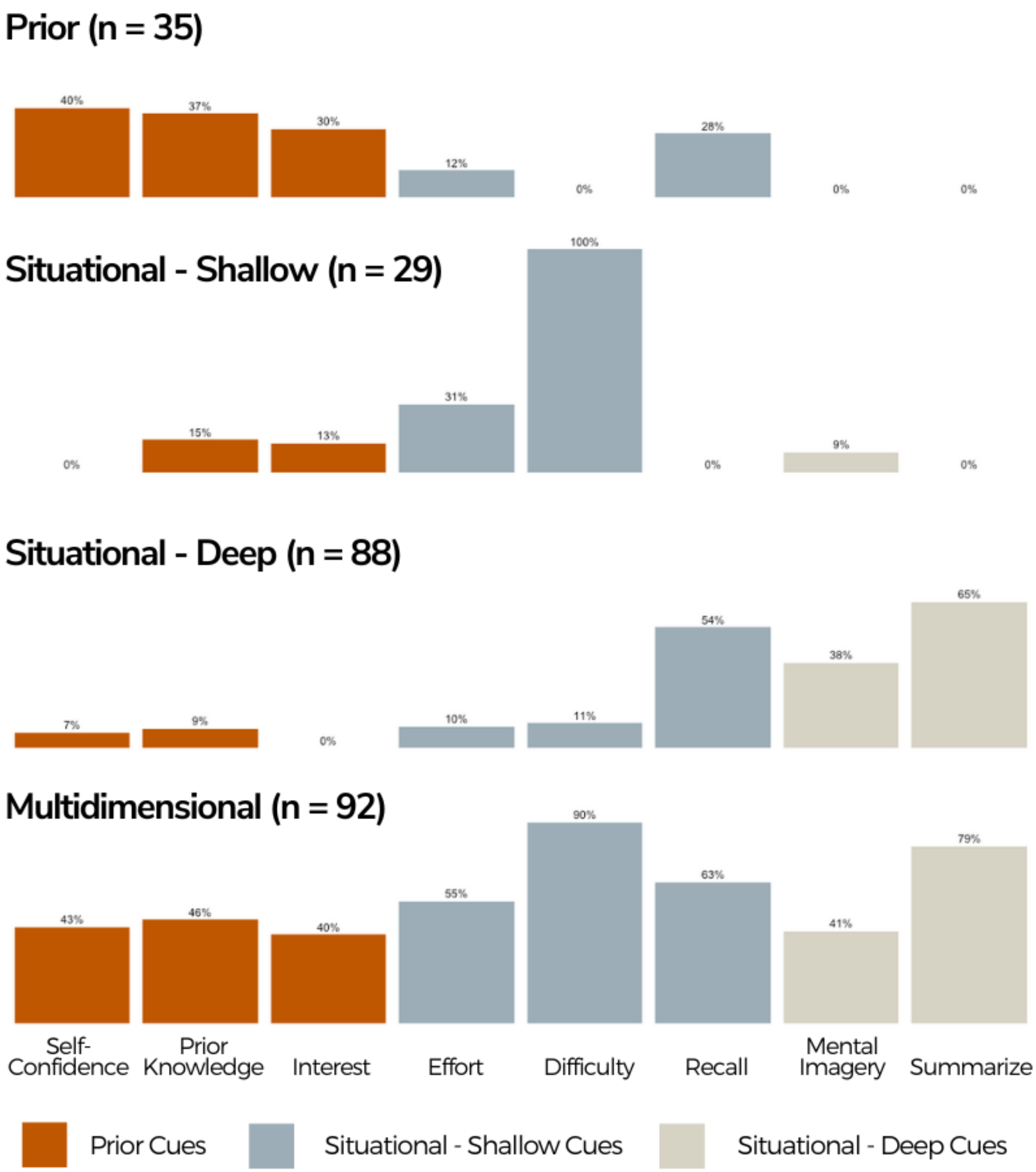
Results of Latent Class Analysis (LCA) for Each Class, by Model

Model	df	AIC	BIC	Log-likelihood	Chi-square goodness of fit	Entropy R-squared
1 Class	236	2437.94	2465.92	-1210.97	668.29	-
2 Classes	227	2368.34	2427.79	-1167.17	337.10	.614
3 Classes	218	2341.34	2432.26	-1144.67	310.54	.707
4 Classes	209	2335.28	2457.68	-1132.64	287.13	.796
5 Classes	200	2326.68	2480.55	-1119.34	242.78	.792
6 Classes	191	2324.92	2510.27	-1109.46	223.18	-
7 Classes	182	2324.49	2541.31	-1100.24	217.21	.803
8 Classes	173	2328.50	2576.80	-1093.25	194.30	.861

In applying the four-class Latent Class Analysis (LCA) model, participants were categorized based on the highest probability of class membership, which facilitated the identification of distinct learner profiles. Figure 4 illustrates this classification, highlighting a diverse set of metacognitive cues utilized by the learners. One class, encompassing 92 participants, was characterized by a broad and varied use of cues spanning the entire spectrum available (termed 'multidimensional'). Conversely, three other learner profiles exhibited more specialized cue utilization patterns. A group of 29 participants predominantly relied on their subjective experience of the learning process, such as the perceived ease or difficulty of the material, indicating a 'situational-shallow' engagement. Another distinct class of 35 learners drew heavily on their pre-existing knowledge and attitudes, including prior understanding, interest, and self-assessed capability in learning from similar materials (labeled 'prior'). The final profile, comprising 88 participants, was defined by a reliance on the content derived directly from the texts. These learners' cues were focused on their capacity to remember details, visualize concepts, and explain the concepts from the text, representing a 'situational-deep' approach to learning.

Figure 4.

Four Labeled Profiles from Latent Class Analysis, with the Proportion Reporting each Cue, by Profile



Participants generally estimated their comprehension to be at a level where they could correctly answer about half of an anticipated set of multiple-choice questions. The average predicted comprehension stood at 51.15%, but there was variability (SD = 23.27%).

To gain a deeper understanding of this, I delved into the role of cue profiles on these comprehension judgments. If learners consistently use certain cues, allowing for a distinct profile to be identified for each, such profiles could help tailor interventions more precisely for individual students. Using the four-profile model to assign each participant to the cue profile they fit closest to (prior, situational-surface, situational-deep, multidimensional), I then analyzed whether there were differences in how each class of learner estimated their comprehension. As shown in Table 9, while group means aligned with expectations based upon depth of processing (those classes typically utilizing deeper processing cues had higher estimates of comprehension (situational-deep, multidimensional), with the inverse true for prior and situational-shallow, an analysis of variance showed no effect of class on judgment of comprehension, $F(3, 240) = 1.715$, $p = .165$.

While comprehension judgments showed no significant differences across classes, actual exam performance did vary. An analysis of variance highlighted that the classes had at least one divergent mean score, $F(3, 250) = 2.895$, $p = .036$. Notably, even though the situational-deep class performed 9.65% better than the situational-shallow class, this difference wasn't statistically significant, likely a consequence of small sample sizes. Furthermore, deeper cue profiles exhibited stronger correlations between comprehension judgments and exam scores: situational-deep ($r = .54$, $p < .001$) and multidimensional ($r = .51$, $p < .001$). In contrast, the prior and situational-shallow profiles recorded correlations of $r = .40$ (p = .030) and $r = .33$ ($p = .097$) respectively.

Table 9.

Analysis of Variance Results for Four Class Cue Profile Model Impacts to Judgments of Comprehension

Measure	Prior N = 35		Situational Shallow N = 29		Situational Deep N = 88		Multi-dimensional N = 92		$F(3, 240)$	η^2
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*		
Judgment of Comprehension	.506	.243	.433	.240	.544	.232	.501	.225	1.715	.021

In this study, the four-class LCA model was instrumental in identifying distinct learner profiles based on cue utilization. However, the subsequent utility of these profiles for predicting self-judged comprehension was not supported by the data. Despite the theoretical appeal of personalized intervention strategies based on cue profiles, the lack of significant differences in comprehension judgments across classes suggests a disconnect between cue utilization and metacognitive awareness.

Despite these shortcomings, the divergence between predicted and actual performance, along with the correlations observed in deeper cue profiles, does suggest nuanced relationships worth exploring. However, the current evidence does not support the use of this model as a reliable tool for predicting comprehension judgment, necessitating alternative approaches to understand and leverage these relationships in future research.

Reporting Certain Cues Enhances Comprehension Judgments, While Others Diminish It

Moving away from cue profiles, further analyses examined the impact of individual cues on judgments of comprehension. To infer the strength of relationship between a cue and its impact to an overall judgment of comprehension independent t-tests were conducted to compare the judgments of comprehension between those who did not report a cue and those who did report a cue. The means and the results of the t-test are reported in Table 10.

Among the cues relating to prior beliefs and knowledge, those who reported using self-efficacy and prior knowledge cues gave higher judgments than those who did not report using those cues; those who reported using interest as a cue gave lower judgments. Reported use of the two situational surface cues—difficulty and effort—was associated with lower judgments of comprehension. I speculate that those who reported these two cues found the passages more difficult and effortful, though direction cannot be determined from the study's measures. Taking a holistic view, every single prior cue and surface cue was related to metacomprehension.

In contrast, only one of the deep cues was related to metacomprehension—those who used mental imagery as a cue gave higher judgments of comprehension than those who did not. Report of recall and summarize cues were not related to judgments of comprehension.

Table 10.

Independent t-test Relationships Between Judgments of Comprehension of Those Reporting the Cue and Those not Reporting the Cue

	Mean (*SD*) Judgment of Comprehension			Significance tests		
Cue	Did Not Report Cue	Reported Cue	Reported-Not Reported Difference	*t*	df	*p*
Self-efficacy	48.97 (23.13)	57.82 (22.61)	8.85	2.618	102.30	0.010
Prior Knowledge	49.35 (24.16)	55.80 (20.24)	6.45	2.110	144.36	0.037
Interest	52.72 (23.40)	45.21 (22.01)	-7.51	-2.139	82.40	0.035
Difficulty	54.39 (23.63)	47.80 (22.51)	-6.59	-2.231	241.94	0.027
Effort	54.76 (22.87)	42.52 (22.07)	-12.25	-3.912	137.63	<.001
Recall	53.66 (23.72)	48.47 (22.57)	-5.19	-1.751	241.94	0.081
Mental Imagery	48.36 (23.30)	57.31 (22.14)	8.95	2.875	151.84	0.005
Summarization	48.79 (23.65)	53.09 (22.86)	4.30	1.435	229.64	0.153

Table 11 shows the results of a multiple regression analysis that predicted the judgment of comprehension based on the presence or absence of the eight cues. The analysis revealed a significant effect between cue reporting and judgment of comprehension ($F(8, 235) = 6.69$, $p < .001$, Adjusted R-squared = .16). The broad pattern of relationships that emerged from the t-tests remains the same, though when controlling for the presence or absence of the other seven cues, only four cues remain statistically significant. Three cues are positively associated with judgments of comprehension: self-efficacy, prior knowledge, and mental imagery. The unstandardized beta values indicated that reporting these cues resulted in a 6% to 9% boost in confidence in understanding. Therefore, when students reported using their thoughts about self-efficacy or their prior knowledge when thinking about how well they understood a text passage, their confidence in having learned the material was likely to be higher, boosting their judgment of comprehension. A similar association was found between the ability to mentally visualize concepts from the text.

However, when learners reported using the effort required to understand the text as a cue, they were more likely to revise their estimates of learning downward, lowering their estimated learning an average of nearly 11 percent. Interestingly, while independent t-tests revealed a significant negative relationship between reporting interest or difficulty as cues and judgments of comprehension, when controlling for the reporting of other cues this association is no longer significant (interest: $t = -1.775$, $p = .077$, difficulty: $t = -1.917$, $p = .056$).

Table 11.

Unstandardized Regression Summary Table, Predicting Judgments of Comprehension from Reported Cue Use

Predictor	*B*	*B* 95% CI	SE	*t*	*p*
(Intercept)	51.16	[45.77, 56.55]	2.737	18.690	<.001
Self-efficacy	9.91	[3.44, 16.38]	3.283	3.018	0.003
Prior Knowledge	8.11	[1.86, 14.36]	3.172	2.555	0.011
Interest	-6.43	[-13.56, 0.71]	3.622	-1.775	0.077
Difficulty	-5.82	[-11.80, 0.16]	3.036	-1.917	0.056
Effort	-10.83	[-17.37, -4.30]	3.317	-3.265	0.001
Recall	-5.62	[-11.34, 0.11]	2.907	-1.932	0.055
Mental Imagery	7.84	[1.83, 13.85]	3.051	2.569	0.011
Summarization	5.41	[-0.57, 11.39]	3.035	1.784	0.076

Reporting Prior Knowledge and Situational Deep Cues Boosts Comprehension Exam Scores

Participants answered 36 multiple choice comprehension questions with a median percent correct of 55.56% (M = 54.75%, SD 18.58%). As illustrated in Table 12, participants who predominantly employed situational-deep cues or a broad spectrum of cues (multidimensional) generally outperformed others. In contrast, those relying heavily on shallow processing cues, specifically the situational-shallow and prior categories, lagged behind. An analysis of variance confirmed a distinct influence of cue profiles on performance, $F(3, 240) = 4.101$, $p = .007$. Specifically, the situational-deep group outscored the prior group ($p = 0.024$). However, other group comparisons did not reveal significant performance differences.

Table 12.

Analysis of Variance Results for Four Class Cue Profile Model Impacts to Exam Performance

Measure	Prior N = 35		Situational Shallow N = 29		Situational Deep N = 88		Multi-dimensional N = 92		*F*(3, 240)	η^2
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*		
Exam Score	.492	.200	.476	.176	.573	.187	.558	.176	2.895*	.035

Turning from profiles of reported cue usage to an examination of individual cues and their influence on exam outcomes, distinct patterns emerged. As shown in Table 13, from the prior cues category, only prior knowledge demonstrated a notable positive impact on performance. In contrast, interest and self-efficacy did not significantly influence results. While surface cues like difficulty and effort showed no considerable linkage with performance outcomes, deep cues told a different story. Participants who reported using mental imagery or summarization as cues notably outperformed those who did not. Recall, however, did not exhibit such an association.

Table 13.

Independent t-Test Relationships between Exam Score of Those Reporting the Cue and Those not Reporting the Cue

	Mean (*SD*) Judgment of Comprehension			Significance tests		
Cue	Did Not Report Cue	Reported Cue	Reported-Not Reported Difference	*t*	df	*p*
Self-efficacy	55.00 (18.24)	53.98 (19.72)	-1.02	-0.353	94.14	.725
Prior Knowledge	52.71 (19.41)	60.01 (15.10)	7.29	3.111	155.59	.002
Interest	55.89 (18.59)	50.44 (18.07)	-5.45	-1.905	80.25	.060
Difficulty	54.93 (19.80)	54.56 (17.30)	-0.37	-0.155	239.55	.877
Effort	56.01 (19.01)	51.74 (17.25)	-4.27	-1.711	145.92	.089
Recall	55.42 (18.88)	54.03 (18.30)	1.40	-0.587	241.70	.558
Mental Imagery	51.29 (18.04)	62.39 (17.54)	11.10	4.538	148.62	<.001
Summarization	50.10 (18.31)	58.56 (17.97)	8.46	3.621	231.07	<.001

When considering the broad impact of cue reporting on comprehension exam scores, a significant relationship emerged. Multiple regression analysis, controlling for the reporting of all available cues, indicated a substantial collective effect between cue reporting and exam score: $F(8, 235) = 6.36$, $p < .001$, R-squared = .18). At the cue level, prior knowledge significantly enhanced performance when reported as a cue, as shown in Table 14. In contrast, cues like interest and self-efficacy as a reader did not significantly impact performance. Among deep cues, mental imagery, and the ability to explain one's understanding led to notably better comprehension exam scores. However, some cues, specifically shallow cues, did not show a significant association with exam performance. Overall, learners reporting the use of prior knowledge or deep cues, especially summarization and mental imagery, consistently outperformed their peers who didn't report these cues.

Table 14.

Unstandardized Regression Summary Table, Predicting Exam Score from Reported Cue Use

Predictor	*B*	*B* 95% CI	SE	*t*	*p*
(Intercept)	50.086	[45.76, 54.41]	2.915	22.822	<.001
Self-efficacy	-1.954	[-7.14, 3.23]	2.632	-0.742	0.459
Prior Knowledge	8.414	[3.40, 13.42]	2.543	3.308	0.001
Interest	-4.655	[-10.38, 1.07]	2.904	-1.603	0.110
Difficulty	-1.065	[-5.86, 3.73]	2.434	-0.437	0.662
Effort	-4.807	[-10.05, 0.43]	2.660	-1.808	0.072
Recall	-3.357	[-7.95, 1.23]	2.331	-1.440	0.151
Mental Imagery	9.027	[4.21, 13.85]	2.446	3.691	<.001
Summarization	8.237	[3.44, 13.03]	2.433	3.385	0.001

DISCUSSION

Estimating our comprehension accurately is important to ensure we learn efficiently, studying the right material for the right amount of time. Research has shown that a variety of cues are used to inform estimates of understanding, but little is known about the variety of these cues, how they interrelate, how many are used to make an estimate, and how each cue may impact estimates and performance on an exam. These results reveal that people use a variety of cues, often more than one at a time, to estimate their understanding of short text passages. Instead of a simple world where learners use a single cue to estimate how well they understand the concepts and information in a learning task, we live in a complex world where learners often use multiple cues to evaluate their comprehension. When reading short text passages, readers report using strategies arising from experience (e.g., difficulty, effort), from thinking deeply about the text (e.g., summarizing), and from their own unique experiences they bring into the task (e.g., prior knowledge, interest).

However, learners do not report these cues evenly. Although distinct patterns in cue reporting emerged among learners—ranging from those who focus primarily on a specific category, such as fluency cues like effort and difficulty, to those who employ a 'kitchen sink' approach, incorporating a wide range of cues—the linkage of these profiles to judgments of comprehension is not statistically significant.

Additionally, each cue can influence a learner's estimate in disproportionate ways, carrying more or less influence than other reported cues. People easily report using strategies they were likely taught in school (summarize, recall, draw visual analogies). Other cues, such as fluency cues (difficulty, effort), appear to hold sway in drawing evaluations of comprehension, but these may need prompting to be specifically reported by learners as they only appeared after being suggested as possible cues one may use.

Further, cues which come to mind appear to influence both the overall estimate of comprehension and learning. When reported, some cues are associated with increased estimates of understanding, while others are associated with lower estimates. This makes intuitive sense. When one has relevant prior knowledge or is confident in their ability to explain the material to a friend, these metacognitive cues should move their estimate of understanding higher, while feeling that a passage was particularly difficult to parse should lower perceived levels of understanding. The same should hold for learning outcomes and did so in this study. When participants reported using cues tied to deep cognitive experiences, such as summarization and producing mental images to understand the text, these participants were more likely to perform better on the subsequent exam. The same held true for prior knowledge, though not for other individual factors which may not be directly applicable to the learning task, even if they feel they are (e.g., interest in the topic, self-efficacy as a reader). Finally, cues tied to the fluency of the learning experience, such as how difficult or effortful learning felt, were commonly reported by participants, and were significantly related to judgments of comprehension, but did not relate to actual performance. Notably, in this experiment, participants simply reported whether they used a cue to estimate their understanding or not. The valence, or direction, of the cue, such as whether those reporting using difficulty as a cue found the learning to be difficult, or easy, cannot be determined in the present study.

Additional limitations may also apply to these results, as the cues reported by learners herein may not match the cues actually used (Undorf et al., 2018). It is possible participants use metacognitive cues in estimating their learning but fail to report them. Relatedly, participants may have reported using cues which, in reality, did not factor into their metacognitive judgments. While these are limitations, the danger of unreported cues and biased cue reporting impacting the results should be limited by the large sample size which should somewhat insulate these results.

