USE OF A SCAFFOLDED METACOGNITIVE INTERVENTION TO IMPROVE TRAINING OUTCOMES THROUGH A REDUCTION IN COGNITIVE LOAD

By

Morgan Brittany Showler

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Psychology-Master of Arts

2016

ABSTRACT

USE OF A SCAFFOLDED METACOGNITIVE INTERVENTION TO IMPROVE TRAINING OUTCOMES THROUGH A REDUCTION IN COGNITIVE LOAD

By

Morgan Brittany Showler

The present study sought to investigate how four metacognitive interventions influenced the degree of cognitive load experienced by learners during a complex training task, as well as how cognitive load subsequently affected learners' self-efficacy, procedural knowledge and skill performance. Additionally, cognitive load was examined as a mediator of the relationship between the metacognitive interventions and learning outcomes. Participants were randomly assigned to one of four metacognitive prompt conditions where they received prompts sequenced over time based on specificity during three training periods. Within each training period, participants received either a generic prompt at Time 1 and a generic prompt at Time 2, a generic prompt at Time 1 and a specific prompt at Time 2, a specific prompt at Time 1 and a generic prompt at Time 2, or a specific prompt at Time 1 and a specific prompt at Time 2. Results of a structural equation path analysis indicated that none of the four metacognitive interventions were significantly related to cognitive load or to any of the learning outcomes. Cognitive load was significantly related to self-efficacy such that participants who experienced greater cognitive load reported lower self-efficacy with regard to the training task; cognitive load was not significantly related to procedural knowledge or skill performance. Implications of these findings are discussed.

This work is dedicated to my family and cohort.

ACKNOWLEDGMENTS

I would like to thank my advisor, J. Kevin Ford, for his continuous patience and guidance throughout this learning process. His encouragement and approach to mentoring graduate students helped me greatly to develop an understanding of how to conceptualize and conduct high quality research, and also helped me to learn something about myself. The many critical pieces of suggestions and advice that he provided along the way were integral to completing this work. Additionally, I would like to thank my other committee members, Neal Schmitt and Christopher Nye, for providing valuable feedback that ultimately greatly improved the quality and rigor of my thesis.

I would also like to thank my wonderful family and friends for their never-ending love and support, although there are no words that can adequately express the depth of my gratitude. To my husband, Chris, my sister, Lauren, and my other family and friends, thank you for all of the unconditional emotional support you've given me over the last three years. To my amazing cohort mates, thank you for sharing in the struggles of graduate school by providing laughter and friendship, intellectual conversation, and always being there when I needed it most. Lastly, thank you to my father, whom I miss dearly, for shaping me into the person I am today, for helping me to believe in myself and teaching me what I am capable of, and for being a truly amazing role model and parent.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The modern workplace is characterized by a shift away from closed skill, consistent and clearly defined job roles to more knowledge-based, variable and open skill jobs. Open skill jobs are characterized by multiple ways of accomplishing a task without a set of concrete or prescribed steps, and varied and changing job roles and responsibilities that require a more general skillset. This transition towards open skill jobs and a more complex work environment has highlighted the need for employees to have an awareness of their skill strengths and weaknesses and to constantly monitor cues or changes in their job tasks and job environment, as well as be able to continually regulate their job-related behavior and strategize in order to perform successfully. These awareness and regulation skills are characteristic of metacognition, which represent a higher order understanding of a task and how an individual's cognitions and/or behaviors are related to progression towards a goal.

Metacognition has been shown to be a critical element of the learning process that helps to facilitate positive learning and performance outcomes (Garner & Alexander, 1989). However, research has shown that individuals typically do not naturally engage in metacognition often or very effectively. In recognition of this issue, some modern organizational training and educational initiatives have incorporated metacognitive interventions meant to encourage individuals to reflect on their learning progress in relation to learning goals and identify areas of strengths and weaknesses. This reflective process can then be utilized to initiate productive self-regulatory behaviors that allow for individuals to progressively learn and improve their skills and performance.

Despite the increasing use of such metacognitive interventions, research in this area has yielded mixed findings with regards to the effectiveness of the interventions in enhancing metacognitive activity and subsequent learning outcomes, as well as with identifying which types of metacognitve interventions are most effective. One potential reason for these inconsistent findings is that most research presently does not take into account the robust finding in the educational literature that learners' needs change as they progress through skill acquisition. Changing learner needs suggest that metacognitive interventions may also need to change over time in order to remain effective. Thus, the critical question should be which prompts work best *when*; consideration of this temporal element is missing from most metacognitive interventions presently used.

The proposed study examines how metacognitive interventions, and specifically metacognitive prompts, can be scaffolded or sequenced according to cognitive load theory to better support and facilitate learners' metacognitive activity through a reduction in cognitive load as their needs and abilities develop and change during the learning process. The study contributes to the literature in three primary ways: 1) by investigating how altering prompts over time may facilitate learning outcomes, 2) by sequencing prompt types to align with learner developmental needs through a novel application of cognitive load theory and scaffolding to the design of metacognitive prompting, and by 3) directly investigating perceptions of cognitive load as a mechanism influencing the relationship between metacognitive interventions and learning outcomes.

The present paper is organized as follows. First, I discuss in more detail the nature of metacognition and its role in the learning process and summarize the various conceptualizations of metacognitive prompts, a specific type of metacognitive intervention, that have been used in the literature. Second, I present a model of the proposed linkages between the metacognitive intervention, cognitive load and learning outcomes. Third, I describe cognitive load theory and highlight why it may be related to training outcomes. Fourth, I explain the concept of scaffolding and how it can be applied to the sequencing of metacognitive prompt interventions to best align prompt types with learner needs over time. And finally, I discuss the role of cognitive load as a mediating mechanism between the metacognitive intervention and learning outcomes. Whereas previous research has tested models connecting metacognitive activity to learning outcomes with the assumption that cognitive load is driving this relationship, the proposed study directly tests this assumption by including cognitive load as a mediator. Additionally, it utilizes a novel metacognitive intervention not previously investigated that scaffolds two types of prompts, generic and specific prompts, in order to facilitate greater metacognition over time.

METACOGNITION

Metacognition has been broadly defined as “knowledge and awareness of one’s own cognitive processes and the ability to actively control and manage those processes” (Berthold, Nuckles & Renkl, 2000). It involves self-regulatory control through both planning and monitoring behavior, as well as evaluation of progress towards a goal (Keith & Frese, 2005). More specifically, it has often been conceptualized in the literature in terms of two dimensions: awareness and regulation (Fiorella, Vogel-Walcutt, & Fiore, 2012; Schmidt & Ford, 2003).

Awareness, which is also sometimes referred to as monitoring, refers to the monitoring aspect of cognitive processes during learning and includes identifying the task at hand and repeatedly assessing and evaluating progress towards a goal (Schmidt & Ford, 2003). Awareness also encompasses what has been called metacognitive knowledge, or knowledge about the specific tasks, strategies or situational variables that are relevant and act to influence cognition (Efklides, 2008; Pintrich, Wolters, & Baxter, 2000). Metacognitive knowledge also includes knowledge about cognitive processes and how proficient an individual is at them. (Efklides, 2008).

Regulation, sometimes referred to as control, refers to the continual identification and selection of strategies and behavior that maximize learning and/or performance, including where to expend effort and what to prioritize (Fiorella et al., 2012; Schmidt & Ford, 2003). Additionally, regulation includes judgments regarding the appropriateness of a strategy or response given situational information (Pintrich et al., 2000). These judgments are then evaluated based on the effectiveness of a given strategy (Efklides, 2008). Regulation has also been referred to as metacognitive skill.

Despite the largely accepted distinction of the dimensions of monitoring and regulation, the definition of metacognition has evolved over time to become somewhat convoluted; in particular, there is inconsistency regarding how to differentiate it from self-regulation, or even how to clearly separate out the two dimensions of metacognition. Some researchers maintain that metacognition consists of knowledge and monitoring (e.g., making judgments about the difficulty level of a task), whereas the actual planning and strategy selection is separate and constitutes self-regulation (Pintrich et al., 2000). To clarify how I will be referring to metacognition in the present study, I define it as consisting of both awareness and knowledge of one's cognitions and the task, as well as the ability to effectively regulate behavior during learning. This is consistent with the current predominant conceptualization of metacognition and also reflects the need for an individual to not only possess knowledge or awareness of tasks or strategies, but also the ability to then apply that information when making decisions or judgments during learning. Without the application or use of metacognitive knowledge, the benefits of metacognition are reduced.

The presence of metacognitive activity indicates that an individual has engaged in some aspect of metacognition. Greater metacognitive activity is thought to be associated with increased learning, which occurs as a result of a greater number of cognitions related to the active self-monitoring and adjustments that individuals engage in during the learning process (Kaufmann, 2004). In addition, this increased amount of cognition is thought to lead to a deeper level of processing related to the task.

As individuals learn a new task or skill, they are processing various pieces of information about the task and their own abilities as well as making decisions as to what to practice and how (Pressley and Afflerbach, 1995). Individuals who engage in more metacognitive activity can develop a better understanding of how these pieces of information are connected and fit together in relation to the task and are also better able to use this information, in combination with knowledge about what they are and are not good at, to select more effective strategies and behaviors to apply towards learning the task. As they learn and engage in practice, they are then better able to evaluate the effectiveness of those chosen strategies and why they may or may not have been useful, as well as connect that information to knowledge about the task (Pintrich et al., 2000). This allows learners to subsequently select better or more appropriate strategies as their skill level changes throughout the learning process. Thus, because of greater awareness of task- and self-related information and regulation of task strategies, individuals who engage in greater metacognitive activity should be more efficient in their learning and should produce better learning outcomes than individuals who engage in less metacognitive activity.

Metacognitive prompts

Metacognitive prompts (also called metacognitive cues and reflection prompts, among other terms) represent one type of metacognitive intervention that is used to induce metacognitive activity in individuals during the learning process. These prompts are meant to overcome superficial processing that learners are likely to engage in (if they do partake in self-reflection at all) and encourage them to more deeply and actively reflect on their learning progress (Berthold, et al., 2007). Various types of metacognitive prompts have been utilized in the literature. These prompts can be classified into two main categories: generic or specific.

Generic prompts are typically more open-ended and simplistic, pushing learners to engage in self-reflection without providing significant direction as to how to do so. Their key function is simply to remind learners to stop and think about the task at hand and/or their developmental progress. Examples of generic prompts include: “Right now we’re thinking…”

(Davis, 2003), "Use the next 15 minutes for reflection. Reflect critically on the course and outcome of your problem-solving process," (Ifenthaler, 2012) and "How can you best organize the structure of the learning content?" (Berthold et al., 2007). There are varying levels of specificity within the category of generic prompts, with some prompts providing no guidance at all while others indicate broad areas in which to focus attention, but the overarching constant is that the content tends to be at a higher level, more ambiguous and more abstract. The main argument for the use of generic prompts is that they allow the learner more control, which can make self-reflections tailored more to the needs of each learner, facilitate engagement in the learning process, and prevent frustration of having to respond to prompts that are not relevant (Lin & Lehman, 1997).

Conversely, specific prompts not only encourage learners to self-reflect, but also provide specific areas in which they should focus their attention and guidance as to how they should be thinking about their learning progress. These prompts are built on the notion that simply asking learners to stop and reflect does not provide sufficient direction because learners do not do a sufficient job of identifying what they should focus their attention and effort on, and that more concrete instruction is needed to facilitate effective metacognitive activity. Examples of specific prompts include: "What did I do wrong here? Does the solution make sense?" (Kramarski & Gutman, 2006), "I believe I solved the problem well because…" (Ifenthaler, 2012), and "What steps are you using to solve the problem?" (Hoffman & Spatariu, 2008). As is evidenced by the content of specific prompts, they tend to include more information, be longer in length and be more detailed. Specific prompts are also sometimes used together in combination to provide a kind of step-by-step guide for learners to follow.

Research on the effectiveness of specific versus generic prompts has not yielded consistent results. While some studies indicate that specific prompts are more effective in enhancing learning outcomes (e.g., Berthold et al., 2007), other studies have indicated that generic prompts are actually more beneficial to learners (e.g., Davis, 2003 and Ifenthaler, 2012). Overall, it appears that generic prompts are more frequently found to be useful, suggesting that they may be somewhat more effective in improving metacognition and learning outcomes; however, this relationship is not always consistently demonstrated.

Within the broader categories of specific and generic prompts, different prompt structures can further be differentiated. Metacognitive prompts may consist simply of instructions or background information on metacognition and its importance and instruct learners to engage in metacognition throughout the learning process (Schmidt & Ford, 2003). Alternatively, prompts may consist of questions about task or learning progress that require learners to think and reflect, and even write down answers (Berthold et al., 2007; Kramarski & Michalsky, 2009). Prompts may also be statements that provide instructions to learners or hints about strategizing, overcoming problems or other metacognition-related topics.

However, despite their growing inclusion in training interventions, previous research has produced conflicting findings regarding the effectiveness of metacognitive prompts that are used to induce metacognition. Research by Berthold, et al. (2007) found that metacognitive prompts in isolation did not improve learning outcomes in undergraduates receiving training in the writing of learning protocols. Other studies have indicated similar results, finding that metacognitive strategies or activity were not related to exam performance (Miller & Geraci, 2011) or software training performance (Caputi, Chan, & Jayasuriya, 2011).

Research has indicated that the instructions and framing that are used during metacognitive prompting may be critical to the effectiveness of these strategies (Keith & Frese, 2005). It follows that one possible reason why metacognitive prompts are sometimes unsuccessful in facilitating meta-cognition and improved learning is that they are not being presented to learners in a way that is most conducive to enabling effective metacognitive activity. During most metacognitive interventions, the same types of prompts are generally applied at one time point or multiple time points during the learning process.

For example, Hoffman and Spatariu (2008) presented only various specific metacognitive prompts (such as, "Have you solved similar problems before," and, "Can the problem be solved in steps,") randomly across a training period. Another study investigated the use of adaptive metacognitive or learning strategy prompts (participants were given specific prompts based on their lowest scores on respective questionnaires) as compared to no prompts or randomly assigned prompts during a learning protocol generation task (Schwonke et al., 2006). Despite their distinction among the different prompt types, all of the prompts were again specific in nature; for example, "Try to relate new terms, concepts, theories, and ideas to known ones." In fact, most of the studies that classify their prompts broadly as metacognitive typically utilize the specific prompt approach; generic prompts tend to be used only in studies that are differentiating between specific and generic prompt types.

Although several studies have compared the utility of generic and specific prompts, only one study was located that combined the two prompt types in some way, although they are referred to as cognitive and metacognitive prompts rather than specific and generic prompts (Berthold et al., 2007). Within this study, cognitive prompts can be thought of as representative of specific prompts; an example prompt was, "Which headings and subheadings enable you to arrange the learning contents in a logical order?" Although the metacognitive prompts were somewhat higher in specificity than most other generic prompts, they are still high level and offer a useful comparison to cognitive prompts in this study; an example prompt was, " Which main points haven't I understood yet?" The study used a no prompts condition, a specific prompts condition, a generic prompts condition and a combined prompts condition. However, the combined prompts condition only randomly mixed the two types of prompts; the prompts were not strategically sequenced.

Thus, a missing element of these traditional metacognitive interventions is the consideration of time and how different prompts might be given purposefully and strategically at different time points based on the development of learners and changes in their cognitive structure. This is a basic principle of scaffolding and is discussed in more detail in later sections. The overarching notion is that it is not a question of whether generic or specific prompts are more effective overall, as both have been shown to be useful in facilitating metacognitive activity in some situations. Instead, the more appropriate question to ask is when each prompt type might be used most effectively to enhance metacognition.

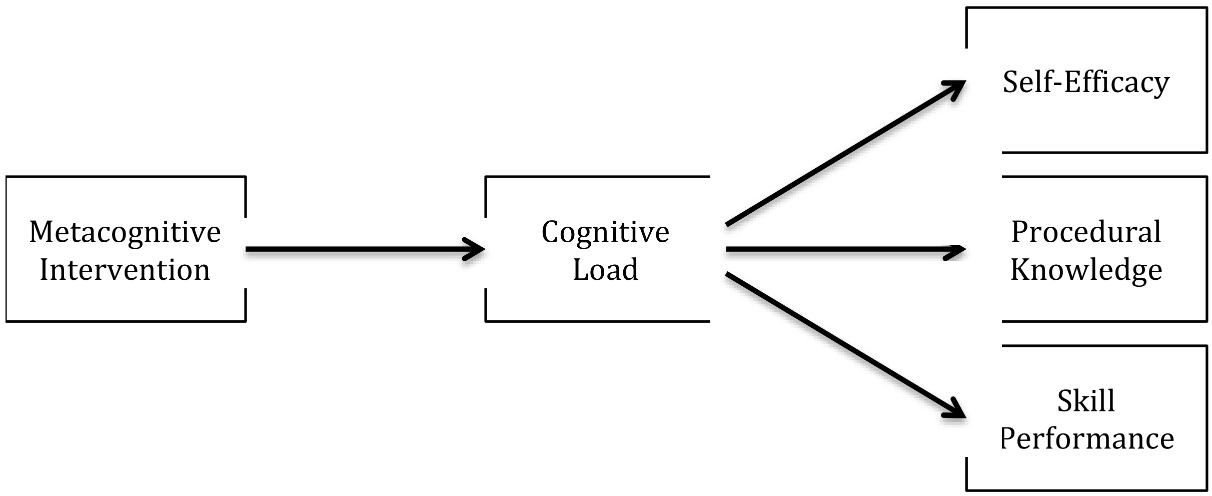
PROPOSED MODEL

To summarize what the literature has indicated to this point, metacognition has been shown to be important during the learning process by allowing individuals to reflect on their learning progress in order to effectively regulate their cognition and behavior. This metacognitive activity can be induced using metacognition interventions such as metacognitive prompts, and many different types of prompts have been used that vary in specificity, content and structure. However, research on metacognitive prompting has not provided consistent findings as to the most effective prompt type for increasing the amount of metacognitive activity learners engage in or for improving learning outcomes. So while it has been shown that metacognitive prompts are sometimes effective, we have not yet discovered why or under what conditions they result in maximized levels of metacognition and performance.

These gaps in the literature will be addressed in the proposed study in three ways. First, learners will be provided with multiple prompts over time in an effort to facilitate metacognition throughout the learning process. Secondly, it is important to note that the lack of consensus regarding which prompts work best is not the most relevant or appropriate question to ask. As individuals progress through the learning process, their knowledge and skills develop and change; it follows that any assistance or aid meant to help with their learning might also need to change, begging the question of *when* certain prompts are effective. The proposed study will investigate this by strategically sequencing different prompt types over time so that they better align with learner needs as those needs change during the learning process. Lastly, the proposed study will seek to address the question of *why* metacognitive prompts relate to learning outcomes. Some researchers have suggested that the mechanism by which metacognitive interventions influence learning outcomes is through a reduction in the amount of cognitive load that learners experience. However, no studies could be located that directly test this notion; the present study will contribute to the literature in this way by investigating cognitive load as the mediating mechanism between metacognitive interventions and learning outcomes.

Figure 1 presents a model of the hypothesized relationships among the metacognitive intervention, cognitive load, and three learning outcomes (self-efficacy, procedural knowledge and skill performance). The metacognitive intervention is expected to be negatively related to cognitive load, which is expected to be negatively related to self-efficacy, procedural knowledge and skill performance. In other words, it is expected that the target prompt sequence will reduce learners' cognitive load, thereby improving their learning outcomes. This is because the more cognitive load participants experience, the worse they should perform on the training task itself, the less knowledge they should accumulate about the training task, and the less confident they should feel about their ability to do well on the training task. The model is an adaptation of Schmidt and Ford's (2003) model relating a metacognitive intervention to metacognitive activity and learning outcomes, with the addition of cognitive load as a mediator of the relationship between the intervention and learning outcomes.

Figure 1
Hypothesized conceptual model of the metacognitive intervention, cognitive load and learning outcomes



The next sections will summarize the literature on cognitive load and scaffolding in order to develop hypotheses related to the proposed model.

Cognitive load theory

Cognitive load theory is based on the principle that individuals have a limited working memory capable of storing roughly seven novel pieces of information plus or minus two pieces, and that they can engage with only two to four pieces at any one given time (van Merrienboer & Sweller, 2005). These limitations are, however, reduced when working memory is engaging with information that is not novel and instead pulled from long-term memory (Ericsson and Kintsch, 1995).

Individuals with increased experience or knowledge tend to have already stored and organized large amounts of information into schemas in long-term memory which allows them to greatly reduce load on their working memory, as they can represent complex schemas containing multiple pieces of information using just one "slot" in working memory (van Merrienboer &

Sweller, 2005). Beginner learners would have to utilize multiple spaces for the same amount of information; this explains why more experienced learners are able to more quickly and effectively process situations and complex information as compared to less experienced learners.

According to cognitive load theory, learning and instructional design can be leveraged to encourage schema construction and application to task aspects that are likely to be present across situations in an effort to minimize cognitive load (van Merrienboer & Sweller, 2005). Two types of cognitive load may operate to constrain working memory: intrinsic cognitive load and extraneous cognitive load.