In this study, I've identified that learners use a variety of cues to estimate their level of comprehension. While these cues appear to fall into distinct categories and carry different levels of influence, more nuanced understanding is needed. For instance, the specific directionality of cues—whether a learner finds a text 'difficult' or 'easy'—remains an open question. The next phase of my research will directly address this gap by incorporating measures for the valence or directionality of each reported cue. Additionally, I will explore potential factors that might affect these cue selections, such as individual differences in learning styles or previous experiences.

CHAPTER 3: MEASURING METACOMPREHENSION CUES

Previous studies have asked learners to report the cues they used to estimate their comprehension through a single open-ended question. However, this method has a limitation: learners may not be able to verbalize their cues easily. In Study 1 (Chapter 2), I provided learners with a checklist of possible cues based on prior research. I hoped that seeing the list would help learners recognize the cues they used. The results showed that learners used multiple cues to estimate their comprehension. These cues influenced both their estimates and their test performance. Specifically, all three prior cues (prior knowledge, self-efficacy, interest), both situational shallow cues (effort, difficulty), and one situational deep cue (mental imagery) were related to judgments of comprehension. The other two situational deep cues (recall, explanation) were not.

Switching from judgments of comprehension to actual exam performance, in Study 1 only three of the eight cues were related to performance: one prior cue (prior knowledge) and two situational deep cues (mental imagery, explanation). These findings suggest a difference between the cues that affect judgments of comprehension and those that affect performance.

While the methods in Study 1 help learners report their cues with less friction, this approach still has limitations. A key limitation of the methods in Study 1 is the inability to understand the direction of influence of each reported cue (i.e., how the cues were used). For example, when a participant reported relying on "How difficult or easy the text was to understand", it is unclear as to whether this difficulty experienced was high or low. To address this limitation, in the following studies I asked participants to self-report cue use and collected information about their judged valence of each cue. This allows me to test both participants' metacognitive beliefs and the accuracy of their metacognitive insight simultaneously.

How does this additional measure help? In the prior study, without an understanding of the direction of the cue I could not accurately evaluate the impact of cues on outcome measures. For instance, while I could measure whether a participant had a higher or lower judgment of comprehension when reporting difficulty as a cue, I could not identify whether this outcome was the result of perceiving low difficulty (ease) or high difficulty (difficult). Not only does this hamper individual analyses, but at a larger level, if half of those reporting difficulty as a cue believed the task was easy, but the other half believed the task was difficult, without additional information these differences will cancel each other out, hiding the evidence for how this cue truly affects comprehension and estimates of comprehension. With a valence measure, I can now gain better insight into how individual, and collective, cues relate to judgments of comprehension.

The addition of valence as a measured variable is not merely a methodological refinement; it addresses a significant gap in the current literature on metacognitive cues. Previous research has largely focused on the identification and categorization of these cues, without addressing the directionality or quality of their influence. By introducing valence as a factor for consideration, these studies offer a more comprehensive, nuanced understanding of metacognitive processes. They enrich our conceptual models by adding another layer of complexity, pushing the field toward a more holistic understanding of how individuals engage in metacognition. Furthermore, knowing the valence of cues can provide critical insights into the reliability and validity of self-reported data on learning experiences, as it adds a dimension to the data that could explain previously unaccounted-for variations in study results.

Furthermore, collecting cue-valence separate from cue-usage also allows me to test whether participants have insight into what cues are influencing their judgments of comprehension. If I find that reported usage does not relate to judgments of comprehension, but cue valence does, then this would imply that participants are unable to explicitly report when a cue influences their judgments. The use of some cues may be easier for participants to recognize than others.

Secondary Questions: Potential Moderators of Cue Valence and Judgments of Comprehension

In Study 1, I did not find actionable profiles of cue usage. Although learners consistently reported distinct 'profiles' of cue usage, these profiles did not translate into statistically significant predictors of either their judgments of comprehension or their exam performance. However, prior research indicates individual differences exist at some level in the reporting of cue-use (Griffin, 2008). For example, Thiede et al. found that college students labeled as at-risk readers relied on different cues than those labeled not-at-risk (2010). Specifically, at-risk reading students were more likely to report cues which are associated with lower metacognitive accuracy (e.g., perceived effort, difficulty). Looking at this from a different angle, there are patterns of behavior which metacognitively ‘good’ readers follow. These learners engage deeply with the materials, searching for relevant prior knowledge to aid in comprehension, asking themselves questions as they read, and developing visual images to aid in understanding. Further, these readers are also less prone to use potentially misleading cues such as feelings of ease or effort (Thiede et al., 2010).

In Studies 2 and 3, I take a different approach to examining whether there may be individual differences in cue-usage. While cue reporting may capture some level of individual differences, the single-shot nature of one reading task may not capture the typical or everyday habits of each participant. Potentially, a better indicator may be everyday habits. Thus, in Study 2 I ask participants to report their typical reading strategy use. The rationale is straightforward: individuals who frequently employ ‘good’ metacognitive strategies—like making cross-text connections or visualizing content—are more likely to rely on cues associated with deeper cognitive processing.

To capture how often a participant utilizes metacognitively 'good' reading strategies I include two different measures of everyday reading behavior in Study 2. One measure, the Metacognitive Reading Strategies Questionnaire (MRSQ) measures how often a learner utilizes 'good' metacognitive reading strategies when engaging in any form of reading (Taraban et al. 2004). A second measure, the Contextualized Reading Strategy Survey (CReSS) measures similar strategies in specific reading contexts (Denton et al. 2015). Together, these individual difference measures will allow me to examine whether the relationships between cue-use, cue-valence, and judgments of comprehension are moderated by the use of metacognitively 'good' reading strategies.

I also speculate that interest may be an important moderator, with people attending to material and to comprehension cues in different ways depending on interest level. Here, the premise is that interest level may alter the quality and depth of engagement with the content. For instance, individuals who find the material engaging may rely more on deeper cognitive cues, like visual imagery and summarization, and report more positive valence for these cues. Conversely, those less engaged may resort to cues related to effort and difficulty. To explore this, Study 3 incorporates an experimental manipulation of interest. Participants are exposed to material of either high or low personal interest, allowing me to investigate how interest level potentially moderates the relationships between cue use, cue valence, and judgments of comprehension.

Studies 2 and 3

In this chapter, I conducted two experiments using a similar paradigm to that of Chapter 2. Participants engaged with the learning materials then reflected on their learning, reporting the metacognitive cues used to gauge their learning, weighing these cues, and calculating a judgment of their comprehension. The difference is that now participants were asked to rate the valence of each cue, the materials were modified to be more complex, and two individual difference measures were added. However, the impacts of these changes from a participant's perspective should be minimal in terms of data collection. The valence items were tucked into the metacognitive cues section, reported as part of the standard series of reflective items.

Further, instead of reading short texts, I asked participants to read an essay (Study 2) and to watch an educational video (Study 3). In Study 1, the short texts (only a few sentences per topic) provided minimal opportunity to engage deeper cues, such as recall and summarization, perhaps explaining why these cues were not as predictive of judgments of comprehension as hypothesized. By expanding the length of the content, whether explained through text or video, participants should have more relevant opportunities to utilize deeper cues.

I also added two individual difference measures (typical reading strategy use and interest in a topic before engaging with the material), whose logic for inclusion are described above.

Research Questions

The main question for the studies in this chapter is whether cue valence predicts judgments of comprehension over and above cue use. More specifically, answering this question involves:

1. Is cue valence independent from cue use?

2. Are the cue valence measures themselves distinct or do they cluster in meaningful ways?

3. Does cue valence predict judgments of comprehension more consistently than cue use?

A secondary question for the studies in this chapter is whether individual differences moderate the relationship between cue valence and judgments of comprehension? More specifically, answering this question involves:

A. Do individual differences in reading strategies (Study 2) and interest (Study 3) relate to cue valence?

B. Do individual differences in reading strategies (Study 2) and interest (Study 3) relate to judgments of comprehension?

C. Do individual differences in reading strategies (Study 2) and interest (Study 3) moderate the relationship between cue valence and judgments comprehension?

Study 2

METHOD

Participants

Two-hundred-forty undergraduate students, recruited from a wide subject pool across a variety of majors (30.80% STEM, 20.25% Social Sciences) completed the study in exchange for course credit. After removing incomplete responses, the final sample contained 237 observations. A majority female (58.23% sample, 54.4% student body), and white (52.74% sample, 38.9% student body), the sample roughly resembled the student body demographics of the large public university at which it took place (Latinx: 17.30% sample, 23.4% student body; Asian: 21.52% sample, 20.2% student body; Black: 4.22% sample, 5.3% student body). The age of participants reflected the university environment, as well (20.68 *M*, 2.57 *SD*).

Materials

Cue Valence

To measure the valence of each cue, participants responded to a valence item for each of the eight metacomprehension cues most reported in prior research. Each item featured a prompt asking for the perceived experience with each cue, with participants rating each cue on a scale from 0-10, detailed in Table 15.

Table 15.

Items for Rating Cue Valences

Cue	Item	Scale (0-10)
Self-efficacy	How confident are you in your ability to learn new material from reading?	[not confident - very confident]
Prior Knowledge	How much knowledge of the topic did you have before you read the text?	[none - a lot]
Interest	How interested were you in the topic of the text?	[not at all - very]
Difficulty*	How easy or difficult did you find it to understand the processes described in the text?	[very hard - very easy]
Effort	How much effort did you put into trying to learn the information contained in the text?	[none - a lot]
Recall	Right now, how many of the facts or details of the text do you think you can recall?	[none - a lot]
Mental Imagery	How often did you come up with mental images or examples to help you understand the processes in the text?	[not at all - a lot]
Summarize	Right now, how well do you think you can remember the concepts and summarize or explain the process?	[not at all - very well]

Note. The valence response to the difficulty item was reversed coded for subsequent analysis such that higher ratings on this item represented greater experience of difficulty.

Reading Strategy Measures

To measure individual differences in typical metacognitive strategy use among participants, two separate measures were completed by participants. The first measure, the Metacognitive Reading Strategies Questionnaire (MRSQ) is a self-report measure of reading strategies used for comprehension and studying (Taraban et al., 2004). The measure utilizes 22 items wherein students report how often they use each of the strategies on a five-point Likert scale ("Never Use" to "Always Use"). Two factors with high reliability comprise the measure: Analytic ($\alpha = .85$) and Pragmatic ($\alpha = .75$). Higher use of both factors has been correlated with higher GPA in US students (Gavora et al., 2019). Sixteen items comprise the Analytic factor, theoretically capturing the frequency with which a student makes inferences and evaluations about a text. Sample items from this factor include: "While reading, I visualize descriptions in order to better understand the text." and "While I am reading, I reconsider and revise my background knowledge about the topic, based on the text's content." Six items comprise the Pragmatic factor, capturing practical ways to find and remember information from reading. Sample items from this factor include: "When I am having difficulty comprehending a text, I re-read the text." and "I make notes when reading in order to remember the information."

The second measure, the Contextualized Reading Strategy Survey (CReSS) is a self-report measure of reading strategy use where participants select whether they tend to use or do not tend to use such strategies when reading (Denton et al., 2015). The measure contains 49 items comprising a mix of empirically beneficial reading strategies, along with theorized less effective strategies common to readers of lower proficiency to reduce the potential for endorsement of all items. The scale is contextualized, asking students to self-report their reading strategies within common learning situations (e.g., "Which of these things do you do to help you understand a story while you are reading or after reading?") instead of checking items from a list. The scale's authors argue such contextualization helps students "more accurately recall their use of strategies" and "might encourage more careful consideration" of items versus lists which may be glossed over (Denton et al., 2015, p. 85). The measure consists of a theorized four factors, each with high reliability: Integration ($\alpha = .90$), Note-taking ($\alpha = .87$), Regulation ($\alpha = .81$), and Help-seeking ($\alpha = .71$).

The first subscale, Integration, focuses on making connections with prior knowledge and summarizing concepts, consists of 14 items. Examples include: "I try to make mental pictures of the information in the story while I read." and "While I am reading, I think about how the parts of the story go together." Six items make up the Regulation subscale, centered around re-reading challenging sections of a text and paying closer attention when perceiving difficulty. Sample items include: "If the text is hard to understand, I pay closer attention to what I am reading." and "I reread the difficult part over and over until it makes sense." Three items comprise the Note Taking subscale, focusing on proclivity to take notes, with items such as "If I can, I highlight or underline important ideas in the text." and three items comprise the Help Seeking subscale, focusing on the tendency to seek clarification or help from peers or instructors, exemplified by items such as "I ask the teacher or a friend for help."

Procedure

Study 2 followed the same method as the experiment reported in Chapter 2 (Study 1), with slight changes to the materials and some additional items. The new differences from the earlier procedure are that (a) participants were assigned to read one of two longer text passages, (b) after they reported the cues used to estimate their comprehension, they were asked to report the valence of each cue, and (c) they completed two individual difference measures that targeted their reading strategies.

At the start of the experiment, participants were randomly assigned to read one of two informative essays adapted from previous state standardized tests in literature arts, normed in the state of Texas. Two different essays were chosen to increase the generalizability of results. Passages were of similar lengths (620 versus 677 words) and were read in similar timeframes (*Mdn* = 1.82 minutes versus *Mdn* = 2.10 minutes). A Welch

Two Sample t-test indicated no significant difference in performance between the two essays, $t(228) = 1.57$, $p = 0.12$, with means of 71.67 (SD=24.60) and 66.24 (SD=28.49). Each passage presented information about a topic, either the confinement of animals in zoos or living alone as an adult, explaining relevant facts, and providing opinions from both sides of the issue.

After reading the essay, participants were asked a series of measures, beginning with an open-response item simply asking for the cues used to evaluate their level of understanding. After completing this item, participants were shown a list of the eight most common cues reported as influential to estimating comprehension. Participants self-reported these cues by checking the box next to the cue or leaving it blank. Next participants weighed the relevant weight of each cue in their overall estimate of comprehension. Then, participants reported their valence levels for every cue, whether reported as utilized or not. For each of the eight cues, participants rated from 0-10 the valence of each cue.

Under the impression that there would be a quiz at the end, participants next provided a judgment of comprehension, after having read the materials. *"Soon, you will be asked a total of 10 multiple choice questions covering the text you just read. Think about what you learned and estimate how many questions you believe you will answer correctly."* Judgment of comprehension was rated with a sliding bar recording the estimated number of correct responses in discrete integers from 0 to 10.

After these metacomprehension items, participants answered the two scales measuring self-reported use of typical reading strategies as a learner, then completed a comprehension exam.

RESULTS

In Study 1, cue usage was measured but cue valence was not. In Study 2, one of the main questions is what measuring cue valence adds to our understanding of metacomprehension. Table 16 displays the proportion of participants who reported using each cue along with the average valence rating for each cue. The most reported cues included experiencing difficulty during reading (difficulty), participants' perception of their ability to recall the passage (recall), and their confidence in explaining the passage's content (summarize). Two cues appeared to elicit strong directional feelings, on average. Most participants felt very comfortable in their ability to learn from reading (M = 7.16, SD = 1.83), which should be expected from college-trained participants. Additionally, most participants felt that understanding the content from the essays was relatively easy (M = 3.11, SD = 2.33), where the item was coded from 0 = easy to 10 = difficult. Valences were more mixed among the remaining six cues, with standard deviations above two and means hovering near the middle of each item's range.

Table 16.

Descriptive Statistics of Cue Use and Cue Valence

Cue	Percent Reporting Cue	Valence *M* (*SD*)
Self-efficacy	20.25	7.16 (1.83)
Prior Knowledge	25.32	4.38 (2.56)
Interest	29.54	5.77 (2.46)
Difficulty	43.88	3.11 (2.33)
Effort	29.54	4.79 (2.05)
Recall	40.51	5.00 (2.07)
Mental Imagery	32.49	5.91 (2.56)
Summarization	64.14	6.08 (2.33)

Looking at the differences in these measures between reading passages reveals more interesting details. Table 17 reports effect size difference between the two passages on each measure as well as whether independent t-tests revealed statistically significant differences. Across the two passages, participants reported similar proportions for most cues, indicating comparable metacognitive processes and strategies. There were only a few exceptions: more people reported using mental imagery as a cue for the zoo essay, and people rated themselves as having higher prior knowledge for the loneliness essay. However, as these differences were few and relatively small, all subsequent analyses will combine the data from the two readings passages.

Table 17.

Descriptive Statistics of Cue Use and Cue Valence

Cue	Proportion Reporting Cue *M (SD)*			Valence *M (SD)*		
	Zoo Essay	Alone Essay	Cohen's *d*	Zoo Essay	Alone Essay	Cohen's *d*
Self-efficacy	.19 (.39)	.22 (.41)	.071	7.02 (1.77)	7.31 (1.88)	.160
Prior Knowledge	.25 (.43)	.26 (.44)	.024	4.85 (2.48)	3.92 (2.56)	.368**
Interest	.32 (.47)	.28 (.45)	.090	5.52 (2.36)	6.01 (2.53)	.199
Difficulty	.46 (.50)	.42 (.50)	.090	3.39 (2.35)	2.84 (2.30)	.237
Effort	.32 (.47)	.28 (.45)	.090	4.91 (1.87)	4.68 (2.21)	.113
Recall	.42 (.50)	.39 (.49)	.055	4.97 (1.93)	5.03 (2.20)	.029
Mental Imagery	.39 (.49)	.26 (.44)	0.290*	6.15 (2.49)	5.68 (2.61)	.188
Summarization	.63 (.48)	.65 (.48)	.036	6.04 (2.24)	6.12 (2.42)	.032

* $p < .05$, ** $p < .01$, *** $p < .001$.

Cue Valence is Mostly Independent of Cue Use

To investigate the relationship between cue-use and cue-valence more generally (essay agnostic), I initially examined the proportion of participants who reported using each cue at different valence levels through visual exploration. I hypothesized a curvilinear or bimodal pattern, with reporting being more likely when a cue is experienced at either very low or very high levels, and less likely in the middle ranges. These findings are depicted in Figure 5. Despite a priori reasons to anticipate bimodal or curvilinear relationships, neither type of relationship appears to exist. Rather, Figure 5 shows that for the majority of the cues, valence is unrelated to cue usage. There were three exceptions: prior knowledge, interest, and mental imagery. In each of these cases, the higher the valence, the higher the reported use. Since none of the panels display a strongly curvilinear or bimodal pattern, I proceeded to conduct a logistic regression analysis for each cue, predicting cue usage based on rated valence, which confirms the patterns identified in Figure 5. The outcomes of these analyses are reported in Table 18.

Figure 5.

Logistic Regression Predicting Cue Use from Cue Valence at each Level of Cue Valence (0-10)

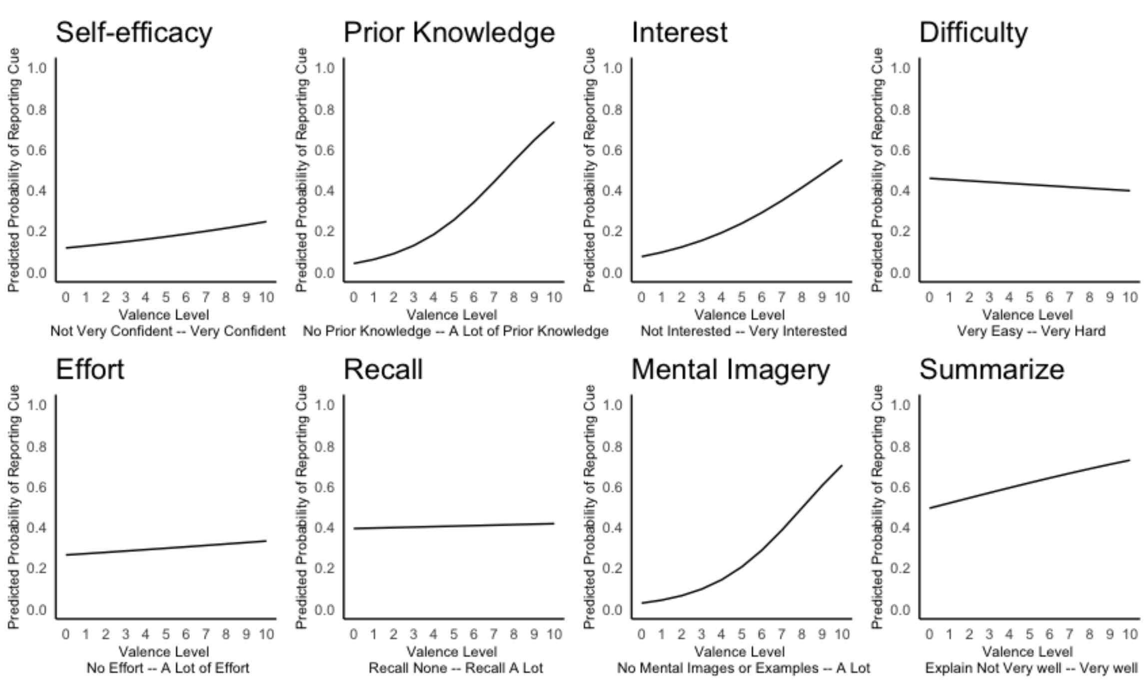


Table 18.