Intrinsic cognitive load refers to the inherent nature of information or task complexity and how it interacts with the expertise level of the learner. The level of intrinsic cognitive load is dependent on the degree to which various pieces of information must be simultaneously processed in working memory, which is known as element interactivity. The greater the amount of element interactivity required by a task, the greater the amount of cognitive load on working memory. Cognitive schemas are needed in order to effectively process the amount of information contained in tasks with high element interactivity, which fosters the creation of a coherent and clear understanding of the entire task or problem space. Because intrinsic load is based on the nature of both the learner and task, it cannot be altered through instructional strategy, according to traditional cognitive load theory. However, recent research sheds some additional light on this proposition and suggests that there may be ways of artificially reducing intrinsic load.

On the other hand, extraneous cognitive load is not an inherent element of learning and is instead related to how tasks are structured and presented to learners, and thus, can be altered through instructional strategy. Extraneous load can be unintentionally introduced if, for example, learning is designed in a way that presents multiple pieces of information to the learner in the same medium (i.e., if all instructions are written); this can overload a specific sensory processor of working memory.

Intrinsic and extraneous cognitive load are different but related concepts; extraneous cognitive load is a hindrance to the learning process mainly when intrinsic load is high. This is because extraneous load is more likely to overload working memory when there is already a high demand on working memory due to the intrinsic load of a task. Thus, instructional design should aim to reduce extraneous cognitive load to aid individuals during the learning process, especially in situations with high intrinsic load resulting from element interactivity. By reducing cognitive load, individuals should have more resources available to engage in metacognitive activity and should demonstrate improved learning outcomes.

Outcomes.

Previous research has shown metacognition to be related to a number of learning outcomes. For example, a positive relationship has been found between the amount of metacognitive activity a learner engages in and learning outcomes such as problem-solving in mathematics and physics (Pennequin, Sorel, Nanty, & Fontaine, 2010; Yarmohammadian & Asli-Azad, 2012), reading comprehension (Meloth, 1990; Xiangyan & Ji, 2007), language learning (Wenden, 1998) and simulation training (Fiorella, et al., 2012). As discussed previously, metacognitive prompts might be structured in ways that potentially reduce cognitive load, thus allowing learners to engage in greater metacognitive activity. Given that a substantive amount of previous research has already tested the relationship between metacognitive activity and learning outcomes, the present study will focus its investigation on cognitive load as the proposed mechanism acting as a link between metacognitive prompts and learning outcomes.

Kraiger, Ford and Salas (1993) developed a taxonomy of training outcomes that emphasizes the importance of conceptualizing learning as a multidimensional construct. Their premise, now universal in the study of learning outcomes, is that learning has affective, skill-based and cognitive components. Cognitive outcomes of learning may be knowledge-related (such as declarative knowledge and knowledge organization) or related to cognitive strategies or processes an individual engages in. Skill-based outcomes are more observable and behaviorally oriented and may be related to the skill acquisition process, such as compilation or automaticity (which represent more advanced stages of skill development). Affective outcomes may be attitudinal or motivational; motivational outcomes consist of constructs such as self-efficacy or goal setting.

In keeping with this foundational taxonomy, the present study investigates self-efficacy (affective component), skill performance (behavioral component) and procedural knowledge (cognitive component) as outcomes of cognitive load, metacognitive activity and the overall learning process.

Self-efficacy.

Self-efficacy has been broadly defined as “beliefs in one’s capabilities to mobilize motivation, cognitive resources, and courses of action needed to meet given situational demands” (Wood & Bandura, 1989). It has generally been conceptualized in the literature in terms of a situation-specific (SSE) or state-like construct that varies across situations, although it more recently has also been thought of as a global (GSE), or trait-like, construct that is consistent across situations as well. Traditional (specific) self-efficacy is thought of as an individual’s belief in his/her ability to perform well in relation to a specific task. However, general self-efficacy is more closely related to individual difference constructs such as goal orientation, and represents a more general belief in an individual's ability to meet demands and typically perform well across tasks and situations (Schwoerer, May, Hollensbe, & Mencl, 2005). This study investigates specific self-efficacy (hereafter referred to as self-efficacy) as a learning outcome because the degree to which metacognition affects self-efficacy in relation to the specific learning task is a more proximal outcome, and also reflects an effort to match the level of outcome specificity with the learning task itself.

Metacognitive activity has been found to be positively related to self-efficacy such that individuals who engage in greater metacognitive activity while learning how to perform a task tend to have higher self-efficacy beliefs (Cera, Mancini, & Antonietti, A., 2013; Landine & Stewart, 1998; Schmidt & Ford, 2003). One reason for this might be that because individuals who engage in greater metacognitive activity are better able to identify when they are struggling and adjust their behavior as a result, they then develop more complete task-related knowledge and more effective strategies that give them a greater confidence in their ability to learn the task and perform well, as has been demonstrated by some researchers (Ford, Smith, Weissbein, Gully, & Salas, 1998).

Reducing cognitive load through metacognitive prompting may improve learners' ability to effectively engage in metacognitive activity. Individuals who engage in greater metacognitive activity are likely to develop more effective practice strategies and are also more likely to perform at a higher level. More recent research expanding upon one of Bandura's originally proposed sources of efficacy information found that when individuals are trained in a way that allows them to experience mastery or success at performing some task aspects while learning, they develop higher self-efficacy (Combs & Luthans, 2007). Thus, individuals who engage in higher levels of metacognition are more likely to perform aspects of a task successfully during learning, which increases their feelings of mastery that then enhance their self-efficacy related to the task. Research has suggested that reducing cognitive load does in fact relate to higher self-efficacy, as it allows learners to improve their performance and build task-related confidence as a result (Zheng, McAlack, Wilmes, Kohler-Evans, & Williamson, 2009). I therefore hypothesize that reduced learner perceptions of cognitive load during the learning process will tend to result in increased task-specific self-efficacy.

Hypothesis 1a: Individuals who report experiencing less cognitive load during the training task will tend to report higher post-training self-efficacy.

Skill-based performance.

Skill-based performance refers to individuals' ability to execute numerous skills associated with the training task they will be learning. Various skill-based outcomes have been researched in the training literature, ranging from engagement and motivation on a medical simulation task to performance during an interpersonal skills role-playing activity (Wesiak, Steiner, Moore, Dagger, & Power, 2014). For example, one study investigated simulation-based training where participants learned to execute as a military fire support team and trained on how to decide when to initiate a "call for fire" command (Fiorella, Vogel-Walcutt, & Fiore, 2012). Trainees were then evaluated based on the learning outcomes of knowledge acquisition, skill in decision-making, and learning efficiency.

Other training studies have further investigated the relationship between metacognition and skill-based outcomes. A study involving web-based training instructing learners on how to create web pages implemented a metacognitive intervention consisting of an explanation of metacognition and its importance, self-questioning strategies asking trainees to identify how various concepts were related, and prompts asking them to evaluate their current understanding level. The study then investigated the amount of metacognitive activity that individuals engaged in, and whether that activity was related to performance on a skill-based measure (Ford & Schmidt, 2003). Findings indicated that learners who reported higher levels of metacognitive activity tended to perform better on the skill measure than learners who reported lower levels of metacognitive activity. Other research has also demonstrated this link between metacognitive activity and skill-related performance (Ford et al., 1998; Kramarski & Gutman, 2006).

Some research has further suggested that perceptions of lower cognitive load are related to the use of better cognitive and metacognitive strategies and higher performance levels (Scheiter, Gerjets, Vollmann, & Catrambone, 2009). One study indicated that for complex tasks where learners must monitor their performance and engage in self-regulation, learners reported increased cognitive load and performed more poorly (Van Gog, Kester, & Paas, 2011). This finding suggests that prompts might help to reduce the cognitive load associated with self-monitoring during complex training tasks, thereby improving performance. The relationship between metacognitive activity and performance is thus likely a result of the more effective strategies and increased knowledge afforded to learners who are able to engage in higher levels of metacognitive activity as a result of reduced cognitive load. As a result, it is expected that reducing cognitive load will lead to improved performance.

Hypothesis 1b: Individuals who report experiencing less cognitive load during the training task will tend to exhibit higher performance of task-related skills during the performance phase of the training task.

Procedural knowledge.

Three distinct types of knowledge are talked about during the process of knowledge acquisition: declarative knowledge, procedural knowledge and tacit knowledge (de Jong &

Ferguson-Hessler, 1996; Reber, 1989). The differences between these forms of knowledge are often characterized as information about *what* (declarative knowledge); *how* (procedural knowledge); and *which, when* and *why* (tacit knowledge; Kraiger et al., 1993). Thus, declarative knowledge represents facts and principles related to a task, procedural knowledge represents behavioral strategies for performing that task, and tacit knowledge represents practical know-how (not explicitly taught) that results in the "intuitive" selection of correct strategies across tasks or situations (Leonard & Insch, 2005).

Knowledge acquisition is typically thought of as a progression where declarative knowledge is learned first followed by procedural knowledge; once an individual learns general facts and principles related to a task, they are then ready to learn about how to apply that information towards actually performing the task itself (Ackerman, 1987). Because declarative knowledge is learned first, it has been suggested that it should be measured in the initial stages of training (Kraiger et al., 1993), whereas procedural knowledge tends to be more closely tied to actual performance on a task, making it more appropriate to be measured towards the end of training. As I am primarily interested in knowledge as a training outcome and in individuals' ability to demonstrate skill-based performance, I will assess procedural knowledge as a knowledge outcome. Tacit knowledge is learned in the next phase after procedural knowledge but differs in that it is not explicitly taught and tends to require high levels of experience in order to develop. The time scale for developing tacit knowledge is much longer as a result; because the present study constitutes a shorter time range than would be needed to develop tacit knowledge, I will focus on more easily defined and quickly developed procedural knowledge.

Although research has generally shown declarative knowledge to be a learning outcome of metacognitive activity (Schmidt & Ford, 2003; Schwonke, Hauser, Nuckles, & Renkl, 2006;

Wu & Looi, 2012), there has been less of a focus on the relationship between metacognition and procedural knowledge. Some support has, however, been found for a positive relationship between the two (Fiorella et al., 2012), and there are also conceptual reasons to expect this relationship. Because metacognition involves awareness of a task and the ability to apply appropriate behavioral strategies towards it, as well as evaluation and adjustment of task-related knowledge and behavior, individuals who engage in more metacognitive activity should be more likely to develop knowledge related to how to actually perform a task. However, if learners must deal with several pieces of complex information without any guidance, they may experience greater cognitive load that may in turn result in fewer cognitive resources available to develop and process knowledge related to the training task. Thus, I expect that individuals who experience reduced cognitive load will develop more knowledge of effective behavioral task strategies and compile greater knowledge of how to perform the task.

> **Hypothesis 1c**: Individuals who report experiencing less cognitive load during the training task will tend to demonstrate more post-training procedural knowledge.

Scaffolding

Scaffolding is an instructional design strategy that has been researched most extensively in the educational literature and refers to a form of strategy or tool for guiding individuals through the learning process in a way that facilitates awareness and self-regulation (Azevedo, Cromley & Seibert, 2004; Najjar, 2008). Scaffolding facilitates learning by supporting learners as they engage with content that may be overwhelming to them if they are left to their own devices without any guidance (Renninger & Granott, 2005). The general notion behind scaffolding relevant to this context is that what is the most effective strategy for structuring learning for a beginner learner or novice may not be most effective for more experienced learners who are further along in the development process. This highlights the need to consider phase-appropriate learning and instructional strategies for individuals relative to their position on the development continuum (van Merrienboer & Sweller, 2005).

Scaffolding strategies can generally be classified into two categories: soft scaffolds and hard scaffolds (Saye & Brush, 2002). Soft scaffolds typically consist of active, real time support provided mainly by a teacher or other guide in response to specific learners' needs as they occur (Berk & Winsler, 1995; Saye & Brush, 2002). In contrast, hard scaffolds are more standardized rather than adaptive, and can be further divided into conceptual and strategic support (Hannafin et al., 1999). Conceptual scaffolds provide hints or cues as to what information the learner should focus on, while strategic scaffolds provide guidance as to how to approach or analyze a situation (Simons & Klein, 2007). Hard scaffolding is more representative of the form of most metacognitive prompts, so I focus my discussion of scaffolding on this particular scaffolding type.

Positive evidence has generally been found for the use of hard scaffolds in supporting learning (Simons & Klein, 2007). More specifically, and relevant to metacognition, hard scaffolds have been shown to positively affect knowledge acquisition (Roehler & Cantlon, 1997), concept integration (Saye & Brush, 2002) and reflection (Davis & Linn, 2000). One study used scaffolding principles to sequence prompts. Specifically, researchers staggered prompts that targeted different aspects of metacognition throughout different learning phases of a pre-service training program for teachers (Kramarski & Michalsky, 2009). Comprehension prompts were administered during the planning phase, strategy question prompts during the action performance phase, and reflection question prompts during the evaluation phase.

Results indicated that reflection prompts presented during the evaluation phase were most effective followed by the comprehension prompts presented during the planning phase, but that strategy prompts presented during the action phase were not useful (Kramarski & Michalsky, 2009). However, it's unclear if the findings are indicative of prompts being more effective during certain learning phases rather than others, or if the content of the learning phases is more important, or if the two interact. This indicates that scaffolding may be a useful application to metacognitive prompts in that prompts are differentially useful across learning phases, but exactly what types of prompts to scaffold has not yet been identified. For the factors highlighted in the section on metacognitive prompts, there are viable reasons why specific and generic prompts might constitute the right type of prompt categories to sequence. While a small number of studies have begun to incorporate scaffolding principles into prompt design and implementation, and other studies discussed in the previous section have compared generic and specific prompts, no studies have scaffolded these two types of prompts in combination together during learning.

In summary, the potential value in applying scaffolding principles to metacognitive interventions has prompted some researchers to identify this as a possible evolution for metacognitive prompting, as well as a fruitful avenue for future research (Kramarski & Michalsky, 2009). However, we do not yet fully understand how to design scaffolding in an optimal way for enhancing learning or why it works when it is effective (Renninger & Granott, 2005). I turn now to research on cognitive load and cognitive load theory for insight as to how to design scaffolding in the proposed study so that it reduces cognitive load through consideration of learners' developmental progress.

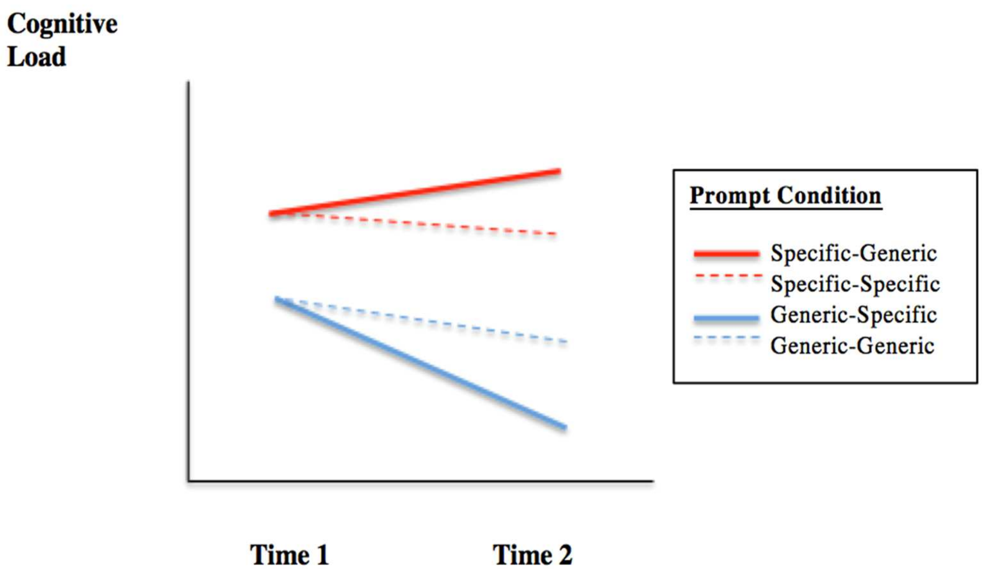
As was mentioned in the section on cognitive load, several methods for minimizing extraneous cognitive load have been researched and found to be effective; one such technique is to reduce the reading level of content by simplifying vocabulary and sentence length. However, much of the research related to cognitive load theory has not investigated cognitive load within highly complex tasks such as the task involved in the proposed study. More recent developments related to cognitive load theory suggest that in highly complex situations, element interactivity and demands on working memory may still remain at a high enough level to prevent learning, even after the reduction of extraneous cognitive load.

One strategy in this area that shows promise in further reducing cognitive load is to sequence instruction into two parts that present individual pieces of information initially, and then present all of the information together later. This progressive, isolated-followed-by-interacting-elements strategy effectively reduces intrinsic load by lowering the amount of element interactivity learners experience initially and has been identified as useful in complex situations for beginner learners who have not yet developed schemas that enable them to effectively deal with the amount of information that needs to be processed. Research in this area is now moving towards identification of which sequencing techniques are most effective in reducing cognitive load; this draws a natural connection to research on scaffolding and also highlights the relevance of the present study. Scaffolding can be used to inform prompt design, and to this end, the present study will employ four different prompt sequences: a generic prompt followed by a generic prompt, a specific prompt followed by a specific prompt, a generic prompt followed by a specific prompt, and a specific prompt followed by a generic prompt.

One reason why specific prompts may not be initially beneficial to learners is related to the finding that given process worksheets to reference during problem-solving, which (similar to the focused and/or step-by-step nature of specific prompts) provide instructions and hints as to what information or cues to attend to and how to go about problem-solving, are not effective in reducing cognitive load (Nadolski, Kirschner & van Merrienboer, 2005). Such process worksheets may not be effective because although they are meant as an aid that could potentially reduce cognitive load by providing information directly to learners, the information provided has high element interactivity and forces learners to divide their attention between the process worksheet and the task at hand (Kester, Kirschner, van Merrienboer, & Baumer, 2001). It may be the case that metacognitive prompting can actually distract learners from the primary learning task (Cannon-Bowers & Bowers, 2009). Similarly, specific or directed prompts may have the same effects on cognitive load; it may therefore be beneficial for learners to engage with simplified, more generic prompts initially before more detailed, specific prompts that have higher information interactivity and produce greater cognitive load. As a result, it is expected that the generic-specific prompt sequencing condition will more closely match learners' developmental needs during the training task and will result in reduced cognitive load.

> **Hypothesis 2:** Prompt sequence will interact with prompt type such that individuals in the generic-specific condition who receive a generic prompt at Time 1 and a specific prompt at Time 2 will report experiencing less cognitive load.

Figure 2
Expected interaction between prompt condition and cognitive load



No other specific directional hypotheses are made with regard to prompt condition, as it is unclear which of the other prompt sequences are likely to produce better or worse learning outcomes. It might be expected that receiving generic prompts at both time points is more effective than receiving specific prompts at both time points because the generic prompts may at least not overload learners during initial stages of learning, whereas specific prompts might for example reduce the amount of metacognitive knowledge that learners develop, which may result in lower overall learning. However, as discussed previously, there is not a clear consensus as to whether generic or specific prompts in isolation are more effective. As a result, all prompt conditions will be compared to determine if the sequencing of the prompts does in fact affect the level of metacognitive activity and subsequent learning outcomes.

As previous research has indicated that there is a relationship between metacognitive interventions and learning outcomes, the metacognitive prompts should also have an indirect effect on the learning outcomes in this study. Because the prompts are directing individuals' attention to the training task and to information related to it, it is reasonable to expect that learning outcomes such as procedural knowledge are likely to increase for many of the reasons discussed previously. The prompts also draw participants' attention to their learning progress and encourage them to think more intentionally about how to improve, which might be expected to increase performance improvements that produce higher overall performance levels. Additionally, numerous studies have shown an increase in reported metacognitive activity by individuals who receive metacognitive prompts as compared to individuals who don't receive prompts (Bannert, 2006; Stadtler & Bromme, 2008). These findings have been suggested to be related to a reduction in cognitive load experienced by learners as a result of prompting, allowing for increased metacognitive activity. Cognitive load should therefore serve as the mechanism by which metacognitive interventions operate to affect learning outcomes.

Hypothesis 3: Cognitive load will mediate the relationship between the metacognitive prompts and the three learning outcomes of self-efficacy, procedural knowledge and skill performance.

METHOD

Sample

Participants were 297 undergraduate students at a large Midwestern university who signed up to participate in the present study in exchange for psychology course research participation credits. The majority of the sample was female (62%) and Caucasian (63%); the remaining portion of the sample was 14% Asian, 9% Hispanic, 8% Black, and 6% other ethnic groups. The average age of participants was 19.69 years of age.

Research design

The study used a 2 (time 1: generic prompt, specific prompt) x 2 (time 2: generic prompt, specific prompt) between subjects design in a laboratory setting. A lab setting was used in order to increase experimenter control and internal validity, allowing for the evaluation of causal linkages. There were 74 participants in the generic-generic and specific-specific conditions, 76 participants in the generic-specific condition, and 73 participants in the specific-generic condition.