Logistic Regression Predicting Cue Use from Cue Valence

Cue	Estimate (Probability of Reporting Cue)	SE	Sig.	CI (lower)	CI (upper)
Self-efficacy	.09	.09	.326	-.09	.28
Prior Knowledge	.42	.07	<.001***	.28	.57
Interest	.27	.07	<.001***	.14	.41
Difficulty	.02	.06	.659	-.14	.09
Effort	.03	.07	.638	-.10	.17
Recall	.01	.06	.877	-.12	.14
Mental Imagery	.44	.08	<.001***	.30	.60
Summarization	.10	.06	.084	-.01	.22

*$p < .05$. **$p < .01$. ***$p < .001$.

Cue Valence Measures Are Related but Distinct

Participants rated their experienced valence on eight different metacognitive cues. Intuitively, some cues should be closely related to one another. Recalling facts or details about a passage, utilizing mental images to understand a concept, and explaining a concept (perhaps with facts or details recalled and visual analogies drawn), are each tools we use to help develop an understanding of what we are learning. As shown in Table 19, reported valences reflect these relationships, as recall is strongly associated with mental imagery (r = .37) and summarization (r =.59), and mental imagery is strongly associated with recall (r =.42) and summarization (r =.34). However, while these cues are very similar, their associations do not rise to the level of appearing to share the same underlying concept. Participants seem to understand some level of distinction between them.

Similarly, at first glance, effort and difficulty might seem interchangeable. However, the data reveal a different story—effort was only weakly correlated with difficulty (*r* = -.16). Sometimes difficulty understanding a passage was associated with putting in high effort, other times it was met with low effort to understand.

The remaining cues, such as interest, prior knowledge, and self-efficacy, tend to associate with other cues (for example, when a person is more interested, they might also report higher levels of recall, mental imagery, and summarization). Nonetheless, none of these cues appear to be entirely subsumed by another. The results support this hypothesis, revealing the strongest association between interest and prior knowledge (*r* = .43), which implies that individuals with an interest in the topic also tended to report higher levels of prior knowledge.

The strongest relationships occur between summarization and three other cues (interest, difficulty, and recall). The more interested participants were in the topic the higher their reported levels of summarization (*r* =.52), a similar association occurred with recall (*r* =.59). The inverse occurred with effort, where the more difficult participants found the article, the lower their reported levels of summarization (*r* = -.57). While these associations are strong, and significant, these cues are generally distinct and can stand on their own based upon their everyday definitions. Thus, all cues appear to be independent, with no cue pairs overlapping at a high level on both a conceptual and statistical association manner.

Table 19.

Pearson Correlations of Cue Valences

Cue	1	2	3	4	5	6	7
1. Self-efficacy							
2. Prior Knowledge	.10						
3. Interest	.31***	.43***					
4. Difficulty	-.35***	-.26***	-.41***				
5. Effort	.10	.19**	.34***	-.16*			
6. Recall	.30***	.38***	.43***	-.37***	.42***		
7. Mental Imagery	.14*	.30***	.35***	-.25***	.37***	.42***	
8. Summarize	.44***	.34***	.52***	-.57***	.26***	.59***	.34***

*$p < .05$. **$p < .01$. ***$p < .001$.

Cue Valence is a Stronger, More Consistent Predictor of Judgments of Comprehension Than Cue Use

Collecting cue-valence separate from cue-usage also allows me to test whether participants have insight into what cues are influencing their judgments of comprehension. If reported cue usage does not relate to judgment of comprehension, but cue valence does, then this would imply that participants are unable to explicitly report when a cue influences their judgments.

For each cue, I examined how its valence and use jointly predicted participants' judgments of comprehension by conducting a series of eight nested regression analyses, one analysis for each cue. At Step 1, I examined whether cue-use was related to judgment of comprehension. This allows me to test whether reporting a cue is associated with changes to judgments of comprehension. At Step 2, I added cue valence; this allows me to test whether the addition of valence as a feature explains more of the association.

The results are summarized in Table 20. At Step 1, results show that only the use of mental imagery was related to judgments of comprehension: when mental imagery was reported as being used, judgments of comprehension were higher, compared to when mental imagery cue use was not reported. This is considerably different from the results of Study 1 (Chapter 2), where many other reported cues were significantly related to judgments of comprehension.

At Step 2, results show that over and above reported cue use, the rated valence of every single cue was significantly related to judgments of comprehension. The higher the self-efficacy, prior knowledge, interest, effort, recall, mental imagery, and explanation, the higher the judgments of comprehension. The higher the difficulty experienced, the lower the judgments of comprehension. Specifically, the addition of valence to the model improved the adjusted R-squared value; however, the magnitude of this improvement varied across different cues, as shown in Table 21. While the increase was negligible for prior knowledge ($\Delta R^2 = 0.014$, $p = 0.040$) and effort ($\Delta R^2 = 0.021$, $p = 0.014$), it was substantial for recall ($\Delta R^2 = 0.242$, $p < 0.001$) and summarization ($\Delta R^2 = 0.381$, $p < 0.001$).

These results suggest that learners may be attuned to their experiences with metacomprehension cues, and that these cue experiences influence their judgments of comprehension; yet at the same time, learners are do not explicitly report using these cues without the aid of additional prompts.

Table 20.

Multiple Independent Regression Analyses Predicting Judgments of Comprehension

Cue		Step 1		Step 2	
		β (SE)	p	*β* (SE)	p
Self-efficacy	Intercept	53.7 (1.41)***	<.001	53.93 (1.32)***	<.001
	Use	0.12 (3.14)	.971	-1.01 (2.93)	0.971
	Valence	--	--	7.12 (1.18)***	<.001
Prior Knowledge	Intercept	52.45 (1.45)***	<.001	53.10 (1.47)***	<.001
	Use	5.05 (2.88)	0.081	2.48 (3.12)	0.428
	Valence	--	--	2.81 (1.36)*	0.040
Interest	Intercept	52.79 (1.50)***	<.001	54.16 (1.40)***	<.001
	Use	3.16 (2.76)	0.253	-1.46 (2.65)	0.581
	Valence	--	--	7.75 (1.21)***	<.001
Difficulty	Intercept	53.76 (1.68)***	<.001	53.95 (1.55)***	<.001
	Use	-0.07 (2.54)	0.977	-0.51 (2.34)	0.826
	Valence	--	--	-7.66 (1.16)***	<.001
Effort	Intercept	54.79 (1.50)***	<.001	54.85 (1.48)***	<.001
	Use	-3.60 (2.76)	0.193	-3.81 (2.73)	0.164
	Valence	--	--	3.09 (1.25)*	0.014
Recall	Intercept	54.14 (1.64)***	<.001	54.22 (1.42)***	<.001
	Use	-1.01 (2.57)	0.694	-1.21 (2.24)	0.590
	Valence	--	--	9.58 (1.10)***	<.001
Mental Imagery	Intercept	51.77 (1.52)***	<.001	52.81 (1.55)***	<.001
	Use	6.02 (2.67)*	0.025	2.82 (2.88)	0.328
	Valence	--	--	3.67 (1.35)**	0.007
Summarization	Intercept	51.76 (2.10)***	<.001	53.58 (1.66)***	<.001
	Use	3.06 (2.62)	0.245	0.23 (2.08)	0.912
	Valence	--	--	12.05 (1.00)***	<.001

Note: valence measures scaled and centered.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 21.

Adjusted R-Squared Difference from Step1 to Step 2

Cue	Step 2 Minus Step 1	F	p
Self-efficacy	.131***	36.285	<.001
Prior Knowledge	.014 *	4.279	.040
Interest	.145***	40.939	<.001
Difficulty	.153***	43.312	<.001
Effort	.021*	6.149	.014
Recall	.242***	75.644	<.001
Mental Imagery	.026**	7.378	.007
Summarization	.381***	145.82	<.001

*$p < .05$. **$p < .01$. ***$p < .001$.

Having established the utility of cue valence, I next examined how the cues collectively relate to judgments of comprehension. Because the prior analysis showed that cue use was not related after controlling for cue valence, I entered only the cue valences as predictors. Table 22 summarizes the results of this regression analysis, showing that the only cues that predicted judgments of comprehension, when controlling for every other cue, were recall and explain; the more participants felt like they could recall or explain the passage, the higher their judgments of comprehension. Each of these cues reflect deeper cues. None of the other cues were significantly related to judgments of comprehension, when controlling for every other cue.

Taken together, these results imply that (a) cue valences matter for judgments of comprehensions, but that (b) participants are insensitive to reporting the cues that they are likely actually using.

Table 22.

Multiple Regression Results Predicting Judgments of Comprehension from Cue Valence

Predictor	*β*	SE	*β* 95% CI [LL, UL]	p
(Intercept)	53.73*	0.96	[51.83, 55.62]	<.001
Valence Self-efficacy	1.59	1.09	[-0.57, 3.74]	0.148
Valence Prior Knowledge	-2.11	1.11	[-4.30, 0.08]	0.059
Valence Interest	1.99	1.25	[-0.46, 4.45]	0.112
Valence Difficulty	-0.59	1.20	[-2.96, 1.78]	0.627
Valence Effort	-1.46	1.11	[-3.65, 0.73]	0.190
Valence Recall	4.61**	1.32	[2.01, 7.22]	0.001
Valence Mental Imagery	-0.48	1.12	[-2.68, 1.72]	0.667
Valence Summarization	8.52**	1.45	[5.66, 11.37]	<.001

Note: valence measures scaled and centered.
$*p < .05$. $**p < .01$. $***p < .001$.

Analytic and Integration Reading Strategies are Related to Cue Valence and Judgments of Comprehension

Two measures of general reading strategies—MRSQ and CReSS—were collected from participants; each measure consisted of six subscales. Descriptive statistics for these subscales are reported in the bottom row of Table 23. Table 23 also shows Pearson correlations between the subscales of the Metacognitive Reading Strategies Questionnaire (MRSQ), Cognitive Regulation Strategies Scale (CReSS), reported cue valence, and judgments of comprehension. Only the Analytic and Integration subscales were related to cue valences and to judgments of comprehension. Higher Analytic reading skills were related to participants reporting higher self-efficacy, interest, mental imagery, and summarization, and lower perceived difficulty. Those who use more integration strategies when reading gave higher valence ratings to mental imagery and summarization. Each of these cue's valences were related to judgments of comprehension. There was just one other cue valence that was related to judgments of comprehension, but which was not related to any of the general reading strategies—recall. Finally, both of these subscales (Analytic and Integration) were also positively related to judgments of comprehension.

Table 23.

Correlation Matrix of Cue Valences, Reading Strategy Subscales, and Judgments of Comprehension

	MRSQ		CReSS				
	Analytic	Pragmatic	Integration	Regulation	Note Taking	Help Seeking	JOC
Self-efficacy	0.23*	0.03	0.17	0.14	0.07	-0.03	0.37***
Prior Knowledge	0.19	-0.05	0.06	0.05	0.07	-0.02	0.17
Interest	0.26**	0.03	0.15	0.13	0.06	-0.03	0.39***
Difficulty	0.22*	-0.03	0.15	0.13	-0.14	0.05	0.39***
Effort	0.19	0.07	-0.01	0.06	-0.07	0.05	0.16
Recall	0.21	0.10	0.08	0.06	0.08	0.05	0.49***
Mental Imagery	0.29***	0.09	0.23*	0.16	0.15	0.13	0.22*
Summarization	0.26**	-0.00	0.22*	0.13	0.18	-0.09	0.62***
JOC	0.22*	-0.03	0.23*	0.13	0.13	0.04	
Mean (*SD*)	51.56 (7.56)	19.40 (4.19)	6.04 (2.57)	4.09 (1.53)	1.78 (1.11)	0.68 (0.85)	

Note. JOC = Judgment of comprehension.
*p < .05. **p < .01. ***p < .001.

Analytic and Integration Strategies Do Not Moderate the Relationship Between Cue Valence and Judgments of Comprehension

To examine whether Analytic and Integration reading strategies moderate the relationship between cue valences and judgments of comprehension, I conducted the following analysis for each cue that was related to judgment of comprehension: self-efficacy, interest, difficulty, mental imagery, and summarization.

For each cue, I conducted a linear regression predicting judgment of comprehension from the cue valence, analytic strategy score, integration strategy score, cue valence x analytic strategy, and cue valence x integration strategy. All variables were standardized (centered and scaled). The results are presented in Table 24.

Table 24.

Summary of Regression Analyses: Standardized Beta Coefficients and Significance Level – Metric (p-value)

Cue	Valence	Analytic	Integration	Valence x Analytic	Valence x Integration	Model Adj. R2
Self-efficacy	6.51	1.74	2.58	-0.05	1.17	.152
	(<.001)***	(0.191)	(0.049)*	(0.971)	(0.392)	(<.001)
Interest	6.85	1.10	2.94	0.75	-2.11	.175
	(<.001)***	(0.405)	(0.022)*	(0.547)	(0.136)	(<.001)
Difficulty	-6.97	1.51	2.72	-0.08	-0.72	.174
	(<.001)***	(0.251)	(0.034)*	(0.950)	(0.574)	(<.001)
Mental Imagery	2.90	2.04	2.79	0.81	2.09	.080
	(0.024)*	(0.144)	(0.041)*	(0.551)	(0.124)	(<.001)
Summarization	11.58	0.59	1.58	-0.45	0.16	.384
	(<.001)***	(0.601)	(0.158)	(0.669)	(0.885)	(<.001)

*$p < .05$. **$p < .01$. ***$p < .001$.

As the results show, there were no significant interactions. That is, the Analytic and Integration individual reading strategies did not moderate the relationship between cue valence and judgments of comprehension.

DISCUSSION

The Importance of Measuring Valence

The results above indicate that adding valence as a measure, on top of reported cues, adds value to understanding complex metacognitive processes. Valences were independent from one another (separate constructs), independent from pure reporting of cues, and these differences were not due to the materials participants engaged with. While valence levels do not directly predict whether or how a cue will be reported, for many cues, as valence levels increase, so does the likelihood the cue will be reported.

So, other than being distinct from cue reporting, what value do valences bring to the table? Examining the relationship with judgments of comprehension revealed that capturing valences helps us see relationships with judgments we would not otherwise see with just reported cues. When only unprompted cue reporting is used, participants lack the ability to recognize all the cues, and their level, that may be impacting their judgments of comprehension. Merely reporting a cue does not necessarily correlate with changes in comprehension judgments. However, adding specifically prompted valences, with levels, helps to capture additional information around the metacognitive processes of learners. In fact, each cue valence demonstrated a significant relationship with comprehension judgments. This implies that the rated valence of cues has a more direct influence on judgments of comprehension than mere cue usage.

Taken as a whole, the results of Study 2 indicate that readers are generally aware of how an individual reading session went (cue valences) but are typically unable to accurately report the cues that influenced this metacomprehension judgment. For instance, we can reflect upon the difficulty level of a reading task, the extent of our recall, the amount of mental imagery we employed, and so forth, in a way that correlates with our comprehension evaluation. However, when asked to directly name the cues that contributed to these evaluations, our awareness appears to falter. People simply aren't good at estimating what goes into their judgments of comprehension, but when asked about actual experiences with cues, these appear to be more valid.

The Role of Individual Differences

The importance of measuring valence is evident, yet the question remains: do individual differences also play a role? While this study does not offer conclusive evidence on how valence interacts with individual differences, it does underscore the influence of these differences on judgments of comprehension. Strategies known to enhance comprehension—such as re-reading, paying close attention during challenging sections, and linking concepts—appear to be effective. Notably, even after accounting for individual differences in metacognitive strategy use, valence metrics continue to be robust predictors of judgments of comprehension.

Continuing on the theme of individual differences, these findings also highlight how general reading strategies shape one's interaction with the material. Specifically, individuals employing more analytic and integrative strategies tend to be more confident and engaged in their reading, report lower levels of difficulty, and activate more effective deep-processing techniques like mental imagery and summarization. Interestingly, while these cues are linked to judgments of comprehension, they are not directly moderated by reading strategies.

This underscores the complexity of the metacognitive landscape, where multiple dimensions—including cues, metacognitive strategies, and comprehension outcomes—are intricately interrelated. By incorporating both CRESS and MRSQ scores as moderators, we gain a holistic view of how individual metacognitive strategy usage influences the interplay between cues and comprehension. The findings of Study 2 suggest that individuals considered to be ‘good’ metacognitive readers frequently employ deep strategies. However, these same individuals may not be as adept at explicitly identifying the cues that inform their judgments.

Future Directions

The findings of Study 2 offer valuable insights into the role of valence and individual differences, particularly in the realm of reading strategies. However, the complexities of metacognition extend beyond textual mediums. In the next phase of my research, I'll pivot to educational videos to explore the nuanced interplay between cues, valence, and a new variable of individual difference—interest. This shift allows me to examine whether interest levels, experimentally manipulated, can affect learners' cue recognition and judgments of comprehension. By transitioning from reading strategies to interest, I aim to expand the boundaries of my inquiry, capturing more dimensions of individual differences in metacognition.

Study 3

Building on the insights gained from Study 2, which explored the complex relationship between valence, cues, and individual differences in reading strategies, Study 3 takes a different tack by switching the medium from written essays to educational videos. This change enables a broader exploration of the kinds of cues that learners rely on, given that the two mediums offer different sets of cues to users (Caspi et al., 2005).

For instance, educational videos inherently offer rich visual imagery designed to facilitate comprehension. This wealth of visual cues could be more readily processed with less cognitive effort, thereby potentially shifting the dynamics of how learners assess their comprehension. However, the availability of a cue does not automatically translate into its recognition or utilization by the learner.

In addition to venturing into this different medium, this study opens up the possibility to examine a new variable for individual differences—interest. Interest can act as a potent agent of change in the learning process, affecting both motivation and engagement. In a departure from focusing solely on reading strategies in Study 2, the current study aims to investigate how learners' interest levels may influence their perception and utilization of cues in the context of video-based learning.

The study utilizes a between-participants design, where interest is experimentally manipulated by asking participants to rank video topics from most to least interesting. For half of the participants, I showed them their top-ranked video and for the other half (high interest condition), I showed their lowest-ranked video (low interest condition). This design allows for analysis of how interest levels may moderate the cues learners report and their subsequent judgments of comprehension. Drawing on previous work that underscores the importance of interest in learning and comprehension (Deci et al., 1991; Patall, 2013;

Renninger et al., 2014), the hypothesis is that patterns of cue use will vary depending on the level of interest in the topic at hand.