Procedure

Participants were recruited using Michigan State University's HPR system, an online experimental management system where students signed up to voluntarily participate in the study. They received four research participation course credits worth a half hour each in exchange for their participation.

Upon entering the computer lab where the experiment took place, participants signed in with the experimenter and were then told to have a seat at an available computer and begin reviewing and completing the consent form, followed by the demographics, goal orientation and cognitive ability measures. After the study start time, the experimenter reviewed the consent form with participants and highlighted the general activities participants would engage in. Once all participants finished these measures, the experimenter played a video of a powerpoint presentation that explained the TANDEM training task and the overall learning goals of the task, along with the specific rules and principles participants should apply. Participants were given instructions as to how to navigate the training program and how to perform the training task. Specifically, the presentation covered how to “hook” contacts, how to zoom in, and how to proceed with decision-making.

Participants then had three minutes to study the training manual material on the computer, after which they completed a one-minute familiarization phase. This short familiarization phase consisted of a single trial that introduced TANDEM and allow participants to become acclimated to the training controls and task environment without affecting their experience or competency levels with the task; participants’ performance was not be recorded during this trial. After the familiarization phase, participants began the training phase that consisted of a series of nine trials grouped into three blocks containing three trials each. At the beginning of each block, participants were given training topics to focus on and learning and task instructions for how to practice the task during the three trials within that block. At the beginning of each of the training trials, participants were given two minutes to look at the training materials and then completed the three-minute practice trials.

Prompts were administered to participants after the first and second trials of each training block (a total of three prompt sequences of two prompts each). This should have allowed participants to have practiced the specific training topics for that block sufficiently enough to understand them and have followed task instructions without mastering the training topics, ensuring that the prompts could help participants to do so in the following trial through reflection and/or strategizing. After each trial, participants were given one minute of feedback providing them with information as to how they performed during that trial. After the training phase, participants engaged in a final four-minute performance trial that their performance score was based on. The total experimental session time was approximately two hours.

The introduction, familiarization, training, and performance phases were structured similarly to how TANDEM has been used in other research (e.g., Kozlowski & Bell, 2008). The three training blocks focused on successively more advanced aspects of the training task, with the first two blocks focusing on key functions. The first training block instructed participants on how to practice hooking targets, understanding cue values and making final decisions for engagement with targets. The second training block focused on zooming, marker targets and defending the penalty perimeters. The third training block emphasized combining skills learned in the first two blocks to prioritize targets and make tradeoff decisions when protecting inner and outer penalty perimeters. The final performance trial required participants to apply all of the skills that they practiced in the previous three training blocks; participants were told at the beginning of the performance trial that their score for the trial would be recorded.

When the performance trial ended, participants completed the cognitive load, self-efficacy and procedural knowledge measures. After completing those measures, they were debriefed and thanked for their participation before exiting the lab.

Materials

Training task.

The training task within this study was a dynamic, computer-based radar simulation known as the Tactical Naval Decision-Making System, or TANDEM. TANDEM provided a complex and challenging learning environment with a high degree of experimenter control.

TANDEM was developed by Dwyer and colleagues (1992) and places learners in the role of a U.S. navy radar operator in charge of making decisions as to whether to fire at targets on a radar screen or clear them without firing. Participants first learn how to identify targets on the screen and classify them as either hostile or peaceful (intent) based on various pieces of information they collect, including a target's type (air, surface or submarine) and class (civilian or military). They must then decide to either fire or clear each target, with the overall goal being to as quickly as possible correctly classify targets and decide whether or not to fire at them. However, participants also have to monitor penalty boundaries around their own ship within the radar screen and lose points if any target enters those zones (Ford et al., 1998). This requires participants to prioritize targets and decide which to focus on initially when making decisions to clear or fire.

TANDEM can be structured to present learners with varying degrees of situational complexity. For example, the number of targets that appear simultaneously and the location where they appear (on the outside of the radar screen versus popping up near penalty circles) can be manipulated to increase or decrease cognitive demands and the amount of information that must be processed. Because the present study is focused mainly on the effects of the metacognitive prompt combination and scaffolding, a standard level of complexity that has been used in other studies involving TANDEM was implemented.

The TANDEM task has been used extensively to research a number of topics, including teamwork stressors (Weaver, Bowers, Salas, & Cannon-Bowers, 1995), learning strategies (Ford et al., 1998), adaptability (Kozlowski, Gully, Brown, Salas, Smith & Nason, 2001; Bell & Kozlowski, 2002) and other research streams. Relevant to the present study, there were several advantages to using a complex and challenging training task such as TANDEM. First, it necessitates the need for metacognitive activity and self-reflection in order for skill progression to occur in a way that allows learners to develop proficiency at the task. All learners are likely to experience difficulty and multiple instances of error that will require them to process information about the task along with information about their particular individual strategies and performance during the course of the training task. Thus, TANDEM should inherently facilitate the need for metacognitive activity. Additionally, because TANDEM is both complex and challenging, but not impossibly so (which could produce low performance variability as a result of most participants having low procedural knowledge and poor performance) and is also not too simplistic (which could produce low performance variability as a result of most participants developing high procedural knowledge levels and performing very well), there was sufficient variability in skill performance on TANDEM. TANDEM also appeared to induce a moderate degree of cognitive load in participants, likely due to the high task complexity and information required to perform TANDEM. This made TANDEM a good environment in which to study the effects of metacognitive prompting on metacognition and learning outcomes.

TANDEM was programmed to fit the proposed study's requirements and pilot testing was conducted with a group of 45 undergraduate students to verify that the TANDEM task functioned properly, the instructions were sufficient, the training phase and performance phase trials created variability in performance, and that all measures demonstrated adequate reliability. No major issues were identified during the piloting phase.

Metacognitive intervention.

The metacognitive intervention consisted of a series of two metacognitive prompts presented to participants during the training task. These metacognitive prompts were meant to facilitate participants' active reflection of their learning and progress on the training task. There were two different prompt types: generic prompts and specific prompts.

Generic prompts consisted of statements that asked participants only to stop and reflect on what they had just done in the training task. These prompts provided no guidance or instructions as to how participants should reflect or what aspects of the training task or their performance they should focus their attention and effort on. Thus, generic prompts provided only a suggestion that reflection should occur while allowing the participants to control what they reflected on. The generic prompt used in this study was an adaptation of prompts used by Ifenthaler (2012) and was as follows: "Use the next minute to reflect on what you just did in the training task".

Specific prompts consisted of statements that not only asked participants to stop and reflect, but also provided specific direction as to what participants should reflect on. Previous research related to specific prompts has tended to use a series of items targeting aspects of both the task and an individual's performance or progress relative to learning that task. Consistent with this approach, the specific prompt format in the present study consisted of three subtasks to direct participants' reflection efforts. These subtasks were adaptations of specific prompts used by Davis (2003) and Ifenthaler (2012) and included: 1) "Right now I'm doing a good job of…", 2) "Right now I'm struggling with…", and 3) "In order to improve, I'm going to...". These prompts asked participants to provide short open-ended answers to each question during a one minute and thirty second time period and were presented immediately following the same set of reflection instructions given to participants in the generic condition. The three prompts were meant to direct learners to reflect on the task structure and their learning progress (monitoring aspects) as well as to consider how they could improve (planning and regulation aspect).

Ratings of prompt specificity were collected from graduate students on a scale ranging from 1 (Very General) to 5 (Very Specific) for 44 metacognitive prompts taken from the literature in order to verify the specificity classifications. Prompts that were longer were generally rated as more specific whereas prompts that were shorter were generally rated as more general regardless of the content contained within the prompt, making the specificity ratings difficult to interpret. The conceptualizations of specific and generic prompts described above, which were based on previous research, were retained for the study.

Prompt conditions were randomly ordered and assigned to study sessions before data collection; all participants within a session received the same prompt manipulation. The four conditions differed in terms of the combination and sequencing of prompts, as depicted in the table below. The cell of interest expected to produce the most metacognitive activity and best learning outcomes was the generic to specific prompt progression denoted in red. The prompts were administered during the training task by means of a Qualtrics survey link that opened a separate window on participants' computer screens. Participants were not able to continue with the training task until the one-minute mandatory reflection period ended. Prompt 1 was administered after the first trial in each block during the training phase (Time 1), while Prompt 2 was administered after the second trial during the training phase (Time 2). Because each trail lasted for only three minutes, Time 1 (three minutes in) should have represented earlier stages of learning and skill acquisition where participants were beginning to be able to perform the task but were still making frequent errors and were relative novices. Time 2 (six minutes in) should represent middle stages of learning where participants were now making fewer mistakes, developing competency, and were able to perform the training task more fluently.

Figure 3
Metacognitive prompt conditions

Prompt Type		Time 1	
		Generic	**Specific**
Time 2	**Generic**	Generic, Generic	Specific, Generic
	Specific	Generic, Specific	Specific, Specific

Cognitive load.

The cognitive load experienced by participants during the training task was assessed using a six-item mental workload scale adapted from the NASA-Task Load Index (Hart & Staveland, 1988; Appendix C). Items were adapted from each of the six NASA-TLX dimensions: mental demand, temporal demand, physical demand, performance, perceived effort, and frustration. Sample items included, “How hurried or rushed did you feel while learning and practicing the training task?” and, “How mentally demanding was learning about and practicing the training task?” Participants indicated their responses on five-point scales ranging from “Not at all” to “Very”.

The reliability of this scale ($\alpha = 0.67$) was initially found to be below an acceptable level of 0.70; examination of item level statistics indicated that the recoded item four was not correlated as highly with the other items in the scale and that reliability would increase to 0.70 if this item were removed from the scale. The content of this item differed from the other items in that it asked for an overall evaluation of success in accomplishing training objectives rather than asking for an indication of, for example, how frustrated or rushed participants felt, which may have contributed to reduced internal scale consistency. This item was thus removed from the cognitive load scale for future analyses, producing a reliability value more consistent with other research (0.83; Hart & Staveland, 1988).

Self-efficacy.

Participants' specific self-efficacy related to the TANDEM task was assessed using a six-item measure adapted from Ford, Smith, Weissbein, Gully and Salas (1998; Appendix C). Participants indicated their responses on a five-point scale ranging from 1, "Strongly disagree", to 5, "Strongly agree". Sample items included, "I can meet the challenges of this simulation," and "I am confident that I can perform the TANDEM task well". This scale had good internal consistency ($\alpha = 0.89$), similar to other research that has shown internal consistency to be at or above 0.9 (Ford et al., 1998).

Procedural knowledge.

The present study was interested in individuals' procedural understanding of the task rather than more basic or declarative knowledge (e.g., knowledge of decision rules, cue identification); if an individual has strong procedural knowledge, they should inherently also possess strong foundational declarative knowledge. Procedural knowledge within the context of TANDEM consisted of information related to the prioritization of actions, target (marker) contacts, and the zooming function of the training task. This differentiated approach to assessing knowledge has been used previously (Bell & Kozlowski, 2002) and has found declarative knowledge and procedural knowledge to be distinct knowledge domains. Procedural knowledge was operationalized as participants' scores on a questionnaire adapted from Baard (2012; Appendix C) that asked them to answer 12 procedural knowledge questions by selecting the best answer from one of four answer choices. Sample items included, "If three contacts are about 10 miles outside your Outer Defensive Perimeter, which of the following should you do to prioritize the contacts," and, "If you Zoom-Out to find three contacts around your Outer Perimeter, how would you determine which contact is the marker contact?"

Skill performance.

Skill performance within TANDEM was contingent on the decision points discussed earlier; specifically, the degree to which participants were able to identify targets, classify them, and prevent them from entering penalty boundaries. For the present study, performance was operationalized consistently with how it is normally assessed within the TANDEM task (e.g., Bell & Kozlowski, 2008). This involved a point system where a participant received 100 points when all of the required decisions (type, class, intent) related to targets were made correctly and lost 100 points if any of the decisions were incorrect, or 10 points if a target crossed a penalty boundary. The total number of points participants obtained during the performance trial at the end of the study constituted their skill performance score.

Manipulation check.

Metacognitive activity. A check for metacognitive activity was conducted in order to examine the nature of the reflection that the four metacognitive prompt conditions induced in participants. Metacognitive activity was assessed using the open-ended prompt responses that participants entered into the Qualtrics surveys. Participants' responses were coded for the total number of words and number of unique ideas described within each prompt response and an overall complexity rating was also provided for the entire set of a participants' prompt responses; coding instructions are located in Appendix D for reference. It was expected that the generic-specific condition would produce the greatest complexity because participants had more freedom to reflect and explore topics of their choice in response to the generic portion of the prompts early on in the learning process, which should then have lead to deeper reflection throughout the specific prompts later on as well. With regard to the number of unique ideas and words, it might be expected that the generic-specific prompt would produce the most ideas and words, again because of the depth of reflection that participants should engage in in response to that prompt scaffolding. However, because the design of the specific prompt included three sub-questions for participants to respond to, it might also be expected that the specific-specific condition would produce the greatest number of ideas and words; thus, there were no particular expectations based on condition for those variables.

Prompt responses were coded by two undergraduate research assistants who received an intial training and then practiced coding using pilot data. Both coders rated the entire set of 247 pilot prompt responses; interrater agreement was computed using intraclass correlation coefficients, which were interpreted as follows: > 0.75 was excellent, 0.40-0.75 was fair to good, and < 0.40 was poor (Fleiss, 1986). ICCs (3,1) for the number of unique ideas [$F(246) = 4.70$, $p < 0.01$] and complexity coding [$F(47) = 3.77$, $p < 0.01$] were found to be at or above the threshold of 0.4 for fair absolute agreement. Coders then coded 1624 prompt responses for the actual study; 75 of the prompt responses were coded by both research assistants so that ICCs could be computed, while the remaining 1549 responses were split up evenly. The ICC for the number of unique ideas was 0.69 [$F(74) = 6.51$, $p < 0.01$], indicating good agreement while the ICC for complexity was 0.40 [$F(12) = 5.91$, $p < 0.01$], indicating fair agreement.

A one-way ANOVA was conducted to determine if the coding indicated that complexity, the number of unique ideas, and the number of words varied by prompt condition. Results of the ANOVA were significant for the number of ideas [$F(3, 1620) = 27.91$, $p < 0.01$] and the number of words [$F(3, 1620) = 12.42$, $p < 0.01$]. Tukey's HSD was used as a post-hoc analysis to determine where differences among the prompt conditions existed. Overall, the generic-generic condition prompt responses contained fewer words and ideas than the other three conditions ($p < 0.01$), while the specific-specific condition contained the most ideas ($p < 0.01$); this is consistent with the design features of the two prompts mentioned previously but may also indicate that the specific-specific prompt facilitated a greater degree of reflection. There were no significant complexity differences among the four conditions [$F(3, 293) = 0.14$, p = ns], which did not support the expectation that learners in the generic-specific condition would engage in deeper and more complex reflection based on Hypothesis 2.

Controls.

Cognitive ability. Cognitive ability is typically defined as an individual's general level of intelligence. It has been shown to be an important individual difference variable during the learning process (Furnham, Swami, Arteche, Chamorro-Premuzic, 2008) and was therefore likely to affect learning during the training task in the present study. In fact, due to the complex nature of the training task, cognitive ability should play an even larger role, as complex tasks require increased attentional resources and information processing that individuals with higher cognitive ability are more proficient at. Previous research has shown that tasks with high complexity (similar to TANDEM) put an increased cognitive load on learners as a result of greater amounts of information processing that are required while engaging with the task, and that individuals with greater cognitive ability are better able to process and deal with larger amounts of information.

Consequently, cognitive ability is likely to influence learning outcomes such as skill performance and procedural knowledge, as both should be dependent on the ability to process task-related information in a complex training environment. I therefore controlled for cognitive ability in the present study using students' ACT or SAT test scores as an indicator. Participants will indicate their highest SAT or ACT test score on a self-report questionnaire. Previous research has suggested that ACT and SAT test scores are a highly reliable (0.96 for the ACT; American College Testing Program, 1989) and valid (e.g., Schmidt, 1988) measure of cognitive ability, and that self-reported SAT and ACT scores are highly correlated with actual test scores ($r = 0.94$; Gully, Payne, Kiechel, & Whiteman, 1999). This indicates that students typically tend to report their scores as close to their actual scores; as such, score inflation is not expected to be a significant issue with using this operationalization of cognitive ability.

Goal orientation. Goal orientation has been identified as an individual difference construct particularly relevant to learning contexts such as a training environment, as it emphasizes the types of goals that individuals tend to prioritize in achievement situations; this reflects trait rather than state goal orientation (Schmidt & Ford, 2003). The underlying notion of goal orientation relevant to training is that different individuals may learn best under different conditions based on their orientation towards either performance or mastery goals. The two main dimensions of goal orientation are learning (or mastery) goal orientation (LGO) and performance goal orientation (PGO). A learning orientation reflects a focus on the understanding of and improvement or development on a task, with emphasis on attaining personal goals and building competency (Schmidt & Ford, 2003). A performance goal orientation reflects a focus on demonstrating ability and receiving recognition, as well as an emphasis on outperforming others.

Previous research has shown that individuals with a LGO tend to adopt learning goals that focus on skill improvement, whereas individuals with a PGO tend to adopt learning goals that focus on avoiding appearing incompetent or on performing better than others (Brett & VandeWalle, 1999). Additionally, individuals who take a mastery approach to learning tend to be more resilient to mistakes and develop effective strategies for dealing with failure while individuals who take a performance approach tend to exhibit decreased performance and inability to respond adequately to failures or challenges (Dweck & Leggett, 1988). As a result, it is possible that in a complex training environment where experimenting with different strategies in increasingly complex scenarios may lead to increased learning but also to making errors, goal orientation may affect the extent to which individuals engage in metacognitive activity and self-reflection, as well as the level of their overall learning outcomes. I controlled for goal orientation in order to more clearly investigate the relationship between the intervention, metacognitive activity and the learning outcomes.

Trait goal orientation will be assessed using a 13-item measure adapted from VandeWalle (1997; Appendix C). The measure consists of three subscales; learning (mastery), performance prove and performance avoid. All three subscales have been demonstrated to have reliabilities above 0.8 (VandeWalle, 1997), which was also found to be the case in the present study (goal orientation prove: $\alpha = 0.82$, goal orientation avoid: $\alpha = 0.86$, goal orientation learning: $\alpha = 0.87$). An example item is, “I’m concerned with showing that I can perform better than my peers.” The measure asked respondents to indicate their answers on a six-point scale ranging from 1, “Strongly disagree” to 6, “Strongly agree”.

RESULTS

Preliminary analyses

Means, standard deviations and zero-order correlations among the variables are presented in Table 1 with coefficient alphas displayed on the diagonal. Correlations among the variables indicate that cognitive load was related to self-efficacy and skill performance in expected ways, but not to procedural knowledge. Additionally, all four control variables were related to at least one of the outcomes, so all four variables were controlled for initially in the path analysis. In particular, cognitive ability was positively related to self-efficacy, procedural knowledge and skill performance; goal orientation prove and avoid were positively related to cognitive load and goal orientation learn was negatively related to cognitive load; goal orientation prove and learn were positively related to self-efficacy and goal orientation avoid was negatively related to self-efficacy; while the goal orientation avoid scale was negatively related to skill performance. The generic-generic dummy coded variable was positively related to goal orientation learning while the specific-specific dummy coded variable was positively related to goal orientation prove and avoid. Additionally, procedural knowledge and skill performance were positively correlated with each other and with self-efficacy.

The scale reliability (Cronbach's alpha) for all original and revised scales was above a satisfactory level (greater than 0.70). A total of 10 participants with either performance scores or procedural knowledge test scores greater than two standard deviations below the mean were removed from the dataset, as those individuals were unlikely to have either sufficiently understood the training task or exerted effort during the performance trial.

Table 1
Means, standard deviations and correlations among variables

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
1. Generic-generic	-	-	-											
2. Generic-specific	-	-	-.34**	-										
3. Specific-specific	-	-	-.33**	-.34**	-									
4. Specific-generic	-	-	-.33**	-.34**	-.33*	-								
5. Cognitive load	3.09	.73	.04	.03	.01	-.08	(.70)							
6. Self-efficacy	3.56	.72	-.04	.03	.01	.01	-.19**	(.89)						
7. Proc. knowledge	7.74	2.33	-.08	-.01	.03	.05	-.03	.13*	-					
8. Skill performance	61.59	405.72	-.01	-.04	.08	-.03	-.14*	.35**	.30**	-				
9. Cog. ability (ACT)	25.42	3.85	-.01	-.02	.04	-.02	-.06	.14*	.30**	.30**	-			
10. Goal orientation, prove	4.32	0.99	-0.03	-0.05	0.12*	-0.04	0.18*	0.15**	0.01	0.03	0.05	(0.82)		
11. Goal orientation, avoid	3.27	1.01	-0.08	-0.01	0.12*	-0.04	0.24**	-0.26**	-0.08	-0.19**	-0.03	0.24**	(0.86)	
12. Goal orientation, learn	4.94	0.77	0.15**	-0.13*	0.02	-0.04	-0.13*	0.26**	-0.04	0.11	0.10	0.23**	-0.33**	(0.87)

Note. *p < 0.05; **p < 0.01. Many participants had negative performance scores that reduced the overall average performance score. Scores for goal orientation and self-efficacy were indicated on a 5-point scale ranging from “Strongly disagree” to “Strongly agree”, scores for cognitive load were indicated on a 5-point scale ranging from “Not at all” to “Very”, test scores for procedural knowledge ranged from 0 to 12, and skill performance scores ranged from -1080 to 920.