METHOD

Participants

Two-hundred-five undergraduate students participated in the study. After removing incomplete responses, the final sample contained 200 observations. Of the total, 83 from the University of Texas at Austin completed the survey in exchange for course credit and 117 completed the survey through the research website Prolific, in exchange for a nominal sum. All participants were college-aged (21.45 *M*, 2.64 *SD*) and while participants differed slightly demographically between the samples, together the sample represents a diverse range of backgrounds. The majority were female (59.50%) and white (45.50%), though a variety of students completed the study (21.50% Asian, 18.50% Latinx, 8.50% Black). Participants were randomly assigned to the high interest or low interest condition, where interest was defined by how highly they ranked a video topic.

Materials

Instead of using text passages, in Study 3, I presented participants with educational videos. The informational videos were segments from the television show "How It's Made", available on YouTube (How It's Made | Watch Full Episodes & More! - Science, n.d.). Each video consisted of a 5-minute segment of the show previously aired on television. Showing the manufacturing process of common consumer products, three topics were used as stimuli: how crayons are made, how hot sauce is made, and how mascara is made. Participants watched only one of the three videos. Videos were similar in length (306, 309, and 318 seconds (about 5 and a half minutes) in duration) and production quality.

Experimental Conditions

Participants began by ranking their level of interest in watching a video on one three different topics: how mascara is made, how crayons are made, or how hot sauce is made. *"Below you will find three different topics. Each topic has a corresponding video explaining how the product is made. Some topics you likely find more interesting than others. Please rank each topic below from most interesting (1) to least interesting (3)."* After ranking, participants were randomly assigned to watch either their top-rated choice (high interest condition) or their lowest rated choice (low interest condition).

Procedure

After random assignment into either the high-interest condition (watching the video rated by the participant as most interesting) or low-interest condition (watching the video rated as least interesting), participants watched their assigned video. Next, participants were asked the same metacognitive questions, in the same order, as those presented to participants in Study 2. Finally, participants answered several manipulation check questions followed by demographic questions. No comprehension exam was administered to participants.

Results

Manipulation Check: Interest Does Vary Between Conditions

First, as a manipulation check, I conducted a t-test comparing the average interest valence—rated after watching the video—in each condition. As expected, those in the high-interest condition reported significantly higher interest ratings (M = 6.54, SD = 2.08) than those in the low interest condition (M = 3.73, SD = 2.72), $t(197) = -8.25$, $p < .001$.

Few Differences Across the Videos

Table 25 displays the proportion of participants who reported each cue and the mean valence rating of each cue across the three different videos. Two one-way MANOVA tests were used to determine whether the three videos differed on either cue reporting or cue valence. The results indicated that the video watched did not have a significant effect on the combined valence measures, $F(16, 382) = .891$, $p = .58$. Similarly, there was no significant effect of the video watched on the reported cues, $F(16, 382) = 1.07$, $p = .38$. As any differences visible in the descriptive results in Table 25 are not substantial, all subsequent analyses will collapse across specific video.

Table 25.

Descriptive Statistics of Cue Use and Cue Valence by Video

Cue	Reporting Cue - Proportion (*SD*)				Valence Rating - *M* (*SD*)			
	Crayons	Hot Sauce	Mascara	Overall	Crayons	Hot Sauce	Mascara	Overall
Self-efficacy	.25 (.44)	.21 (.41)	.24 (.43)	.24 (.44)	7.50 (1.69)	6.85 (1.74)	7.40 (1.79)	7.26 (1.76)
Prior Knowledge	.15 (.36)	.13 (.34)	.13 (.34)	.14 (.34)	2.23 (2.39)	1.85 (2.00)	2.59 (2.59)	2.26 (2.37)
Interest	.33 (.48)	.35 (.48)	.17 (.38)	.28 (.45)	4.75 (3.05)	5.00 (2.82)	4.96 (2.69)	4.91 (2.83)
Difficulty	.47 (.50)	.35 (.48)	.44 (.50)	.42 (.50)	2.90 (1.98)	3.27 (2.40)	2.62 (2.28)	2.90 (2.24)
Effort	.20 (.40)	.29 (.46)	.15 (.36)	.21 (.41)	5.53 (2.38)	5.27 (1.97)	5.49 (2.27)	5.44 (2.21)
Recall	.23 (.43)	.21 (.41)	.30 (.46)	.25 (.43)	5.65 (2.24)	5.29 (2.40)	5.53 (1.96)	5.49 (2.18)
Mental Imagery	.58 (.50)	.61 (.49)	.62 (.49)	.61 (.49)	4.55 (2.97)	4.47 (3.06)	4.88 (2.86)	4.66 (2.95)
Summarization	.45 (.50)	.52 (.50)	.58 (.50)	.52 (.50)	6.05 (2.32)	5.66 (2.34)	5.71 (2.18)	5.80 (2.27)

Modest Correlations Between Valence and Use for Some Cues

As with Study 2, I conducted a logistic regression analysis for each cue, predicting cue usage from rated valence, condition, and the interaction between the two predictors. The results are presented in Figure 6 and Table 26. For many of the cues (self-efficacy, prior knowledge, interest, and summarization), a higher valence rating was associated with higher report of cue usage. A higher rating of difficulty was associated with lower cue usage. Ratings of effort, recall, and mental imagery were not associated with cue usage.

These results differ from that of Study 2, where only prior knowledge, interest, and mental imagery related to cue usage. In the present study, mental imagery did not relate to cue usage, which may be due to the fact that viewing a video may negate the need to rely on mental imagery. In general, however, the valence of a broader range of cues predicted cue usage in Study 3 than in Study 2. It is unclear if the difference is due to the medium (text vs. video) or content.

Figure 6.

Logistic Regression Predicting Cue Use from Cue Valence at each Level of Cue Valence (0-10)

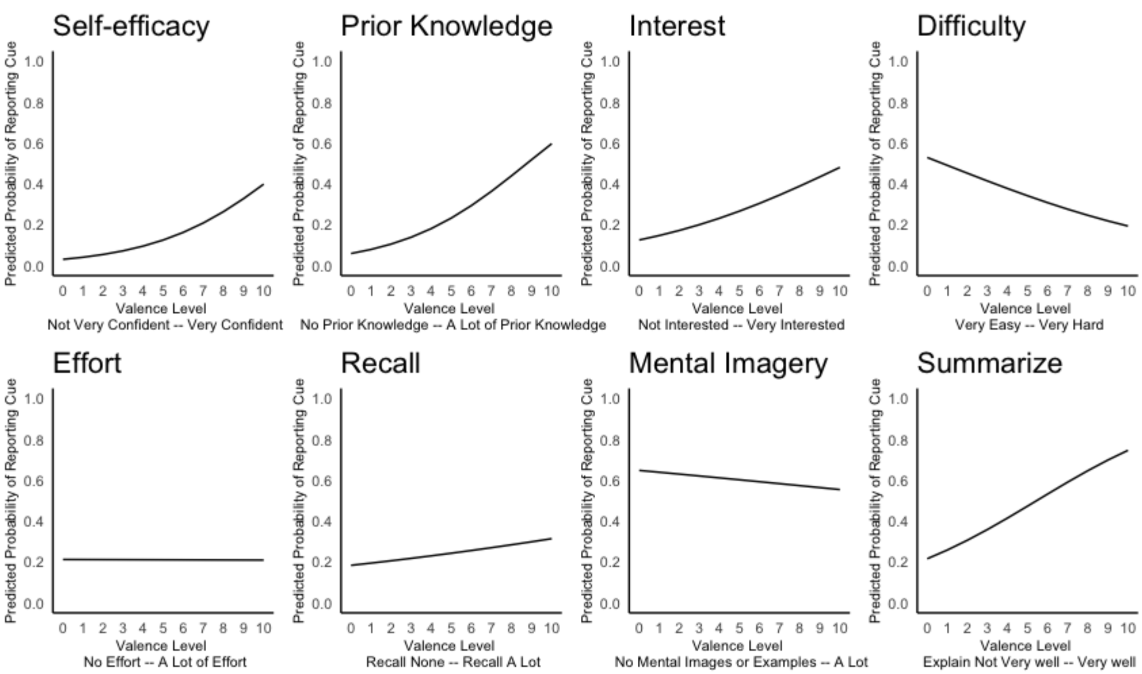


Table 26.

Logistic Regression Predicting Cue Use from Cue Valence

Cue	Estimate (Probability of Reporting Cue)	SE	Sig.	CI (lower)	CI (upper)
Self-efficacy	.31	.11	.004**	.01	.53
Prior Knowledge	.32	.08	<.001***	.16	.48
Interest	.19	.06	.002**	.07	.31
Difficulty	.15	.07	.022*	.02	.29
Effort	.00	.08	.983	-.16	.15
Recall	.07	.08	.349	-.08	.22
Mental Imagery	.00	.05	.425	-.14	.06
Summarization	.24	.07	<.001***	.11	.38

*$p < .05$. **$p < .01$. ***$p < .001$.

Cue Valence is a Stronger, More Consistent Predictor of Judgments of Comprehension Than Cue Use

In Study 2, when reading a text, reported cue usage did not relate to judgment of comprehension for seven of the eight cues. This implies that participants are unable to explicitly report when a cue influences their judgments while reading a text. Does this same pattern appear when watching a video? Similar to the analysis in Study 2, for each cue I examined how its valence and use jointly predicted participants' judgment of comprehension by conducting a series of nested regression analyses, one analysis for each cue. At Step 1, I examined whether cue use was related to judgment of comprehension, testing whether reporting a cue is associated with changes to judgments of comprehension. At Step 2, I added cue valence, testing whether the addition of valence as a feature explains more of the association.

The results are summarized in Table 27. At Step 1, results show that only the use of summarization was related to judgments of comprehension; when participants reported being able to summarize the content, judgments of comprehension were higher, compared to when summarization was not reported. The general lack of a relationship between reported cue-use and judgments of comprehension is similar to that found in Study 2, where only use of mental imagery was related to judgments of comprehension. This difference may relate to the learning materials. Both mental imagery and summarization of content are deeper cues and when reading a text, mental imagery can be the next step of processing to cue one that they understand, or fail to understand, a passage. When watching a video, mental imagery is already available from the content itself, potentially making summarization of the imagery the next step of processing to cue that you understand, or fail to understand, the content.

At Step 2, results show that over and above reported cue use, the rated valence of every single cue was significantly related to judgments of comprehension. The higher the self-efficacy, prior knowledge, interest, effort, recall, mental imagery, and explanation, the higher the judgments of comprehension. The higher the experienced difficulty, the lower the judgments of comprehension.

Table 27.

Multiple Independent Regressions Predicting Judgment of Comprehension

Cue		Step 1		Step 2	
		β (SE)	*p*	*β* (SE)	*p*
Self-efficacy	Intercept	56.08 (1.66)***	<.001	56.73 (1.61)***	<.001
	Use	4.13 (3.43)	.229	1.35 (3.38)	0.689
	Valence	--	--	5.71 (1.44)***	<.001
Prior Knowledge	Intercept	57.57 (1.56)***	<.001	58.13 (1.54)***	<.001
	Use	-3.87 (4.26)	0.365	-7.97 (4.36)	0.069
	Valence	--	--	4.72 (1.49)**	0.002
Interest	Intercept	57.38 (1.71)***	<.001	58.55 (1.58)***	<.001
	Use	-1.20 (3.26)	0.714	-5.47 (3.06)	0.076
	Valence	--	--	8.62 (1.37)***	<.001
Difficulty	Intercept	55.69 (1.91)***	<.001	57.01 (1.71)***	<.001
	Use	3.24 (2.94)	0.273	0.09 (2.66)	0.974
	Valence	--	--	-9.55 (1.32)***	<.001
Effort	Intercept	58.04 (1.63)***	<.001	58.03 (1.54)***	<.001
	Use	-4.70 (3.56)	0.188	-4.68 (3.36)	0.165
	Valence	--	--	7.00 (1.37)***	<.001
Recall	Intercept	55.87 (1.67)***	<.001	56.42 (1.18)***	<.001
	Use	4.73 (3.35)	0.159	2.51 (2.37)	0.292
	Valence	--	--	14.54 (1.03)***	<.001
Mental Imagery	Intercept	53.92 (2.30)***	<.001	53.64 (2.26)***	<.001
	Use	5.17 (2.96)	0.082	5.64 (2.91)	0.054
	Valence	--	--	4.07 (1.43)**	0.005
Summarization	Intercept	53.75 (2.08)***	<.001	57.59 (1.53)***	<.001
	Use	6.35 (2.88)*	0.029	-1.04 (2.15)	0.629
	Valence	--	--	14.58 (1.08)***	<.001

Note: valence measures scaled and centered.
*$p < .05$. **$p < .01$. ***$p < .001$.

As shown in Table 28, for every cue, by adding valence to the model the adjusted R-squared value improved. Some increases were pronounced, such as recall ($\Delta R^2 = 0.498$, p <.001) and summarization ($\Delta R^2 = 0.470$, p <.001). Further, those that appear negligible from raw values are still significant, such as self-efficacy ($\Delta R^2 = 0.070$, p <.001), prior knowledge ($\Delta R^2 = 0.043$, $p = .002$), and mental imagery ($\Delta R^2 = 0.034$, $p = .005$). In general, these results match those from text-based content: valence improved R-squared for all cues, but largest improvements were with deep cues: recall and summarization. These results again suggest that learners may be attuned to their experiences with metacomprehension cues, and that these cues influence their judgments of comprehension, yet at the same time, learners are not adept at explicitly reporting these cues without the aid of additional prompts.

Table 28.

Adjusted R-Squared Difference from Step to Step 2

Cue	Step 2 Minus Step 1	F	p
Self-efficacy	.070***	15.851	<.001
Prior Knowledge	.043**	9.974	.002
Interest	.163***	39.489	<.001
Difficulty	.207***	52.700	<.001
Effort	.112***	26.037	<.001
Recall	.498***	199.04	<.001
Mental Imagery	.034**	8.127	.005
Summarization	.470***	183.34	<.001

**p* < .05. ***p* < .01. ****p* < .001.

The previous analysis focused on one cue at a time. Next, I examined how the cues collectively relate to judgments of comprehension in a multiple regression. Because the prior analysis showed that cue use was not related after controlling for cue valence, I entered only the cue valences as predictors. Table 29 summarizes the results of this regression analysis. In Study 2, results showed that the only cues that predicted judgments of comprehension, when controlling for every other cue, were recall and explain; the more participants felt like they could recall or explain the passage, the higher their judgments of comprehension. In Study 3, similar patterns appeared. Again, the more participants felt like they could recall or explain the passage, the higher their judgments of comprehension, on average. However, in Study 3, a third cue was also a significant predictor: mental imagery. The more mental images or examples participants felt they generated, the lower the judgment of comprehension. This trend was also seen in Study 2 but was not significant in the multiple regression analysis.

In Study 3, as in Study 2, the stepwise regression and multiple regression results imply that (a) cue valences matter for judgments of comprehension, but that (b) participants are very insensitive to reporting the cues that they are likely actually using.

Table 29.

Multiple Regression Results Predicting Judgments of Comprehension from Cue Valence

Predictor	*β*	SE	*β* 95% CI [LL, UL]	p
(Intercept)	57.05***	0.94	[55.20, 58.90]	<.001
Valence Self-efficacy	-1.34	1.12	[-3.56, 0.88]	0.236
Valence Prior Knowledge	1.63	0.97	[-0.27, 3.54]	0.093
Valence Interest	1.45	1.12	[-0.75, 3.66]	0.195
Valence Difficulty	-2.10	1.20	[-4.46, 0.26]	0.081
Valence Effort	1.07	1.11	[-1.12, 3.25]	0.337
Valence Recall	7.61***	1.53	[4.59, 10.63]	<.001
Valence Mental Imagery	-2.47*	1.06	[-4.55, -0.39]	0.02
Valence Summarization	7.89***	1.55	[4.84, 10.94]	<.001

Note: valence measures scaled and centered.
*$p < .05$. **$p < .01$. ***$p < .001$.

Interest is Related to Cue Valence and Judgments of Comprehension

In Study 2, individual reading strategy difference measures were compared to cue valence. In Study 3, a key question is whether interest in the topic affects cue valences and metacomprehension. Independent t-tests were conducted to investigate whether they differed between condition. The descriptive and inferential statistics are reported in Table 30. As would be expected, the two groups differed on interest. They also differed on prior knowledge—potentially the topics people found more interesting to begin with were also the ones where they had more interest. More interestingly, those in the higher interest condition also reported greater amounts of recall, and marginally higher effort and mental imagery. These results suggest that interest may drive deeper effort and strategy use.

Finally, I examined whether interest conditions differed in mean judgments of learning. Those in the high interest condition gave higher judgments of comprehension (*M*

= 58.9, *SD* = 20.4) than those in the low interest condition (*M* = 55.7, *SD* =20.4), *t*(198) = 1.10, *p* = .271, Cohen's *d* = 0.158, though the difference was not significant.

Table 30.

Cue Valences by Condition

Cue	High Interest *M (SD)*	Low Interest *M (SD)*	*t* (df)	*p*	Cohen's *d*
Self-efficacy	7.12 (1.70)	7.36 (1.81)	-.959 (198)	.339	.136
Prior Knowledge	3.02 (2.68)	1.70 (1.94)	3.84 (198)	<.001	.579
Interest	6.54 (2.08)	3.73 (2.72)	8.28 (198)	<.001	1.14
Difficulty	2.95 (2.25)	2.87 (2.24)	.249 (198)	.804	.036
Effort	5.75 (2.01)	5.21 (2.32)	1.76 (198)	.081	.246
Recall	5.86 (2.12)	5.22 (2.19)	2.08 (198)	.039	.296
Mental Imagery	5.11 (3.05)	4.33 (2.84)	1.84 (198)	.068	.266
Summarization	6.04 (2.21)	5.62 (2.30)	1.30 (198)	.194	.186

Interest Does Not Moderate the Relationship between Cue Valence and Judgments of Comprehension

To examine whether interest condition moderated the relationship between cue valences and judgments of comprehension, a linear regression was conducted for each cue, predicting judgment of comprehension (standardized) from cue valence (standardized), interest condition (0 = low interest, 1 = high interest), and the interaction between the two. The results are summarized in Table 31. There was not a single significant interaction.

Table 31.

Summary of Regression Analyses: Standardized Beta Coefficients and Significance Level – Metric (p-value)

Cue	Valence	Interest Condition	Valence x Interest Condition	Model Adj R2
Self-efficacy	5.17 (0.004)**	4.12 (.148)	2.06 (.475)	.078 (<.001)
Prior Knowledge	5.17 (.026)*	1.14 (.706)	-2.46 (.418)	.025 (.046)
Interest	8.79 (<.001)***	-7.02 (.031)*	2.72 (.419)	.161 (<.001)
Difficulty	-9.87 (<.001)***	3.59 (.172)	0.68 (.797)	.212 (<.001)
Effort	7.16 (<.001)***	1.58 (.575)	-0.73 (.802)	.104 (<.001)
Recall	14.82 (<.001)***	-1.01 (.634)	-0.33 (.878)	.498 (<.001)
Mental Imagery	3.72 (0.060)	2.24 (.447)	0.10 (.974)	.024 (.050)
Summarization	13.62 (<.001)***	0.53 (.802)	2.02 (.346)	.489 (<.001)

**p* < .05. ***p* < .01. ****p* < .001.

DISCUSSION

The results from Study 3 reinforce and extend the findings of Study 2 by emphasizing the predictive power of cue valence across both text-based and video-based educational mediums. While cues such as self-efficacy, prior knowledge, interest, and summarization were reported more with higher valence ratings, difficulty was negatively associated with cue usage. More notably, the valence of every single cue was significantly related to judgments of comprehension, independently of whether the cue was explicitly reported or not. These results are critical in asserting the overriding influence of cue valence over mere cue reporting in shaping judgments of comprehension.

In line with Study 2, deeper cues like recall and summarization were especially predictive of judgments of comprehension. This suggests that deeper cues may be more reliable indicators of comprehension across different learning contexts. Such cues are particularly useful given that learners are not adept at explicitly reporting the cues they use, as evidenced by the lack of a strong relationship between reported cue use and judgments of comprehension.

The nature of the educational medium seems to impact cue usage, albeit it's unclear to what extent. For instance, while mental imagery did not relate to cue usage in the video-based Study 3, it was a significant predictor in the text-based Study 2. This differential outcome could signify that visual mediums may negate the need for learners to rely on mental imagery as an additional cognitive resource. However, the current study design does not enable a direct comparison between text and video; hence, future research should aim to disentangle any medium-specific effects on cue usage and valence.

While individual reading strategies were the focus in Study 2, Study 3 examined the role of interest in learning outcomes. Higher interest levels were associated with greater amounts of recall and marginally higher effort and mental imagery. However, interest did not moderate the relationship between cue valence and judgments of comprehension. This lack of moderation indicates that while interest may contribute to the depth of cognitive engagement, it doesn't change how cue valence influences one's judgments about what they have learned. This finding complicates the notion of tailoring educational content based solely on initial interest levels, as interest alone may not significantly influence the valence or effectiveness of metacognitive cues.

The results highlight that instructional design should not merely focus on initial interest as a magical solution for deeper cognitive engagement or accurate metacognition. It is not merely enough to match educational material with learners' professed interests.

Given the findings, educators should also focus on fostering conditions that promote deeper cues, such as summarization and recall. These cues, strongly related to both cue valence and judgments of comprehension, could prove invaluable for educational outcomes.

The prominence of cue valence across both studies underlines the need for further research on its intricate role in educational contexts. Specifically, it would be valuable to explore the efficacy of interventions designed to make learners more aware of deeper cues and their valence. This could help learners become more attuned to effective strategies for comprehension, transcending individual differences and medium-specific limitations.

SYNTHESIZING STUDIES 2 AND 3

The principal findings from Studies 2 and 3 converge on the pivotal role of cue valence in shaping judgments of comprehension. Across varied educational settings, whether text-based or video-based, valence stands out as a more consistent and reliable predictor of judgments of comprehension compared to the mere use of reported cues. Consequently, any metacomprehension study that fails to account for the valence of cues—essentially, their directional and magnitude impact—is missing out on capturing the full scope of metacognitive activity.

Both studies shed light on the role of individual differences—specifically, reading strategies and personal interest—in modifying cue valence. While interest seemed to induce learners to engage in deeper or more effective learning strategies, these factors did not alter the foundational relationship between cue valence and judgments of comprehension. This reinforces the idea that while strategies and interest are influential, they are not moderators of the core relationship.

Consistent with Study 1 and prior research, deeper strategies offer dual benefits: they enhance comprehension and provide learners with more effective cues for metacomprehension. This suggests a clear educational pathway: encourage students to engage in deeper cognitive strategies that naturally give rise to beneficial cues.

The findings from Studies 2 and 3 have made it evident that individual differences—while impactful in other areas—do not serve as moderators between cue valence and judgments of comprehension. This raises an intriguing question: if individual differences aren't moderating this relationship, could contextual factors play a more decisive role? In other words, while individual learners' characteristics may not alter the fundamental relationship between cue valence and judgments of comprehension, perhaps the settings or conditions under which learning occurs could have such an impact. Consequently, as we shift our focus from individual differences to situational variables, the next chapter will aim to explore how different learning contexts might not only influence the reporting of metacognitive cues but also potentially moderate the relationship between cue valence and judgments of comprehension.

CHAPTER 4: CONTEXT AND METACOMPREHENSION CUES

Context profoundly influences learning outcomes. From physical surroundings (Tessmer & Richey, 1997) to psychological and motivational elements (Boekaerts, 2010), contextual factors shape the cues learners attend to and the quality of learning achieved. Chapters 2 and 3 of this series revealed that while learners display idiosyncratic cue usage, cues and judgments of comprehension are malleable, influenced by both stimulus changes and individual differences. Building on these insights, the current experiment aims to examine how altering the context of the learning task impacts the metacognitive cues that participants report. The methodology replicates prior studies, consisting of new material learning, metacomprehension rating, cue reporting, and valence reporting. The distinguishing factor here is the introduction of specific learning goals and a comprehension exam.

One promising avenue for contextual influence is the paradigm of "learning by preparing to teach." Changing the learning motivation can shift metacognitive attention, as evidenced by learners focusing on different types of cues depending on whether they prepare for a comprehension exam or one based on specific details (Thiede et al., 2011). The task at hand—be it a final exam, written report, or presentation—often guides learners' goals, which in turn could alter both cognitive and metacognitive processes. The notion is that specific learning outcomes might lead to more effective learning experiences.

Particularly effective for when you are trying to understand something, focusing your learning on the goal of being able to explain it to others can lead to more learning than simply aiming to learn the material for yourself (Kornell, 2018). Defined as studying learning material with the expectation of teaching or providing explanations, prior research has shown that students can learn better if they imagine themselves preparing to teach information compared to if they are just preparing to take a test (Fiorella & Mayer, 2013;

Nestojko et al., 2014). There is something about having to not only map concepts and connect them to your own knowledge but making additional connections and analogies to ensure they can be comprehended by others that seems to make your learning deeper. This aspect of learning by preparing to teach leads to not only more thorough free recall of content but also enhanced performance on a test compared to those in a standard learning condition (Nestojko et al., 2014).

Thus far, however, the literature on learning by preparing to teach or learning by teaching (Fiorella & Mayer, 2014, Nestojko et al., 2014) has only examined cognitive consequences (i.e., what leads students to learn more), but not the metacognitive consequences (i.e., how learners experience different metacognitive cues, and how these cues relate to judgments of comprehension). This deeper level of engagement could extend to metacognitive processes, influencing both the cues learners use for self-evaluation and how these cues relate to judgments of comprehension.

In the present set of studies, I am building on this prior research to examine whether people are thinking about the content differently when they have these different goals in mind. By altering the instructions presented to participants, setting the context for learning, participants with the goal of "preparing to teach" may engage with the material in different ways than those preparing for a standard test. Those "preparing to teach" may be more aware of their own understanding, as they will need to ensure they fully comprehend the material before teaching it to others. These students may be more likely to develop mental imagery and engage in self-questioning or explaining the content to themselves to check their understanding. Further, those in the "prepare to teach" condition may also elaborate more and spend more time making connections between different ideas in anticipation of having to clarify complex or confusing concepts to their "students".

Such deeper or more focused instances of elaboration and comprehension monitoring should manifest as stronger correlations between cue valence and judgments of comprehension for cues typically associated with deeper processing, such as explanation and mental imagery. Similarly, these cues connected with deeper processing should be more frequently reported in the "prepare to teach" condition. Consequently, changing the context may not only influence patterns of cue reporting and valences but also affect overall judgments of comprehension, even though the actual learning task remains the same.

In this chapter, I aim to probe whether learners engage differently with material when given distinct learning goals. By varying the context through instructional adjustments, I hypothesize that those "preparing to teach" will engage more deeply with the content. This deeper engagement should, in turn, lead to a stronger correlation between cue valence and judgments of comprehension for cues associated with deeper cognitive processes, like explanation and mental imagery.

The current chapter seeks to extend our understanding from purely cognitive consequences to metacognitive processes. If successful, this research could offer educators a simple yet effective instructional tweak to improve both learning and metacognition in real-world educational settings.

Present Studies

In Studies 4 and 5, a "prepare to teach" condition was compared against a more traditional "prepare to take a test" control condition. In Study 4, those in the prepare to teach experimental condition were told that they would video themselves giving the lecture after studying. In Study 5, instead of preparing for a video lecture, participants in the prepare to teach condition were told to draft a lesson plan for someone else to use to teach future students. Each experiment used a similar paradigm and materials to those used in

Study 2. However, I made a simple modification to the instructions presented before participants started reading: they were instructed to either learn the materials with the goal of preparing for a test or preparing to teach the material to future learners. My analyses will focus on answering:

> 1. Do metacognitive cue valences differ significantly between the prepare to teach condition(s) and control condition?
>
> 2. Do metacomprehension judgments and test score differ significantly between the prepare to teach condition(s) and control condition?
>
> 3. Does condition moderate the relationship between cue valence and metacomprehension judgment?
>
> 4. Does condition moderate the relationship between cue valence and test score?

Study 4

METHOD

Participants

Two-hundred-seventy-nine undergraduate students completed the study in exchange for course credit. After removing incomplete responses, the final sample contained 263 observations. Participants were recruited from a wide subject pool across a variety of majors (48.29% STEM, 18.25% social sciences). A majority female (73.76% sample, 54.4% student body), and white (38.78% sample, 38.9% student body), the sample roughly resembled the student body demographics of the large public university at which it took place (Latinx 14.83% sample, 23.4% student body; Asian 27.38% sample, 20.2% student body; Black 5.32% sample, 5.3% student body). The age of participants reflected the university environment as well (20.21 *M*, 2.07 *SD*).

Materials

A short essay, expressing the author's opinion about zoos, was read by participants. This essay, along with six comprehension questions, were obtained from previously released state standardized exam materials, normed in the state of Texas. The essay consisted of 677 words should be easily read and understood by a high school graduate.

Experimental Conditions

Participants were randomly assigned to one of two conditions: learning the material by preparing to teach or learning the material by preparing for a test. This phrasing of the goal for learning, setting the context for the learning task, was the only difference between conditions. Those in the prepare to teach condition were instructed to learn the material with the intention of designing and recording a written summary which would be used to teach another undergraduate student the material. Those in the prepare for a test condition were instructed to learn the material with the intention of later taking a comprehension test over the material.

The text of the instructions in the prepare to teach (record lesson) condition was: *"In this experiment, you will read a short expository essay and then be asked to teach the material that you learned. You will have 6 minutes to read the passage. Then, we will ask you to record yourself giving a lecture, as if you were teaching an undergraduate student. This lecture should include as much of the original detail of the passage as possible, and make it so a viewer can comprehend the author's position. We will use this videotaped lecture to teach the passage to another participant in the study, who will then be tested on the passage. As you are reading the passage, please DO NOT take any notes."*

The text of the instructions in the prepare for test condition was, *"You may recall the information in any format within the box below (e.g., paragraph form, bullet point, etc.). The goal is to remember as much as you can. You will have 6 minutes to read the passage. You will then be given a memory and comprehension test on the passage, as if you were a student being tested as part of a class. You will be tested on how much of the original detail of the passage you can remember, and whether you comprehend the author's position. Please DO NOT take any notes."*

Procedure

Table 32.

Instructions for Participants

Phase	Prompt
Read text	In the following section you will be given a text passage to read. You will have 6 minutes total to study the passage.
Report metacomprehension	How did you decide whether you understood the passage? What thoughts did you have that helped you gauge how well you understood the text? (additional questions follow)
Distractor task	You have five minutes to answer the questions below. The time remaining is shown below. You will not be able to move to the next question until the time is up. From memory, please list as many countries as you can in the box below. From memory, please list as many colors (e.g., chartreuse) as you can in the box below.
Additional instructions	Next, you will complete a memory and comprehension test of the essay you read. Please answer the questions to the best of your ability.
Free recall	Please type as much information (e.g., details, ideas, messages) from the passage as you can recall from the essay. You may recall the information in any format within the box below (e.g., paragraph form, bullet point, etc.). The goal is to remember as much as you can.
Comprehension test	Please answer the following questions concerning the text passage you read earlier. (comprehension questions follow)
Debrief	Thank you for your responses and contribution to our research. Despite what we said at the beginning of the study, you will not need to record yourself giving a lecture. Prior research has shown that students can actually learn better if they imagine themselves preparing to teach information compared to if they are just preparing to take a test (see Fiorella & Mayer, 2013; Nestojko, Bui, Kornell, & Bjork, 2014). In this present study, we are building on this prior research to examine whether people are thinking about the content differently when they have these different goals in mind.

RESULTS

Preparing to Take a Test Evokes Higher Cue Valences Compared to Preparing to Teach

Table 33 shows how cue valences differed between conditions. Four cues were experienced at higher levels by those learning with the goal of preparing to take a test. Participants preparing to take a test reported experiencing higher levels of self-efficacy and higher levels of ability to recall details from the text passage, and summarize the content, on average. In contrast, participants in the prepare to teach condition reported a higher experience of difficulty. These cue valence differences may reflect the familiarity that college students have with preparing for a test, a very common task, versus the more unusual, and stress-inducing task of teaching the content to strangers via video.

Table 33.

Descriptive Statistics of Cue Valence by Condition

Cue	Prepare to Teach *M (SD)*	Prepare to Test *M (SD)*	Cohen's *d*	*p*
Self-efficacy	6.84 (2.03)	7.39 (1.88)	.283	.029
Prior Knowledge	4.89 (2.36)	5.14 (2.64)	.099	.424
Interest	5.25 (2.57)	5.70 (2.66)	.172	.173
Difficulty	3.75 (2.26)	2.72 (2.00)	.487	<.001
Effort	5.88 (2.13)	5.84 (2.18)	.020	.872
Recall	5.26 (2.02)	6.22 (1.85)	.497	<.001
Mental Imagery	6.17 (2.49)	6.66 (2.39)	.205	.110
Summarization	5.90 (1.95)	7.02 (2.00)	.567	<.001

Preparing to Test Resulted in Higher Test Scores and Higher Metacomprehension Judgments Compared to Preparing to Teach

In contrast to prior studies, I did not replicate a benefit of preparing to teach. In fact, those in the preparing to take a test condition (*M* = 80.85%, *SD* = 21.59%) scored significantly higher than those in the preparing to teach condition (*M* = 74.35%, *SD* = 23.53%), $t(261) = 2.25$, $p = .025$, Cohen's $d = 0.291$.

This difference in test scores was also reflected in metacomprehension judgments. Participants in the prepare to teach condition gave significantly lower judgments (*M* = 60.46, *SD* = 18.46) than those in the prepare to test condition (*M* = 65.53, *SD* = 17.89), $t(261) = -2.197$, $p = .029$, Cohen's $d = 0.280$.

Condition Does Not Moderate the Relationship Between Cue Valence and Metacomprehension Judgment

To examine whether condition moderated the relationship between cue valences and judgments of comprehension, a linear regression was conducted for each cue, predicting judgment of comprehension (standardized) from cue valence (standardized), condition (0 = prepare to test, 1 = prepare to teach), and the interaction between the two. The results are summarized in Table 34.

There was only a single interaction—with the difficulty cue. This interaction is depicted in Figure 7. In both conditions, there is a negative relationship between difficulty rating and judgment of comprehension, but this association is stronger in the prepare to teach condition. At lower levels of difficulty, participants were in fact more confident in the prepare to teach condition, but this pattern is reversed at higher levels of difficulty. This intriguing finding suggests that when participants are preparing to teach, they may be more attuned to the challenges presented by the material, leading them to make more nuanced judgments of comprehension based on the perceived difficulty of the content.

Table 34.

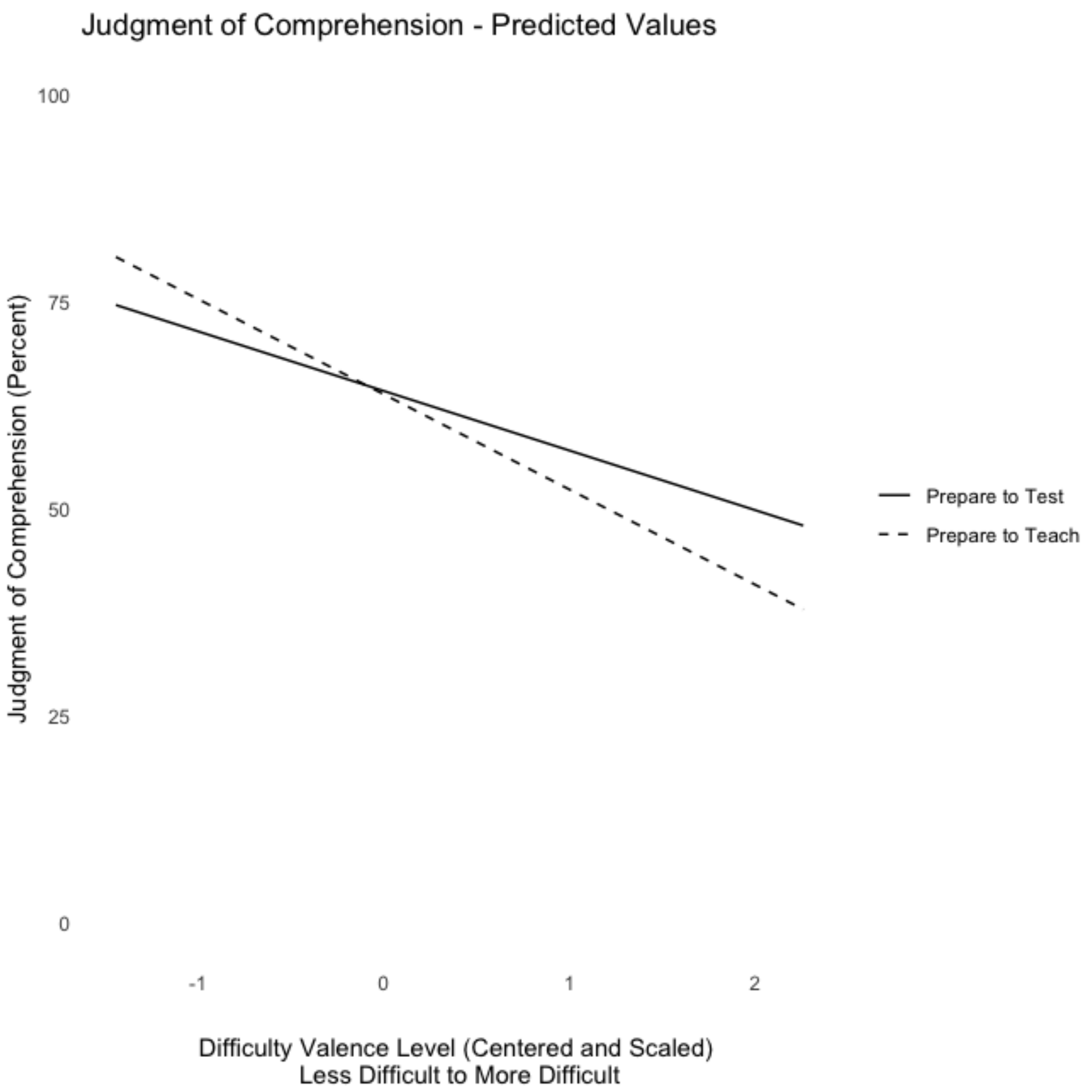
Summary of Regression Analyses Predicting JOC: Standardized Beta Coefficients and Significance Level – Metric (Standard Error)

Cue	Valence	Condition	Valence x Condition	Model Adj R2
Self-efficacy	8.44 (1.26)***	-2.21 (1.97)	2.84 (1.94)	.290 (<.001)
Prior Knowledge	4.65 (1.30)***	-4.40 (2.17)*	3.43 (2.24)	.117 (<.001)
Interest	7.73 (1.29)***	-3.76 (2.09)	-0.11 (2.11)	.186 (<.001)
Difficulty	-7.20 (1.34)***	-0.41 (2.05)	-4.30 (2.01)*	.259 (<.001)
Effort	4.50 (1.35)***	-5.19 (2.19*	2.53 (2.21)	.102 (<.001)
Recall	11.99 (1.23)***	0.04 (1.91)	-2.35 (1.88)	.357 (<.001)
Mental Imagery	3.21 (1.43)*	-4.21 (2.26)	1.67 (2.24)	.055 (.001)
Summarization	11.60 (1.11)***	1.92 (1.81)	1.89 (1.81)	.437 (<.001)

Note. Standardized beta coefficients are reported in the table, with standard error reported in parentheses.
**p* < .05. ***p* < .01. ****p* < .001.

Figure 7.