Table 2
Means, standard deviations and standardized mean differences (*d*) for outcome variables by prompt condition

	Cognitive Load	Self-Efficacy	Procedural Knowledge	Skill Performance
			Generic-Specific	
Mean	3.13	3.60	7.71	37.63
SD	0.61	0.65	2.35	404.77
			Generic-Generic	
Mean	3.15	3.52	7.42	53.38
SD	0.76	0.79	2.45	383.68
d	-0.03	0.11	0.12	-0.04
			Specific-Generic	
Mean	2.99	3.56	7.96	35.89
SD	0.72	0.72	2.19	458.33
d	0.19	0.06	-0.11	0.01
			Specific-Specific	
Mean	3.10	3.57	7.88	116.06
SD	0.81	0.71	2.34	388.18
d	0.04	0.04	-0.07	-0.19

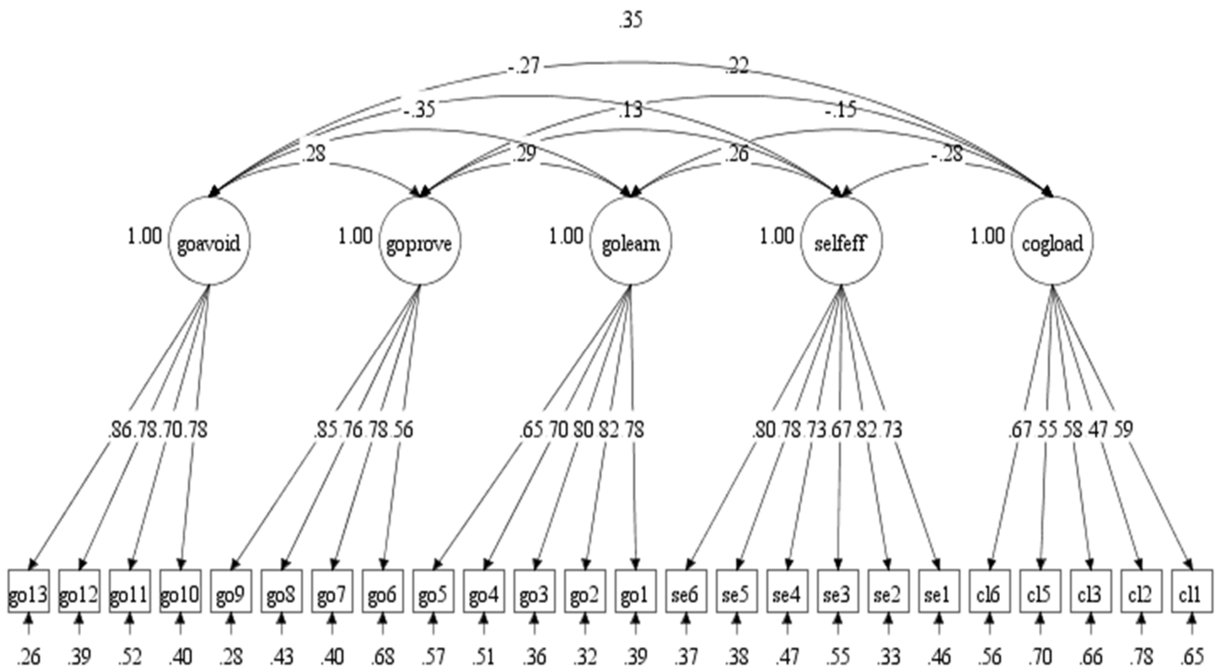
Note. Cohen's d values represent standardized mean differences between scores for the generic-specific condition and each of the other three prompt conditions.

Measurement model analysis

Before examining the hypothesized relationships among the constructs in the proposed structural model, a confirmatory factor analysis was conducted using Mplus Version 7.4 (Muthen & Muthen, 1998-2015) to test the measurement model and confirm the factor structure of the latent variable measures of cognitive load, self-efficacy and goal orientation. The hypothesized five factor model specified six items loading on the cognitive load factor, six items loading on the self-efficacy factor, five items loading on the learning goal orientation factor, four items loading on the prove goal orientation factor and four items loading on the avoid goal orientation factor.

Model fit is generally assessed as acceptable when chi-square values are non-significant (with caveats), CFI values are greater than 0.95, RMSEA values are less than 0.06 and SRMR values are less than 0.08 (Hu & Bentler, 1999). The results of the CFA indicated that model fit was marginally acceptable, $\chi^2(242) = 449.83$, $p < 0.01$; CFI = 0.93; RMSEA = 0.05; SRMR = 0.05. The SRMR and RMSEA values were within the desired rule of thumb values of 0.08 and 0.06, respectively. The statistical significance of the chi-square value indicates that the model does not fit well, although previous research has indicated this to be an inaccurate indicator of model fit because it is affected by sample size, model complexity and other factors. The CFI value of 0.93 is slightly outside the rule of thumb value of 0.95, indicating a marginal overall model fit. Relationships between each item and its respective factor were all significant ($p < 0.01$), with standardized coefficients ranging from 0.47 to 0.86, providing support for the hypothesized measurement model. An examination of item content and modification indices did not reveal any theoretically justifiable respecifications to the model; this integrated approach to respecification is recommended in order to prevent capitalizing on chance when changes are made to a model post-hoc (Anderson & Gerbing, 1988; Kenny, 2011). Overall, the measurement model adequately fit the data for the purposes of this study and no substantial issues were identified in terms of factor loadings for each of the items and their respective scales, so all scales were retained in their proposed form in future analyses.

Figure 4
Confirmatory factor analysis results for hypothesized measurement model



Structural model analysis

To test the hypothesized model and determine whether cognitive load fully mediates the relationship between the metacognitive intervention and procedural knowledge, skill performance and self-efficacy, a structural equation path model was tested using Mplus Version 7.4 (Muthen & Muthen, 1998-2015). This allowed for the simultaneous testing of all paths in the hypothesized model. Maximum likelihood estimation was used to test the proposed hypotheses. Cognitive ability and goal orientation were controlled for in the analyses by specifying additional paths in the model from the control variables to the mediator (cognitive load) and the outcome variables (self-efficacy, procedural knowledge, skill performance). In order to test the hypotheses, three dummy-coded variables were created to test the effects of the four conditions of the metacognitive intervention on cognitive load with the generic-specific prompts set as the referent group. The generic-specific prompt condition was hypothesized to be the most effective condition for reducing cognitive load, and it was thus of interest to have an indication of the effects of the other three conditions in relation to this target condition.

A preliminary path model testing the mediating effect of cognitive load on the relationship between the metacognitive intervention and learning outcomes without taking control variables into account was analyzed and was demonstrated to fit the data well, $\chi^2(9) = 5.13$, $p = ns$; CFI = 1.00; RMSEA = 0.00; SRMR = 0.02. The hypothesized model that also specified control paths from all control variables to all dependent variables was tested next and also exhibited good fit, $\chi^2(9) = 3.40$, $p = ns$; CFI = 1.00; RMSEA = 0.00; SRMR = 0.01; all non-significant control paths were then eliminated. Remaining control paths specified relationships from goal orientation avoid and prove to cognitive load and self-efficacy; from cognitive ability and goal orientation learn to procedural knowledge; and from cognitive ability and goal orientation avoid to skill performance. Overall, fit for the revised model was good as indicated by fit indices, $\chi^2(17) = 15.90$, $p = ns$; CFI = 1.00; RMSEA = 0.00; SRMR = 0.03.

Values of 1.00 for the CFI and 0.00 for the RMSEA may at first glance seem problematic but can occur as a result of the chi-square formula despite the model being over-identified, with fit statistics that are based on the chi-square formula affected (Kenny, 2015). Specifically, the numerators in the RMSEA and CFI formulas (Appendix E) require the chi-square degrees of freedom to be subtracted from the chi-square value, which produces values of 1.00 for the CFI and 0.00 for the RMSEA when a chi-square value is less than the degrees of freedom, as it is in the hypothesized path model. Interpretations of these fit statistics indicate that the model is actually quite well-fitting (Muthen, 2010). Overall, all fit indices were within the desired rules of thumb and the chi-square statistic was non-significant, indicating that the hypothesized model fit the data well. The model accounted for 8% of the variance in cognitive load ($p < 0.05$), 12% of the variance in self-efficacy ($p < 0.01$), 10% of the variance in skill performance ($p < 0.05$) and 11% of the variance in procedural knowledge ($p < 0.01$).

An alternative partial mediation model was tested that included the same paths as the hypothesized model but also specified direct paths between the dummy-coded metacognitive intervention variables and the three training outcomes of self-efficacy, procedural knowledge and skill performance. Overall fit for this alternative model was also good with similar fit indices to the hypothesized model: $\chi^2(8) = 12.55$, $p = ns$; CFI = 0.96; RMSEA = 0.05; SRMR = 0.03 and accounted for 8% of the variance in cognitive load ($p < 0.05$), 12% of the variance in self-efficacy ($p < 0.01$), 10% of the variance in skill performance ($p < 0.05$) and 12% of the variance in procedural knowledge ($p < 0.01$).

Because the hypothesized fully mediated model was nested within the alternative partially mediated model, a chi-square differences test was computed to compare overall fit between the two models. The value produced by the chi-square differences test was less than the critical value of $\chi^2(9) = 16.92$ listed at the $p < 0.05$ significance level in a chi-square table, indicating that there was no statistically significant difference in fit to the data between the two models, $\chi^2(9) = 3.35$, $p = ns$. Further, the addition of the nine direct path effects from the dummy-coded variables to the three outcomes were all non-significant, indicating that those modeled relationships did not fit the data. Thus, the hypothesized fully mediated model was more parsimonious and demonstrated similar fit to the alternative partially mediated model, so it was retained for hypothesis testing.

Hypothesis 1a indicated that individuals who report experiencing less cognitive load during the training task would tend to report higher post-training self-efficacy. Results of the path analysis indicate that hypothesis 1a was supported ($\beta = -0.18$, $p < 0.01$). Cognitive load was negatively associated with self-efficacy, indicating that participants who reported experiencing greater cognitive load also reported lower self-efficacy.