Difficulty Valence Interaction with Judgments of Comprehension



Given that there were so few interactions (and the one interaction was not particularly strong), I conducted a second multiple linear regression analysis entering all the cue valences and condition together as predictors of judgments of comprehension. These results are summarized in Table 35. After controlling for cue valences, condition still does not predict judgments of comprehension. The cue valences that uniquely predict judgments of comprehension are self-efficacy, interest, difficulty, recall, and summarization. In other words, participants rely on multiple types of cues.

Table 35.

Multiple Regression Results Predicting Judgments of Comprehension from Cue Valence and Condition

Predictor	β	SE	β 95% CI [LL, UL]	p
(Intercept)	65.00	1.26	[62.52, 67.48]	<.001
Condition (1 = Teach, 0 = Test)	-2.35	1.64	[-5.58, 0.88]	0.154
Valence Self-efficacy	3.26	0.93	[1.43, 5.09]	0.001
Valence Prior Knowledge	0.93	0.88	[-0.81, 2.67]	0.295
Valence Interest	2.43	0.92	[0.61, 4.24]	0.009
Valence Difficulty	-2.04	0.96	[-3.93, -0.16]	0.034
Valence Effort	0.51	0.84	[-1.14, 2.16]	0.544
Valence Recall	3.83	1.05	[1.76, 5.91]	<.001
Valence Mental Imagery	0.48	0.82	[-1.12, 2.09]	0.554
Valence Summarization	5.63	1.18	[3.32, 7.95]	<.001

Note: valence measures scaled and centered.
*$p < .05$. **$p < .01$. ***$p < .001$.

Condition Does Not Moderate the Relationship Between Cue Valence and Test Score

I took a similar analytic approach to examining test scores as I did for judgments of comprehension. First, a linear regression was conducted for each cue, predicting test score (standardized) from cue valence (standardized), condition (0 = prepare to test, 1 =

prepare to teach), and the interaction between the two. The results are summarized in Table 36. There was not a single significant interaction.

Table 36.

Summary of Regression Analyses Predicting Exam Score: Standardized Beta Coefficients and Significance Level

Cue	Valence	Condition	Valence x Condition	Model Adj R2
Self-efficacy	4.98 **(.005)	-4.69 (.091)	2.41 (.378)	.082 (<.001)
Prior Knowledge	1.65 (.334)	-6.34 *(.026)	-0.07 (.981)	.014 (.087)
Interest	2.00 (.250)	-5.86 *(.039)	2.88 (.313)	.031 (.011)
Difficulty	-6.50 ***(<.001)	-3.11 (.265)	-1.04 (.704)	.100 (<.001)
Effort	1.71 (.325)	-6.59 *(.019)	3.98 (.159)	.036 (.006)
Recall	9.63 ***(<.001)	-2.72 (.319)	-2.98 (.267)	.144 (<.001)
Mental Imagery	-1.11 (.539)	-6.40 *(.026)	2.61 (.357)	.012 (.109)
Summarization	7.83 *** (<.001)	-2.07 (.459)	0.43 (.877)	.127 (<.001)

Note. Standardized beta coefficients are reported in the table, with standard error reported in parentheses.
**p* < .05. ***p* < .01. ****p* < .001.

As there were no interactions with valence, I next conducted a multiple linear regression entering all cue valences and condition together to predict test score. These results are summarized in Table 37. After controlling for cue valences, condition still does not predict judgments of comprehension. The cue valences that uniquely predict judgments of comprehension are recall, with a positive effect, and difficulty, with a negative effect on judgments of comprehension.

Table 37.

Multiple Regression Results Predicting Comprehension Exam Score from Cue Valence and Condition

Predictor	*β*	SE	*β* 95% CI [LL, UL]	*p*
(Intercept)	77.63***	2.11	[73.48, 81.78]	<.001
Condition (1 = Teach, 0 = Test)	1.13	2.75	[-4.28, 6.55]	0.680
Valence Self-efficacy	1.38	1.55	[-1.68, 4.44]	0.376
Valence Prior Knowledge	-1.01	1.48	[-3.92, 1.91]	0.497
Valence Interest	0.25	1.55	[-2.80, 3.30]	0.873
Valence Difficulty	-3.21*	1.60	[-6.36, -0.05]	0.047
Valence Effort	0.52	1.41	[-2.25, 3.29]	0.713
Valence Recall	4.74**	1.77	[1.26, 8.21]	0.008
Valence Mental Imagery	-1.82	1.37	[-4.52, 0.87]	0.184
Valence Summarization	2.93	1.97	[-0.95, 6.81]	0.138

Note: valence measures scaled and centered.
*$p < .05$. **$p < .01$. ***$p < .001$.

DISCUSSION

Changing the context of the learning environment by establishing a goal for learning led to significantly different cue valences and performance on a subsequent comprehension exam. However, these group differences were not in the anticipated direction. Setting the context as learning with the goal of preparing to teach by recording a lesson was associated with lower judgments of comprehension and lower performance on a comprehension exam. Nearly across the board, instructing students to learn with the goal of preparing for a comprehension exam led to higher levels of cue valence, higher judgments of comprehension, and higher scores on the exam.

Furthermore, while lower judgments of comprehension can be advantageous, such as when a student recognizes difficulty in learning the material and gives a more accurate estimate), that is not what happened in the prepare to teach condition. While the judgments were lower, so was performance. This gap was nearly identical across both conditions. In short, I did not replicate the benefit of preparing to teach.

So, what happened? What about the context led to these divergent results? Considering the instructions from the position of the participant, a few things come to mind. First, one of the two conditions is far more familiar to college students. Reading a text passage with the goal of learning to perform well on a multiple-choice exam is perhaps the most common academic activity in college. With years of practice under their belt, the participants in this study were intimately familiar with the cues which are useful to help them monitor and control their learning under this context, having succeeded and failed in this context thousands of times before. However, reading a passage with the goal of teaching the content to a stranger is not a typical task in college. Participants in this condition were likely unsure of which metacognitive cues were valuable and how to rate their impact to their learning.

Second, preparing to record a lesson for others, while likely engaging deeper cues, also requires engaging strategies to complete the task that are less effective for the exam. With limited time to complete the task, encouraging participants to focus attention on different aspects of the material that are less directly applicable to the exam may put those in this condition at a disadvantage.

Third, the nature of the preparing to teach condition may not have been as neutral or benign as planned. Again, from the shoes of a participant, being asked to record yourself teaching a lesson, to a stranger nonetheless, can be a stress-inducing event, even if only simulated (Lacy et al., 1995). Instead of encouraging deeper processing with the material, the manipulation likely created a setting similar to test-anxiety for many participants. In fact, the results show similar patterns to those suffering from test-anxiety: lower judgments of comprehension and lower performance (Miesner & Maki, 2007).

Overall, results indicate that changing the context for learning, through something as simple as the goal for learning, can have a significant effect on how metacognitive cues are interpreted, judgments of comprehension, and performance on comprehension questions, though educators will take note to consider how participants may interpret the context. With this limitation in mind, in the next study participants were asked to complete the same tasks, but those in the prepare to teach condition were instructed to undertake a less stress-inducing activity.

Study 5

In Study 5, I altered the description of the prepare to teach condition. Instead of facing the threat of being recorded, which may make participants uneasy, the instructions were modified to ask participants to write a lesson plan for someone else to use to teach the content. As in Study 4, the idea of the manipulation is that those preparing to teach may invest more mental resources and make stronger connections with the material than those merely preparing for an exam. Thinking about how to teach the content to a stranger should require deeper engagement with the material (considering the potential levels of other students' prior knowledge, what may or may not make them interested, analogies that may be useful, etc.) (Fiorella & Mayer, 2015, Nestojko et al., 2014). This deeper engagement should be associated with higher levels of cue valence, higher judgments of comprehension, and, ultimately, higher performance on the comprehension exam. However, if participants feel threatened or uneasy, such as when facing a surprise recording of themselves, this context may interfere with the desired effect of the condition (Miesner & Maki, 2007). Thus, in Study 5, I removed the threat of recording and switched to simply preparation of a written lesson. The tasks, materials, and other procedures remained entirely the same; only the instructions were altered between Study 4 and Study 5.

METHOD

Participants

Two-hundred-eighty-six undergraduate students completed the study in exchange for course credit. After removing incomplete responses, the final sample contained 284 observations. Participants were recruited from a wide subject pool across a variety of majors (42.61% STEM, 18.66% Social Sciences). A majority female (60.92% sample, 54.4% student body), and white (40.49% sample, 38.9% student body), the sample roughly resembled the student body demographics of the large public university at which it took place (Latinx: 17.96% sample, 23.4% student body; Asian: 20.77% sample, 20.2% student body; Black: 4.93% sample, 5.3% student body). The age of participants reflected the university environment, as well (20.03 *M*, 1.47 *SD*).

Materials

Participants read the same short essay and completed the same six comprehension questions as used in Study 4.

Experimental Conditions

Participants were randomly assigned to one of two conditions: learning the material by preparing to teach or learning the material by preparing for a test. Like Study 4, this phrasing of the goal for learning, setting the context for the learning task, was the only difference between conditions. However, in contrast to Study 4, the prepare to teach condition contained slightly different instructions. To provide a preparing to teach context with a more neutral setting, Study 5 differs by instructing those in the preparing to teach condition to learn the material with the intention of designing a written summary which would be used to teach another undergraduate student the material. Identical to Study 4, those in the prepare for a test condition were instructed to learn the material with the intention of later taking a comprehension test over the material.

Prepare to teach (write lesson) condition: *"In this experiment, you will read a short expository essay and then be asked to teach the material that you learned. You will have 6 minutes to read the passage. Then, we will ask you to design and write a lecture, as if you were teaching an undergraduate student. This lecture should include as much of the original detail of the passage as possible, and make it so a reader can comprehend the* *author's position. We will use this lecture you wrote to teach the passage to another participant in the study, who will then be tested on the passage. As you are reading the passage, please DO NOT take any notes."*

Prepare to test condition: *"In this experiment, you will read a short expository essay and then take a test on what you remember and comprehend. You will have 6 minutes to read the passage. You will then be given a memory and comprehension test on the passage, as if you were a student being tested as part of a class. You will be tested on how much of the original detail of the passage you can remember, and whether you comprehend the author's position. Please DO NOT take any notes."*

Procedure

The procedure matched that of Study 4 in all respects. Only the instructions for the prepare to teach condition differed.

RESULTS

Again, Preparing to Take a Test Evokes Higher Cue Valences Compared to Preparing to Teach

Table 38 shows how cue valences differed between conditions. Three cues were experienced at higher levels by those learning with the goal of preparing to take a test. Participants preparing to take a test reported experiencing higher levels of self-efficacy, ability to recall details from the text passage, and ability to summarize the content, on average. In contrast, to Study 4, participants in the prepare to teach condition no longer reported a higher experience of difficulty. Changing from preparing to record a video to writing a lesson plan reduced perceptions of difficulty, though all other cue valence patterns remained similar. Again, these cue valence differences may reflect the familiarity that college students have with preparing for a test, a very common task, versus the more unusual task of teaching the content to strangers via lesson plans.

Table 38.

Descriptive Statistics of Cue Valence by Condition

Cue	Prepare to Write *M (SD)*	Prepare to Test *M (SD)*	Cohen's *d*	*p*
Self-efficacy	7.01 (1.89)	7.44 (1.66)	.240	.045
Prior Knowledge	5.13 (2.44)	4.99 (2.49)	.056	.637
Interest	5.77 (2.23)	5.71 (2.43)	.027	.823
Difficulty	2.93 (2.06)	2.57 (2.02)	.179	.132
Effort	6.10 (2.05)	6.06 (2.13)	.020	.866
Recall	5.78 (2.04)	6.31 (2.06)	.259	.030
Mental Imagery	6.59 (2.41)	6.48 (2.51)	.044	.711
Summarization	6.57 (1.98)	7.16 (2.09)	.289	.015

* $p < .05$, ** $p < .01$, *** $p < .001$.

Preparing to Test Did Not Result in Higher Test Scores or Higher Metacomprehension Judgments Compared to Preparing to Teach

Again, in contrast to prior studies, I did not replicate a benefit of preparing to teach. Those in the preparing to take a test condition (M = 81.98%, SD = 23.04%) performed nearly identically to those in the preparing to teach condition (M = 82.60%, SD = 21.26%), t(282) = 0.236, p = .814, Cohen's d = .028.

Despite nearly identical performance, those in the prepare to test condition were slightly more optimistic. Participants in the prepare to teach condition gave slightly lower judgments (M = 65.56, SD = 19.79) than those in the prepare to test condition (M = 67.57, SD = 17.69), but the difference was not significant, $t(282)$ = .900, p = .370, Cohen's d = 0.107.

Again, Condition Does Not Moderate the Relationship Between Cue Valence and Metacomprehension Judgment

To examine whether condition moderated the relationship between cue valences and judgments of comprehension, a linear regression was conducted for each cue, predicting judgment of comprehension (standardized) from cue valence (standardized), condition (0 = prepare to test, 1 = prepare to teach), and the interaction between the two. The results are summarized in Table 39.

While in Study 4 there was an interaction with difficulty, when switching from preparing to record a video to preparing to write a lesson plan, this interaction disappeared. Across the other seven valences there were also no significant interactions. Condition did not moderate the relationship between cue valence and metacomprehension judgments.

Table 39.

Summary of Regression Analyses Predicting JOC: Standardized Beta Coefficients and Significance Level – Metric (Standard Error)

Cue	Valence	Condition	Valence x Condition	Model Adj R2
Self-efficacy	7.84 (1.42)***	0.28 (1.92)	3.34 (1.93)	.265 (<.001)
Prior Knowledge	1.44 (1.52)	-2.15 (2.21)	2.08 (2.22)	.012 (.093)
Interest	3.64 (1.43)*	-2.14 (2.14)	3.08 (2.16)	.070 (<.001)
Difficulty	-7.75 (1.39)***	-0.43 (1.98)	-2.02 (1.98)	.214 (<.001)
Effort	0.71 (1.51)	-2.04 (2.22)	2.22 (2.23)	.004 (.239)
Recall	10.28 (1.22)***	1.00 (1.78)	2.63 (1.78)	.373 (<.001)
Mental Imagery	1.82 (1.48)	-2.18 (2.18)	4.04 (2.19)	.043 (.002)
Summarization	9.59 (1.25)***	1.13 (1.85)	2.59 (1.86)	.324 (<.001)

Note. Standardized beta coefficients are reported in the table, with standard error reported in parentheses.
*$p < .05$. **$p < .01$. ***$p < .001$.

Given that there were no interactions, I conducted a second multiple linear regression analysis entering all the cue valences and condition together as predictors of judgments of comprehension. These results are summarized in Table 40. After controlling for cue valences, condition still does not predict judgments of comprehension. The cue valences that uniquely predict judgments of comprehension are self-efficacy, difficulty, recall, and summarization.

Table 40.

Multiple Regression Results Predicting Judgments of Comprehension from Cue Valence and Condition

Predictor	β	SE	β 95% CI [LL, UL]	p
(Intercept)	67.74***	1.19	[65.39, 70.09]	<.001
Condition (1 = Teach, 0 = Test)	-2.17	1.67	[-5.46, 1.12]	.195
Valence Self-efficacy	3.28**	1.06	[1.19, 5.37]	.002
Valence Prior Knowledge	-0.85	0.92	[-2.66, 0.97]	.361
Valence Interest	-0.89	1.04	[-2.93, 1.16]	.394
Valence Difficulty	-2.95**	1.03	[-4.97, -0.92]	.004
Valence Effort	-0.49	0.92	[-2.30, 1.32]	.594
Valence Recall	6.42***	1.17	[4.11, 8.72]	<.001
Valence Mental Imagery	0.25	0.93	[-1.58, 2.08]	.784
Valence Summarization	3.99**	1.16	[1.70, 6.27]	.001

Note: valence measures scaled and centered.
*$p < .05$. **$p < .01$. ***$p < .001$.

Again, Condition Does Not Moderate the Relationship Between Cue Valence and Exam Score

I took a similar analytic approach to examining test scores as I did for judgments of comprehension. First, a linear regression was conducted for each cue, predicting test score (standardized) from cue valence (standardized), condition (0 = prepare to test, 1 =

prepare to teach), and the interaction between the two. The results are summarized in Table 41. There was not a single significant interaction.

Previously, in Study 4, while no interactions were significant, for several cues condition was a significant predictor of exam performance. However, with the new prepare to teach instructions (write lesson plan, instead of record video), all effects of condition on performance disappeared.

Table 41.

Summary of Independent Regression Analyses Predicting Exam Score from Each Cue

Cue	Valence	Condition	Valence x Condition	Model Adj R2
Self-efficacy	5.10 (1.93)**	1.79 (2.60)	-0.34 (2.61)	.039 (.003)
Prior Knowledge	1.43 (1.82)	0.57 (2.65)	-1.04 (2.65)	.001 (.869)
Interest	3.78 (1.75)*	0.56 (2.63)	-3.10 (2.65)	.006 (.187)
Difficulty	-3.82 (1.83)*	1.32 (2.62)	-0.24 (2.62)	.021 (.030)
Effort	3.17 (1.79)	0.60 (2.63)	-4.96 (2.64)	.004 (.259)
Recall	6.71 (1.78)	2.06 (2.58)	-2.12 (2.58)	.058 (<.001)
Mental Imagery	2.64 (1.79)	0.60 (2.63)	-4.54 (2.65)	.001 (.366)
Summarization	8.91 (1.70)***	2.76 (2.51)	-2.73 (2.52)	.112 (<.001)

Note. Standardized beta coefficients are reported in the table, with standard error reported in parentheses.
*$p < .05$. **$p < .01$. ***$p < .001$.

As there were no interactions with valence, I next conducted a multiple linear regression entering all cue valences and condition together to predict test score. These results are summarized in Table 42. While condition remained non-significant, the findings diverged from Study 4. In Study 4, both recall and difficulty emerged as significant predictors of exam outcomes. Yet, in this analysis, only summarization significantly influenced performance, showing a positive effect.

Table 42.

Multiple Regression Results Predicting Comprehension Exam Score from All Cue Valences and Conditions Simultaneously

Predictor	beta	SE	beta 95% CI [LL, UL]	p
(Intercept)	84.00***	1.82	[80.41, 87.59]	<.001
Condition (1 = Teach, 0 = Test)	-3.31	2.55	[-8.32, 1.71]	0.195
Valence Self-efficacy	1.11	1.62	[-2.08, 4.30]	0.495
Valence Prior Knowledge	-0.39	1.41	[-3.17, 2.38]	0.781
Valence Interest	-0.33	1.58	[-3.44, 2.78]	0.834
Valence Difficulty	-0.20	1.57	[-3.29, 2.88]	0.897
Valence Effort	0.15	1.40	[-2.62, 2.91]	0.917
Valence Recall	1.34	1.79	[-2.18, 4.86[	0.455
Valence Mental Imagery	-1.75	1.42	[-4.54, 1.04]	0.219
Valence Summarization	6.78***	1.77	[3.29, 10.26]	<.001

Note: valence measures scaled and centered.
**p* < .05. ***p* < .01. ****p* < .001.

Discussion

Altering the prepare to teach instructions to a more benign task removed differences in exam performance between the conditions. However, there were still differences in how participants in each condition perceived their metacomprehensive experiences. Valence levels for self-efficacy, recall, and summarization were all significantly higher in the prepare to test condition, perhaps continuing to reflect comfort and familiarity with the task of prepping for an exam versus developing a lesson plan.

However, despite these differences in valence reporting between conditions, there was no significant effect of condition on either judgments of comprehension or exam performance.

Overall, results indicate that changing the context of the learning environment can affect the cues learners use to evaluate their learning. An unfamiliar task that may be anxiety-inducing (preparing to teach by recording a lesson) can increase perceptions of difficulty and decrease reported levels of important cues, such as summarization and recall. Removing the stress of recording oneself but retaining an unfamiliar task (writing a lesson plan), can lower perceptions of difficulty, but is still interpreted metacognitively in a different way than a more familiar task, such as preparing for a multiple-choice comprehension exam.

Despite these differences in interpretation, when the ultimate task is still something familiar (like a multiple-choice comprehension exam), changing the learning context may not change perceived judgments of comprehension or even exam performance.

CHAPTER 5: GENERAL DISCUSSION

Metacomprehension serves not merely as a cognitive afterthought but as a cornerstone in the complex architecture of self-regulated learning. Its role in modulating study strategies and resource allocation positions it as a crucial determinant of both academic success and lifelong learning. Previous research has often focused on isolated cues and relied on single, open-ended responses to probe learners' judgments. However, this dissertation represents a significant shift in how we approach the exploration of metacomprehension cues and their nuanced interactions with individual differences and contextual factors.

The primary objective here was to develop and implement a multidimensional methodological framework that allows for a more comprehensive understanding of how learners form judgments of comprehension. This innovative paradigm introduces a nuanced way to measure cue valence and its directional impact on metacomprehension, going beyond mere cue identification. This approach counters the limitations inherent in traditional open-response methodologies, which often underestimate the diversity and intensity of cues that learners use in making metacomprehension judgments.

Summary of Main Findings

This dissertation journeyed through a series of interconnected studies, each contributing to the overarching narrative on metacomprehension. One of the most significant findings was the unveiling of multiple cues—self-efficacy, interest, difficulty, recall, and summarize—as predictors of judgments of comprehension. This contradicts the traditional notion that learners rely on singular or isolated cues, emphasizing instead that metacomprehension is influenced by a myriad of factors, each with its unique valence.

Moreover, this research unveiled the nuance between cues related to comprehension judgments and those related to learning. For instance, while several cues were associated with metacomprehension judgments, those that reflected the use of deeper cognitive strategies like summarization and recall were more uniquely related to actual comprehension. This distinction is critical because it suggests that fostering metacognitive awareness around such deeper cues could improve comprehension, an insight with far-reaching educational implications.

A further notable outcome was the potency of deeper cognitive strategies like summarization and recall in influencing metacomprehension. These deeper cues were uniquely related to both judgments of comprehension and actual comprehension, emphasizing their critical role in effective learning. While individual differences like analytical skills and interest in the topic did have an impact, they were not the dominant factors. Additionally, the learning context—such as preparing to test versus preparing to teach—showed no significant moderating effect on how cue valence influenced judgments of comprehension.

My findings challenge the assumption of accurate meta-metacomprehension, demonstrating that learners struggle to explicitly report their experiences with metacognitive cues unless guided. By measuring multiple cues and their valences and synthesizing this information in the context of individual differences and learning settings, this dissertation offers not just an incremental addition to the literature but a paradigm shift. It provides a more granular lens through which we can understand how multiple cues interact to shape judgments of comprehension and learning outcomes, how these cues can vary in importance across different individuals, and how their efficacy can be context dependent.

Implications

The overarching achievement of this dissertation is its significant shift in metacomprehension research from a single cue focus to a nuanced, multi-dimensional understanding of multiple cues and their respective valences. This shift has several critical implications for both research and practice.

One of the most groundbreaking contributions is the inclusion of valence as a vital measure. Previously, studies in metacomprehension research centered on the presence or absence of cues. However, this dissertation illuminates the importance of understanding the directional impact of these cues. This nuance allows researchers to probe more deeply into questions that were hitherto elusive, such as the differential impacts of negative and positive valences of the same cue on metacomprehension. Incorporating valence measurements can significantly enrich our understanding and could be crucial for designing educational interventions that alter cue valences for more effective learning outcomes.

In addition to valence, this dissertation also emphasizes the need for a more comprehensive approach that considers multiple cues concurrently. The research reveals that metacomprehension can no longer be understood adequately by looking at individual, isolated cues. Instead, it demands an appreciation of the interplay between these cues. This insight is particularly crucial when considering educational interventions. For example, understanding how learners use multiple cues can lead to the development of personalized educational plans that target the most influential cues for each individual learner. This multi-cue approach also allows for the creation of enhanced study strategies that consider how a combination of cues can either reinforce or negate each other.

Furthermore, these findings have immediate applicability to diverse learning environments. Whether in academic settings, workplace training, or lifelong learning initiatives, understanding the complex interplay of multiple cues and their valences can equip educators and instructional designers with the tools needed to create more effective learning scenarios that facilitate self-regulated learning.

Limitations

While this dissertation offers significant contributions to our understanding of metacomprehension, some limitations must be acknowledged to provide a complete picture and offer avenues for future research.

One evident limitation is the dissertation's reliance on self-report measures to gather insights on participants' thoughts and cue usage. Though invaluable, self-report measures have their drawbacks. For instance, participants might not be fully aware of all the cues affecting their judgments or could find it challenging to articulate them precisely. While the novel methodology used here aims to mitigate some of these concerns, the limitations inherent to self-reporting still apply.

The materials employed for the studies were complex but not exhaustive in scope or length. In real-world educational settings, students are likely to interact with much more extensive and multifaceted learning materials. As such, generalizing the findings to prolonged and complicated educational scenarios demands caution. Future studies employing longer, more complex materials and following participants over extended periods will be essential for validating the cues identified as shaping metacomprehension judgments.

Relatedly, there are limitations in the research design methods incorporated here, specifically the predominant use of between-subjects paradigms. Allowing participants to experience multiple conditions, and report through various measures any changes in how they approached the task or metacognitively engaged with it, could be a beneficial avenue of research.

In addition, the set of cues examined here, although based on comprehensive literature reviews, may not capture all possible influencers of metacomprehension. It remains plausible that there exist additional significant cues that research has yet to identify. Thus, subsequent studies should expand the list of cognitive and emotional cues investigated to generate a more comprehensive picture of metacomprehension's contributing factors.

The current sample may not be fully representative, thus limiting the generalizability of the findings. Future work should focus on including a more diverse set of participants that span different educational levels, cultures, and age groups to provide a richer understanding of the dynamics of metacomprehension.

By acknowledging and addressing these limitations in future research, the field can build on the strong foundation established by this dissertation.

Contribution and Future Research

I hope that this dissertation serves as a launching pad for a new generation of metacomprehension research. The innovative methods presented here should enable researchers to design more effective educational interventions that consider a learner's holistic experience rather than reducing it to isolated cues or moments. This has implications beyond the academic setting; a better understanding of metacomprehension has the potential to improve various forms of self-directed learning, including traditional educational settings, workplace training, and lifelong educational pursuits.

In closing, this dissertation serves as a key waypoint in the journey to understand metacomprehension. It offers both a look back at the groundwork laid by previous research and a vision for the future. By paving the way for more nuanced and comprehensive studies, this dissertation stands as an enduring contribution to the field of educational psychology.

APPENDIX

STUDY 1 – MATERIALS READ BY PARTICIPANTS

Earth's Rotation: The spinning Earth rotates in the solar system at a constant tilt. This tilted spinning, which is called "precession," is due to a balance between the spin of the Earth which acts to straighten the Earth up, and gravity which acts to pull the Earth down.

Earth's Orbit: The Earth orbits the sun in accordance with a basic principle. The orbit involves striking a balance between two forces that equal each other, but act in ways opposite to each other. With the first of these forces, which is called "centripetal force," the sun acts to pull the Earth in towards itself. However, with a second, equally powerful force, the Earth acts to pull out away from the sun. This second force is greater the more massive the planet.

Planet's Orbit: The planets orbit the sun on nearly-circular pathways and they orbit at different speeds. Those planets that are on paths closer to the sun move at a faster speed because the force of gravity is strong; those planets farther from the sun move at a slower speed because the force of gravity is weak.

Collapsing Stars: Collapsing stars spin faster and faster as they fold in on themselves and their size decreases. This phenomenon of spinning faster as the star's size shrinks occurs because of a principle called "conservation of angular momentum."

Pulsars: A pulsar is a collapsing star. As the star collapses, it rotates steadily and it also emits streams of radiation. The stream of radiation from the pulsar flows continuously. However, by virtue of the fact that collapsing stars rotate, the pulsar's stream of radiation appears to us here on Earth as though it was flashing intermittently.

Galaxies: The galaxies in the universe constantly move farther and farther apart from one another. The reason galaxies move apart is that the universe itself constantly expands farther and farther outward. It is also important to note that the surface of the universe has no one particular place that can rightly be called its "center."

Black and White Holes: A black hole continuously draws in all the different types of material that surround it in space. The black hole will let no material escape from its powerful pull, and that even includes starlight. The black hole can thus be contrasted with the white hole. With the white hole, all material is spewed back into the universe.

Telescopes: The astronomer who tries to catch starlight in a telescope must position the telescope forward. In other words, she must tilt the telescope in such a way that the starlight will fall all the way through it. The astronomer must do this in order to compensate for the movement of the Earth in its orbit. Otherwise, the starlight will land on the side of the astronomer's telescope.

Light's Travels: As light travels into a piece of glass it instantly changes its angle of travel. It also begins to move at a slower speed as it passes into the glass, because glass is a denser substance. As the light exits the glass and enters less dense air, it instantly changes direction and speeds up again.

STUDY 2 – MATERIALS READ BY PARTICIPANTS

One Kingdom: Our Lives with Animals by Deborah Noyes

Ideally a zoo visit is a mentorship—emotional, intellectual, spiritual—a way to interact with nature on a concentrated scale. It's the wide, wild world in miniature, and its human architects are conscientious stewards. The zoo experience should be, first and foremost, meaningful. There's too much at stake for it not to be. But zoos are a paradox. Even as children, many of us feel there—together with our interest and curiosity—a muffled unease. People do have meaningful encounters in zoos, or they wouldn't flock to them in record numbers. I've had my share, usually late in the day when the crowds have thinned or gone, when the heat of the sun has waned, when the evening's meal is imminent and the animals know it, when I'm willing to sit alone—blank and patient and outside myself—and sit some more. But more often my experience has been representative. If the animals are visible at all—not off exhibit or obscured by the very greenery installed to protect their privacy (and who would begrudge them?)—we watch through wire or glass, aching for a connection that rarely comes. Some children (adults, too) rap on the window or otherwise urge the animals on with funny faces and undignified attempts at cross-species communication. Do something, we think, and they do precisely what they will or won't. Natural antics—monkeys grooming or swinging in play—delight us, but familiar zoo behaviors like pacing, swaying, or regurgitating food evoke a vague embarrassment, as of some unwelcome intimacy. We may half-heartedly read the sign stationed to inform us of the captive's natural habitat and behaviors, but by now its unnatural fate may well have disheartened us. We seek solace in interaction, buttons to push and levers to pull, or we fix on some other distraction: tired toddlers wailing for ice cream, the heat, a blaring boombox. In the end, novelty wins out over a tangle of emotions we can't name; that, or the tug to move on to the next exhibit, to the gift shop, to the snack bar, before our legs give out. In this way we carry on... consumers at odds with our own motives. Perhaps we're uneasy because the animals withhold from us the one thing we would have: their consent. It would ease my spirit (prepare for some shameless anthropomorphism here), to be sure, if the wolf suddenly ceased its pacing, looked up, met my eye, and said, "Welcome, friend, and thanks for being here today. You see me, and it has changed you. I now see the worth of my sacrifice." But he will not pause. He does not look. I am unforgiven. These are my own feelings, not representative of anyone else's, I realize. But zoos do seem to leave many people uncomfortable, maybe because—despite our best intentions—it's a lopsided exchange. We consume a healthy sampling of the world's biodiversity in an afternoon, and we make it home for dinner. But what do humans give zoo animals in return? The benefits of captive breeding. Basic care and protection (survival is no picnic in the wild; animals certainly have it easier in captivity). But is survival enough, and why is it ours to give? It's an old controversy, and a circular one. There are as many ways to justify captivity (the animals have nowhere else to go... they're safer and better nourished in zoos... they're educational ambassadors here on behalf of their kind) as there are arguments against it. I have, with difficulty, weighted the scale first one way and then the other, and in the end I'm no closer to knowing than I ever was. I fear someday we'll regret what could not be helped, but for now I continue to see the intrinsic worth of zoos.

Living Alone Is the New Norm by Eric Klinenberg

The extraordinary rise of solitary living is the biggest social change that we've neglected to identify, let alone examine. Consider that in 1950, a mere 4 million Americans lived alone, and they made up only 9% of households. Back then, going solo was most common in the open, sprawling western states—Alaska, Montana and Nevada—that attracted migrant workingmen, and it was usually a short-lived stage on the road to more conventional domestic life. Not anymore. According to 2011 census data, people who live alone—nearly 33 million Americans—make up 28% of all U.S. households, which means they are now tied with childless couples as the most prominent residential type, more common than the nuclear family, the multigenerational family and the roommate or group home. These aren't just transitional living situations: over a five-year period, people who live alone are more likely to remain in their current state than anyone else except married couples with children. They're concentrated in big cities throughout the country, from

Seattle to Miami, Minneapolis to New Orleans. Living alone, being alone and feeling lonely are hardly the same, yet in recent years experts have routinely conflated them, raising fears that the rise of soloists signals the ultimate atomization of the modern world. The theme of declining communities entered popular culture with Bowling Alone, political scientist Robert D. Putnam's book, which was published in 2000. It argued that social splintering had diminished the quality of life in the U.S. More recently, in The Lonely American, Harvard psychiatrists Jacqueline Olds and Richard S. Schwartz warn that "increased aloneness" and "the movement in our country toward greater social isolation" are damaging our health and happiness. Their evidence: First, a widely disputed finding published in the American Sociological Review that from 1985 to 2004, the number of Americans who said they had no one with whom they discussed important matters had tripled, to nearly a quarter of the population. (One of the study's authors later acknowledged that there was a problem with the data and that the findings were unreliable.) Second, an interpretation: that the record number of people who live alone is a sign of how lonely and disconnected we have become. In fact, there's little evidence that the rise of living alone is making more Americans lonely. Reams of published research show that it's the quality, not the quantity, of social interactions that best predicts loneliness. As University of Chicago social neuroscientist John T. Cacioppo concluded in the book he co-authored, Loneliness, what matters is not whether we live alone but whether we feel alone. There's ample support for this idea outside the laboratory. As divorced or separated people often say, there's nothing lonelier than living with the wrong person. My research—which included more than 300 interviews with people who live alone and careful scrutiny of the scientific literature on the social connections of solo dwellers—shows that most singletons are not lonely souls. On the contrary, the evidence suggests that people who live alone compensate by becoming more socially active than those who live with others and that cities with high numbers of singletons enjoy a thriving public culture. The truth is, nearly everyone who lives alone had other, less expensive options, from finding roommates to living with family. But today most people presented with those choices will opt to go solo. Wouldn't you? After all, living alone serves a purpose: it helps us pursue sacred modern values—individual freedom, personal control and self-realization—that carry us from adolescence to our final days. Living alone allows us to do what we want, when we want, on our own terms. It liberates us from the constraints of a domestic partner's needs and demands and permits us to focus on ourselves. Today, in our age of digital media and ever expanding social networks, living alone can offer even greater benefits: the time and space for restorative solitude. This means that living alone can help us discover who we are as well as what gives us meaning and purpose. Paradoxically, living alone might be exactly what we need to reconnect.

STUDY 4 AND STUDY 5 – MATERIALS READ BY PARTICIPANTS

One Kingdom: Our Lives with Animals by Deborah Noyes

Ideally a zoo visit is a mentorship—emotional, intellectual, spiritual—a way to interact with nature on a concentrated scale. It's the wide, wild world in miniature, and its human architects are conscientious stewards. The zoo experience should be, first and foremost, meaningful. There's too much at stake for it not to be. But zoos are a paradox. Even as children, many of us feel there—together with our interest and curiosity—a muffled unease. People do have meaningful encounters in zoos, or they wouldn't flock to them in record numbers. I've had my share, usually late in the day when the crowds have thinned or gone, when the heat of the sun has waned, when the evening's meal is imminent and the animals know it, when I'm willing to sit alone—blank and patient and outside myself—and sit some more. But more often my experience has been representative. If the animals are visible at all—not off exhibit or obscured by the very greenery installed to protect their privacy (and who would begrudge them?)—we watch through wire or glass, aching for a connection that rarely comes. Some children (adults, too) rap on the window or otherwise urge the animals on with funny faces and undignified attempts at cross-species communication. Do something, we think, and they do precisely what they will or won't. Natural antics—monkeys grooming or swinging in play—delight us, but familiar zoo behaviors like pacing, swaying, or regurgitating food evoke a vague embarrassment, as of some unwelcome intimacy. We may half-heartedly read the sign stationed to inform us of the captive's natural habitat and behaviors, but by now its unnatural fate may well have disheartened us. We seek solace in interaction, buttons to push and levers to pull, or we fix on some other distraction: tired toddlers wailing for ice cream, the heat, a blaring boombox. In the end, novelty wins out over a tangle of emotions we can't name; that, or the tug to move on to the next exhibit, to the gift shop, to the snack bar, before our legs give out. In this way we carry on... consumers at odds with our own motives. Perhaps we're uneasy because the animals withhold from us the one thing we would have: their consent. It would ease my spirit (prepare for some shameless anthropomorphism here), to be sure, if the wolf suddenly ceased its pacing, looked up, met my eye, and said, "Welcome, friend, and thanks for being here today. You see me, and it has changed you. I now see the worth of my sacrifice." But he will not pause. He does not look. I am unforgiven. These are my own feelings, not representative of anyone else's, I realize. But zoos do seem to leave many people uncomfortable, maybe because—despite our best intentions—it's a lopsided exchange. We consume a healthy sampling of the world's biodiversity in an afternoon, and we make it home for dinner. But what do humans give zoo animals in return? The benefits of captive breeding. Basic care and protection (survival is no picnic in the wild; animals certainly have it easier in captivity). But is survival enough, and why is it ours to give? It's an old controversy, and a circular one. There are as many ways to justify captivity (the animals have nowhere else to go... they're safer and better nourished in zoos... they're educational ambassadors here on behalf of their kind) as there are arguments against it. I have, with difficulty, weighted the scale first one way and then the other, and in the end I'm no closer to knowing than I ever was. I fear someday we'll regret what could not be helped, but for now I continue to see the intrinsic worth of zoos.

Study 4 and Study 5 – Comprehension Questions Answered by Participants

Correct answers marked in bold.

1. Which statement best expresses the main idea of the essay?

- As uncomfortable as zoos can make us feel, they are useful and appealing.

- Zoos benefit everyone by helping animals and educating people.

- When visiting zoos, people should respect the animals.

- Zoos have initiated many improvements that benefit animals.

2. Early in the article, the author wrote "It's the wide, wild world in miniature, and its human architects are conscientious stewards." In this quotation, the author means that -

- zoos should provide visitors with opportunities to feed and care for the animals

- zoos need to be redesigned to accommodate most of the world's animal species

- people are responsible for properly feeding the animals in zoos

- people have built and control the unnatural world the animals in their care now inhabit

3. What evidence does the author primarily use to support her message?

- Facts about biodiversity

- Expert opinions and interviews

- Personal observation and reflection

- Details about zoo practices

4. In the middle paragraphs, the author seeks to -

- explain why visitors to zoos often experience troubling emotions

- describe the zoo habitats of the animals that children most like to visit

- advocate for zoo programs that will provide more-stimulating activities for the animals

- inform visitors to zoos how to behave when viewing the animals

5. Why does the author include an imaginary conversation with a wolf?

- To contrast the experience of animals in a zoo with that of wild animals

- To suggest that animals are worthy of moral consideration

- To ridicule the behavior of zoo visitors who do not respect the animals

- To present a subtle argument for the improved treatment of animals

6. Which of the following provides the best summary of the essay?

- Zoos exhibit the world's animals to millions of visitors every year. But since the animals' habitats do not provide the freedom and stimulation found in the wild, the animals often act in unnatural ways that are disturbing to the people observing them.

- Zoos provide meaningful opportunities for people to observe animals they might not otherwise see. However, seeing animals in captivity can be disheartening and can make people wonder whether the benefits of zoos outweigh the disadvantages.

- Zoos allow people the opportunity to observe wild animals at close range. Although these opportunities are beneficial to people, the animals in zoos live in unnatural habitats that restrict their movement and cause them permanent harm.

- Zoos give visitors a view of the world's animals in settings similar to their natural habitat. By providing care and food to these animals, zoos improve the survival chances for many species facing extinction in the wild.

REFERENCES

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*(3), 219–235. https://doi.org/10.1177/1088868309341564

Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology*, *81*(1), 126–131. https://doi.org/10.1037/h0027455

Bandura, A. (1997). *Self-Efficacy: The exercise of control*. Macmillan.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development*, *67*(3), 1206. https://doi.org/10.2307/1131888

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, *3*(2), 77–85. https://doi.org/10.1111/j.2044-8317.1950.tb00285.x

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, *28*(5), 610–632. https://doi.org/10.1016/0749-596x(89)90016-8

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of*

Experimental Psychology: General, *127*(1), 55–68. https://doi.org/10.1037/0096-3445.127.1.55

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-Regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*(1), 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Blake, A. B., & Castel, A. D. (2018). On belief and fluency in the construction of judgments of learning: Assessing and altering the direct effects of belief. *Acta Psychologica*, *186*, 27–38. https://doi.org/10.1016/j.actpsy.2018.04.004

Boekaerts, M. (2010). The crucial role of motivation and emotion in classroom learning. In *Educational Research and Innovation* (pp. 91–111). OECD. http://dx.doi.org/10.1787/9789264086487-6-en

Bröder, A., & Undorf, M. (2019). Metamemory viewed through the judgment lens. *Acta Psychologica*, *197*, 153–165. https://doi.org/10.1016/j.actpsy.2019.04.011

Brown, A. L. (1977). *Knowing When, Where, and how to Remember: A Problem of Metacognition*. Center for the Study of Reading.

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin & Review*, *7*(1), 26–48. https://doi.org/10.3758/bf03210724

Caspi, A., Gorsky, P., & Privman, M. (2005). Viewing comprehension: Students? learning preferences and strategies when studying from video. *Instructional Science*, *33*(1), 31–47. https://doi.org/10.1007/s11251-004-2576-x

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. Springer US. http://dx.doi.org/10.1007/978-1-4684-2262-7

Cattell, R. B., & Vogelmann, S. (1977). A comprehensive trial of the scree and kg criteria for determining the number of factors. *Multivariate Behavioral Research*, *12*(3), 289–325. https://doi.org/10.1207/s15327906mbr1203_2

Channel, S. (2020). How it's made: Mascara [Video]. In *YouTube*. https://www.youtube.com/watch?v=xK3gkLMZdIo

Channel, S. (2022). How it's made: Hot sauce [Video]. In *YouTube*. https://www.youtube.com/watch?v=-Jj5MTtfJfQ

de Bruin, A. B. H., Thiede, K. W., Camp, G., & Redford, J. (2011). Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology*, *109*(3), 294–310. https://doi.org/10.1016/j.jecp.2011.02.005

Deci, E., Vallerand, R., Pelletier, L., & Ryan, R. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist*, *26*(3), 325–346. https://doi.org/10.1207/s15326985ep2603&4_6

Denton, C. A., Wolters, C. A., York, M. J., Swanson, E., Kulesz, P. A., & Francis, D. J. (2015). Adolescents' use of reading comprehension strategies: Differences related to reading proficiency, grade level, and gender. *Learning and Individual Differences*, *37*, 81–95. https://doi.org/10.1016/j.lindif.2014.11.016

Dunlosky, J., & Bjork, R. A. (2014). The integrated nature of metamemory and memory. In *Handbook of Metamemory and Memory*. Routledge. http://dx.doi.org/10.4324/9780203805503.ch2

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension. *Current Directions in Psychological Science*, *16*(4), 228–232. https://doi.org/10.1111/j.1467-8721.2007.00509.x

Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. SAGE Publications.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*(4), 271–280. https://doi.org/10.1016/j.learninstruc.2011.08.003

Eakin, D. K., & Moss, J. (2018). The methodology of metamemory and metacomprehension. In *Handbook of Research Methods in Human Memory* (pp. 125–153). Routledge. http://dx.doi.org/10.4324/9780429439957-8

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, *38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Fiorella, L., & Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology*, *39*(2), 75–85. https://doi.org/10.1016/j.cedpsych.2014.01.001

Fiorella, L., & Mayer, R. E. (2015). Eight ways to promote generative learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

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

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology*, *1*(4), 324–340. https://doi.org/10.1016/0010-0285(70)90019-8

Flavell, J. H., & Wellman, H. M. (1975). Metamemory. *American Psychological Association, Chicago*.

Frank, D. J., & Kuhlmann, B. G. (2017). More than just beliefs: Experience and beliefs jointly contribute to volume effects on metacognitive judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(5), 680–693. https://doi.org/10.1037/xlm0000332

Glenberg, A. M., & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition*, *15*(1), 84–93. https://doi.org/10.3758/bf03197714

Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General*, *116*(2), 119–136. https://doi.org/10.1037/0096-3445.116.2.119

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*(1), 75–90. https://doi.org/10.1037/0033-295x.109.1.75

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). DISCOURSE COMPREHENSION. *Annual Review of Psychology*, *48*(1), 163–189. https://doi.org/10.1146/annurev.psych.48.1.163

Granado, E. A. (2015). *Comparing three effect sizes for latent class analysis.* [Dissertation]. University of North Texas.

Griffin, T. D., Jee, B. D., & Wiley, J. (2009). The effects of domain knowledge on metacomprehension accuracy. *Memory & Cognition*, *37*(7), 1001–1013. https://doi.org/10.3758/mc.37.7.1001

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, *36*(1), 93–103. https://doi.org/10.3758/mc.36.1.93

Hagenaars, J. A., & McCutcheon, A. L. (2002). *Applied latent class analysis*. Cambridge University Press.

Hertwig, R., Herzog, S. M., Schooler, L. J., & Reimer, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(5), 1191–1206. https://doi.org/10.1037/a0013025

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(1), 22–34. https://doi.org/10.1037/0278-7393.29.1.22

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jaeger, A. J. (2012). *Can Self-Explanation Improve Metacomprehension Accuracy for Illustrated Text?* [Thesis]. University of Illinois at Chicago.

Jaeger, A. J., & Wiley, J. (2014). Do illustrations help or harm metacomprehension accuracy? *Learning and Instruction*, *34*, 58–73. https://doi.org/10.1016/j.learninstruc.2014.08.002

Jaeger, A. J., & Wiley, J. (2015). Reading an analogy can cause the illusion of comprehension. *Discourse Processes*, *52*(5–6), 376–405. https://doi.org/10.1080/0163853x.2015.1026679

Jia, X., Li, P., Li, X., Zhang, Y., Cao, W., Cao, L., & Li, W. (2016). The effect of word frequency on judgments of learning: Contributions of beliefs and processing fluency. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01995

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. https://doi.org/10.1037/0003-066x.58.9.697

Kahneman, D., & Frederick, S. (2005). A Model of Heuristic Judgment. In *The Cambridge Handbook of Thinking and Reasoning*. Cambridge University Press.

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, *35*(4), 401–415. https://doi.org/10.1007/bf02291817

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark Iv. *Educational and Psychological Measurement*, *34*(1), 111–117. https://doi.org/10.1177/001316447403400115

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295x.95.2.163

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

Kintsch, W., & Rawson, K. A. (2005). Comprehension. In M. J. Snowling (Ed.), *The Science of Reading: A Handbook* (pp. 211–226). Blackwell Publishing Ltd. http://dx.doi.org/10.1002/9780470757642.ch12

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*(4), 609–639. https://doi.org/10.1037/0033-295x.100.4.609

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, *124*(3), 311–333. https://doi.org/10.1037/0096-3445.124.3.311

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*(4), 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, *9*(2), 149–171. https://doi.org/10.1006/ccog.2000.0433

Koriat, A. (2012). The subjective confidence in one's knowledge and judgements: Some metatheoretical considerations. In *Foundations of Metacognition* (pp. 213–233). Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199646739.003.0014

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 34–53. https://doi.org/10.1037/0278-7393.27.1.34

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, *52*(4), 478–492. https://doi.org/10.1016/j.jml.2005.01.001

Kornell, N. (2014). Where is the "meta" in animal metacognition? *Journal of Comparative Psychology*, *128*(2), 143–149. https://doi.org/10.1037/a0033444

Kornell, N. (2018, January 17). *Three words that helped my 13-year-old study better*. Psychology Today. https://www.psychologytoday.com/us/blog/everybody-is-stupid-except-you/201801/three-words-helped-my-13-year-old-study-better

Kornell, N., & Bjork, R. A. (2009). A stability bias in human memory: Overestimating remembering and underestimating learning. *Journal of Experimental Psychology: General*, *138*(4), 449–468. https://doi.org/10.1037/a0017350

Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias. *Psychological Science*, *22*(6), 787–794. https://doi.org/10.1177/0956797611407929

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Lacy, C. R., Contrada, R. J., Robbins, M. L., Tannenbaum, A. K., Moreyra, A. E., Chelton, S., & Kostis, J. B. (1995). Coronary vasoconstriction induced by mental stress (simulated public speaking). *The American Journal of Cardiology*, *75*(7), 503–505. https://doi.org/10.1016/s0002-9149(99)80590-6

Li, P., Jia, X., Li, X., & Li, W. (2016). The effect of animacy on metamemory. *Memory & Cognition*, *44*(5), 696–705. https://doi.org/10.3758/s13421-016-0598-7

Lipko, A. R., & Dunlosky, J. (2007). Using feedback to improve grade school students judgments of text learning. *PsycEXTRA Dataset*. https://doi.org/10.1037/e658672007-001

Luna, K., Albuquerque, P. B., & Martín-Luengo, B. (2018). Cognitive load eliminates the effect of perceptual information on judgments of learning with sentences. *Memory & Cognition*, *47*(1), 106–116. https://doi.org/10.3758/s13421-018-0853-1

Maki, R. H. (1998). Test predictions over text material. In *Metacognition in Educational Theory and Practice* (pp. 131–158). Routledge. http://dx.doi.org/10.4324/9781410602350-13

Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 609–616. https://doi.org/10.1037/0278-7393.16.4.609

Maki, R. H., & Serra, M. (1992). The basis of test predictions for text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(1), 116–126. https://doi.org/10.1037/0278-7393.18.1.116

McDaniel, M. A., & Donnelly, C. M. (1996). Learning with analogy and elaborative interrogation. *Journal of Educational Psychology*, *88*(3), 508–519. https://doi.org/10.1037/0022-0663.88.3.508

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), 1–43. https://doi.org/10.1207/s1532690xci1401_1

McNamara, D. S., & Magliano, J. (2009). Chapter 9 toward a comprehensive model of comprehension. In *The Psychology of Learning and Motivation* (pp. 297–384). Elsevier. http://dx.doi.org/10.1016/s0079-7421(09)51009-2

Mendes, P. S., Luna, K., & Albuquerque, P. B. (2019). Word frequency effects on judgments of learning: More than just beliefs. *The Journal of General Psychology*, *148*(2), 124–148. https://doi.org/10.1080/00221309.2019.1706073

Metcalfe, J., & Dunlosky, J. (2008). Metamemory. In *Learning and Memory: A Comprehensive Reference* (pp. 349–362). Elsevier. http://dx.doi.org/10.1016/b978-012370509-9.00159-5

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*(1), 174–179. https://doi.org/10.3758/pbr.15.1.174

Miesner, M. T., & Maki, R. H. (2007). The role of test anxiety in absolute and relative metacomprehension accuracy. *European Journal of Cognitive Psychology*, *19*(4–5), 650–670. https://doi.org/10.1080/09541440701326196

Morris, C. C. (1990). Retrieval processes underlying confidence in comprehension judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(2), 223–232. https://doi.org/10.1037/0278-7393.16.2.223

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

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In *Psychology of Learning and Motivation* (pp. 125–173). Elsevier. http://dx.doi.org/10.1016/s0079-7421(08)60053-5

Nestojko, J. F., Bui, D. C., Kornell, N., & Bjork, E. L. (2014). Expecting to teach enhances learning and organization of knowledge in free recall of text passages. *Memory & Cognition*, *42*(7), 1038–1048. https://doi.org/10.3758/s13421-014-0416-z

Patall, E. A. (2013). Constructing motivation through choice, interest, and interestingness. *Journal of Educational Psychology*, *105*(2), 522–534. https://doi.org/10.1037/a0030307

Pitts, M. (1983). Comprehension Monitoring: Definition and Practice. *Journal of Reading*, *26*(6), 516–523.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111. https://doi.org/10.2307/271063

Rawson, K. A., & Dunlosky, J. (2002). Are performance predictions for text based on ease of processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 69–80. https://doi.org/10.1037/0278-7393.28.1.69

Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition*, *28*(6), 1004–1010. https://doi.org/10.3758/bf03209348

Redford, J. S., Thiede, K. W., Wiley, J., & Griffin, T. D. (2012). Concept mapping improves metacomprehension accuracy among 7th graders. *Learning and Instruction*, *22*(4), 262–270. https://doi.org/10.1016/j.learninstruc.2011.10.007

Renninger, K. A., Hidi, S., Krapp, A., & Renninger, A. (2014). *The Role of interest in Learning and Development*. Psychology Press.

Rhodes, M. G. (2015). *Judgments of learning*. Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199336746.013.4

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*(4), 615–625. https://doi.org/10.1037/a0013684

Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review*, *16*(3), 550–554. https://doi.org/10.3758/pbr.16.3.550

Sarmento, D. (2018). *Does Realism Harm Metacomprehension Accuracy?* [Thesis]. University of Illinois at Chicago.

Schraw, G., & Bruning, R. (1996). Readers' implicit models of reading. *Reading Research Quarterly*, *31*(3), 290–305. https://doi.org/10.1598/rrq.31.3.4

Schraw, G., & Roedel, T. D. (1994). Test difficulty and judgment bias. *Memory & Cognition*, *22*(1), 63–69. https://doi.org/10.3758/bf03202762

Schwartz, B. L., & Metcalfe, J. (2017). Metamemory: An update of critical findings. In *Learning and Memory: A Comprehensive Reference* (pp. 423–432). Elsevier. http://dx.doi.org/10.1016/b978-0-12-809324-5.21056-0

Shell, D. F., Colvin, C., & Bruning, R. H. (1995). Self-efficacy, attribution, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement-level differences. *Journal of Educational Psychology*, *87*(3), 386–398. https://doi.org/10.1037/0022-0663.87.3.386

Soderstrom, N. C., & Bjork, R. A. (2014). Testing facilitates the regulation of subsequent study time. *Journal of Memory and Language*, *73*, 99–115. https://doi.org/10.1016/j.jml.2014.03.003

Soderstrom, N. C., Yue, C. L., & Bjork, E. L. (2015). *Metamemory and education*. Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199336746.013.6

Son, L. K., & Kornell, N. (2008). Research on the allocation of study time. In *Handbook of Metamemory and Memory*. Routledge. http://dx.doi.org/10.4324/9780203805503.ch17

Tapiero, I. (2007). *Situation models and levels of coherence*. Routledge. http://dx.doi.org/10.4324/9780203810040

Taraban, R., Kerr, M., & Rynearson, K. (2004). Analytic and pragmatic factors in college students' metacognitive reading strategies. *Reading Psychology*, *25*(2), 67–81. https://doi.org/10.1080/02702710490435547

Tessmer, M., & Richey, R. C. (1997). The role of context in learning and instructional design. *Educational Technology Research and Development*, *45*(2), 85–115. https://doi.org/10.1007/bf02299526

Thiede, K. W., & Anderson, M. C. M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology*, *28*(2), 129–160. https://doi.org/10.1016/s0361-476x(02)00011-5

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*(1), 66–73. https://doi.org/10.1037/0022-0663.95.1.66

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 1024–1037. https://doi.org/10.1037/0278-7393.25.4.1024

Thiede, K. W., Dunlosky, J., Griffin, T. D., & Wiley, J. (2005). Understanding the delayed-keyword effect on metacomprehension accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1267–1280. https://doi.org/10.1037/0278-7393.31.6.1267

Thiede, K. W., Griffin, T. D., Wiley, J., & Anderson, M. C. M. (2010). Poor metacomprehension accuracy as a result of inappropriate cue use. *Discourse Processes*, *47*(4), 331–362. https://doi.org/10.1080/01638530902959927

Thiede, K. W., Griffin, T. D., Wiley, J., & Redford, J. S. (2009). Metacognitive Monitoring During and After Reading. In *Handbook of Metacognition in Education*. Routledge.

Thiede, K. W., Wiley, J., & Griffin, T. D. (2011). Test expectancy affects metacomprehension accuracy. *British Journal of Educational Psychology*, *81*(2), 264–273. https://doi.org/10.1348/135910710x510494

Thomas, A. K., & Mcdaniel, M. A. (2007). Metacomprehension for educationally relevant materials: Dramatic effects of encoding-retrieval interactions.

Psychonomic Bulletin & Review, *14*(2), 212–218. https://doi.org/10.3758/bf03194054

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10. https://doi.org/10.1007/bf02291170

Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology*, *21*(1), 437–484. https://doi.org/10.1146/annurev.ps.21.020170.002253

UK, D. (2017). How Crayons are Made [Video]. In *YouTube*. https://www.youtube.com/watch?v=lmiRjmbnn8Q

Undorf, M., Söllner, A., & Bröder, A. (2018). Simultaneous utilization of multiple cues in judgments of learning. *Memory & Cognition*, *46*(4), 507–519. https://doi.org/10.3758/s13421-017-0780-6

Undorf, M., & Zimdahl, M. F. (2019). Metamemory and memory for a wide range of font sizes: What is the contribution of perceptual fluency? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(1), 97–109. https://doi.org/10.1037/xlm0000571

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

Velicer, W. F., & Jackson, D. N. (1990). Component Analysis versus Common Factor Analysis: Some issues in Selecting an Appropriate Procedure. *Multivariate Behavioral Research*, *25*(1), 1–28. https://doi.org/10.1207/s15327906mbr2501_1

Vössing, J., & Stamov-Roßnagel, C. (2016). Boosting metacomprehension accuracy in computer-supported learning: The role of judgment task and judgment scope. *Computers in Human Behavior*, *54*, 73–82. https://doi.org/10.1016/j.chb.2015.07.066

Wiley, J., Griffin, T. D., Jaeger, A. J., Jarosz, A. F., Cushen, P. J., & Thiede, K. W. (2016). Improving metacomprehension accuracy in an undergraduate course context. *Journal of Experimental Psychology: Applied*, *22*(4), 393–405. https://doi.org/10.1037/xap0000096

Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *The Journal of General Psychology*, *132*(4), 408–428. https://doi.org/10.3200/genp.132.4.408-428

Wiley, J., Jaeger, A. J., Taylor, A. R., & Griffin, T. D. (2018). When analogies harm: The effects of analogies on metacomprehension. *Learning and Instruction*, *55*, 113–123. https://doi.org/10.1016/j.learninstruc.2017.10.001

Witherby, A. E., & Tauber, S. K. (2016). The concreteness effect on judgments of learning: Evaluating the contributions of fluency and beliefs. *Memory & Cognition*, *45*(4), 639–650. https://doi.org/10.3758/s13421-016-0681-0

Yan, V. X., Bjork, E. L., & Bjork, R. A. (2016). On the difficulty of mending metacognitive illusions: A priori theories, fluency effects, and misattributions of the interleaving benefit. *Journal of Experimental Psychology: General*, *145*(7), 918–933. https://doi.org/10.1037/xge0000177

Zimmerman, B. J. (2000). Self-Efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82–91. https://doi.org/10.1006/ceps.1999.1016

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 920–933. https://doi.org/10.1037/0278-7393.20.4.920