Hypothesis 1b indicated that individuals who report experiencing less cognitive load during the training task would tend to exhibit higher performance of task-related skills during the performance phase of the training task. Hypothesis 1b was not supported ($\beta = -0.05$, $p = ns$), indicating that cognitive load was not related to participants' levels of skill performance.

Hypothesis 1c indicated that individuals who report experiencing less cognitive load during the training task would tend to demonstrate more post-training procedural knowledge. Hypothesis 1c was not supported ($\beta = -0.06$, $p = ns$); thus, cognitive load was unrelated to participants' procedural knowledge scores.

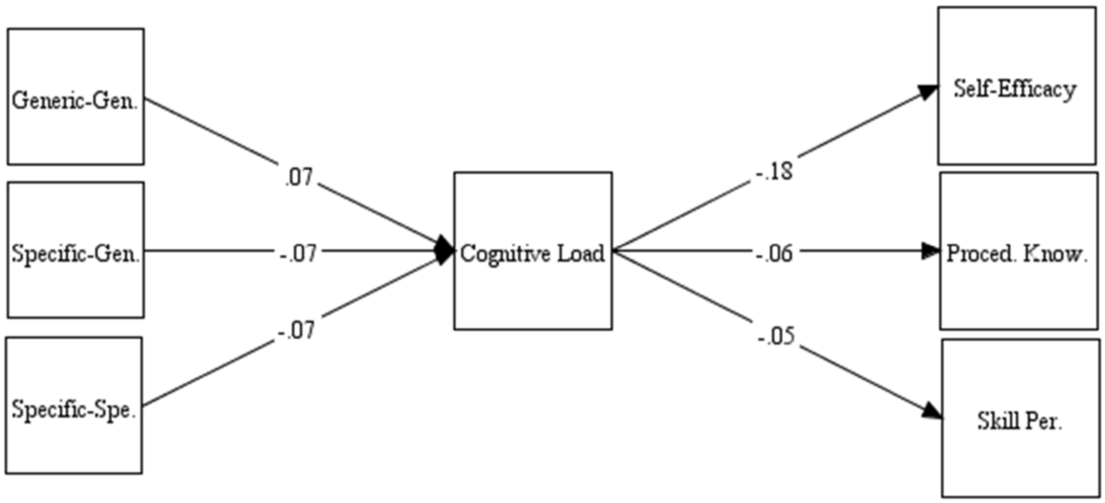
Hypothesis 2 indicated that time would interact with prompt type such that individuals in the generic-specific condition who received a generic prompt at Time 1 and a specific prompt at Time 2 would report experiencing less cognitive load. Hypothesis 2 was not supported; comparison of the beta estimates for the three dummy-coded condition variables, which indicate their effects on cognitive load relative to the referent group (generic-specific condition), were all non-significant (generic-generic condition, $\beta = 0.07$, $p = ns$; specific-generic condition, $\beta = -0.07$, $p = ns$; and specific-specific condition, $\beta = -0.07$, $p = ns$). This suggests that prompt condition was unrelated to the amount of cognitive load experienced by participants.

Hypothesis 3 indicated that cognitive load would mediate the relationship between the metacognitive prompts and the three learning outcomes of self-efficacy, procedural knowledge and skill performance. Bootstrapping was used to test the mediation effect of cognitive load. All indirect effects were non-significant and all 95% confidence intervals included zero; results are displayed in Table 3. Hypothesis 3 was thus not supported; cognitive load did not mediate any of the relationships between the metacognitive intervention variables and the outcomes.

Table 3
Indirect mediation effects of cognitive load

Mediation Effect	Standardized Indirect Effect	SE	p	95% CI
Self-Efficacy				
GG→Cognitive load→Self-efficacy	-.01	.01	.37	(-.04, .02)
SG→Cognitive load→Self-efficacy	.01	.01	.34	(-.01, .04)
SS→Cognitive load→Self-efficacy	.01	.02	.40	(-.02, .04)
Procedural Knowledge				
GG→Cognitive load→Procedural knowledge	-.01	.01	.55	(-.02, .01)
SG→Cognitive load→Procedural knowledge	.01	.01	.56	(-.01, .02)
SS→Cognitive load→Procedural knowledge	.01	.01	.58	(-.01, .02)
Skill Performance				
GG→Cognitive load→Skill performance	-.01	.01	.62	(-.02, .01)
SG→Cognitive load→Skill performance	.01	.01	.64	(-.01, .02)
SS→Cognitive load→Skill performance	.01	.01	.66	(-.01, .02)

Figure 5
Diagram of path analysis results with standardized beta estimates for hypothesized structural model



DISCUSSION

The present study sought to investigate whether scaffolding metacognitive prompts so that learners receive generic prompts early on when learning a new task and then specific prompts later on in the learning process would result in less cognitive load experienced by learners and better learning outcomes. The results of the SEM analyses do not support this hypothesized fully mediated model. There were no differences among the four metacognitive prompt conditions in terms of their effects on cognitive load, and cognitive load was not found to be a mediating mechanism between the metacognitive intervention and the training outcomes of self-efficacy, procedural knowledge and skill performance. Cognitive load was found to be negatively related to self-efficacy, but was not related to procedural knowledge and skill performance.

Metacognitive prompting and cognitive load reduction

One of the main contributions of the proposed study was to investigate whether the sequencing of different metacognitive prompts over time could be used to facilitate learning during a complex training task. Research on metacognition has demonstrated that it is related to various learning outcomes (e.g., Pennequin et al., 2010; Xiangyan & Ji, 2007), although research on the effectiveness of metacognitive prompts in facilitating metacognition and thereby improve learning outcomes has yielded inconsistent results (e.g., Berthold et al., 2007; Caputi et al., 2011). Further, research investigating which types of prompts are most effective has also produced unclear findings (e.g., Berthold et al., 2007; Ifenthaler, 2012). The concept of scaffolding was used in the present study to sequence specific and generic metacognitive prompts across time, creating four metacognitive prompt conditions. It was predicted that the generic-specific condition would produce the lowest cognitive load and the best learning outcomes. Contrary to what was expected, results indicate that none of the four metacognitive intervention conditions significantly differed from each other in terms of their relationship with cognitive load; thus, the metacognitive prompt intervention did not work.

One potential reason for this might be that the metacognitive prompt intervention was not structured effectively. The design of the study emphasized a new training topic at the beginning of each training block, and each training block consisted of three trials with prompts given after the first and second trial; prompts were thus distributed a total of six times across three training blocks. It may be that the two types of prompts given did not provide a strong enough stimulus for encouraging effective metacognition, and participants therefore did not think as deeply about the training task as was expected. Specifically, asking participants to reflect about what they did (generic prompt), what their strengths and weaknesses are and/or what they will do to improve (specific prompt) may not provide enough information or cues as to how they should actually engage in reflection. Prompts may be more effective by directly encouraging connections among learner's knowledge, strategies and progress across trials rather than assuming that participants will be able to do so on their own from information generated about what they just did; an example prompt of this kind might be "How did your performance on this trial relate to the information in the training manual?" This direct approach is consistent with research indicating that individuals are not very good at engaging in metacognition and self-regulation effectively on their own, and that they may need more specific guidance regarding how to do so (Zimmerman, 2000).

It is also possible that the metacognitive prompts may have been viewed as an interruption by participants as they were learning the training task, which could have prevented them from engaging fully with the content and as a result did not reduce cognitive load or increase metacognitive activity; in support of this, some research has shown that prompts can be ineffective when perceived as unnecessary or intrusive (Lin & Lehman, 1999). Similarly, research has also suggested that utilizing prompts during training can actually distract learners from the training task rather than aid them during the learning process (Cannon-Bowers & Bowers, 2009). Taking into consideration learner perceptions about prompts so that they are structured in a way that is face valid may help to mitigate this effect, in addition to explaining the usefulness and benefits of the prompts to learners before they are given prompts.

Another possible explanation for why the prompts did not work is based on one of the main tenets of scaffolding: that learners require assistance during critical early stages of the learning process, but that this support should be removed as learner expertise grows (MacInnes, Santosa, Kronenfeld, McCuaig, & Wright, 2008). Prompts were scaffolded *within* training blocks in the present study, but they may have also needed to have been scaffolded *across* training blocks to be effective. It may be that metacognitive prompts were needed only during the first training block or first two to three training trials because even though learners were instructed to focus on acquiring skills with regard to different topics in each training block, they were engaging with the same training task throughout and thus did not need prompts after gaining sufficient experience with the training simulation. This is supported by research on the expertise reversal effect, which shows that utilizing the same instructional methods with both novices and experts can have a null effect on or may even actually reduce the performance levels of experts as a result of redundant or unnecessary information that increases extraneous cognitive load (van Merrienboer & Sweller, 2005). Thus, the present study may not have effectively taken into account the interactive effects of learner expertise and characteristics of the task that comprise the degree of intrinsic cognitive load experienced.

Another potential factor that may explain the finding that prompt conditions did not differ from each other is that the distinction made between the specific and generic prompts could have been in name only while the content did not actually function differently; in other words, the strength of the manipulation may not have been adequate. Specificity ratings obtained for a list of metacognitive prompts extracted from journal articles, including those used in the present study, did not clearly distinguish the specific and generic prompts targeted as highly rated for their respective traits, suggesting that the prompts themselves may not have been structured correctly. However, the specific and generic prompts utilized in this study were similar in structure to other specific and generic prompts that were developed through extensive pilot work and have been found to be effective in other studies (Davis, 2003). Prompt specificity levels are therefore less likely to be as central of an issue in this study as the prompt scaffolding itself, and it would be worthwhile for future research to examine different scaffolding techniques for specific and generic metacognitive prompts.

Lastly, it is also possible that participants became fatigued over time throughout the course of the study, as the study lasted for two hours and involved a complex, challenging and somewhat repetitive simulation task. Additionally, the prompts themselves were repetitive in nature, with either half or all of the six prompts identical depending on the condition; this may have further contributed to the experience of fatigue and caused participants to stop mentally engaging fully with the prompts over time. Some evidence of this was found in participants' responses to the open-ended prompts; for example, "I did the same thing in that training simulation as I did in the last. However, my brain's becoming tired and my eyes are drooping shut. So it's very possible to not have performed as well as I could", "…this is too much. Too long, boring, repetitive" and "…I'm losing focus because this study is really long, and super repetitive." Given that several participants indicated that the training task was long and repetitive and that they were having a difficult time concentrating throughout the duration, future research in this area could utilize fewer training trials; perhaps two blocks of six total training trials rather than three blocks of nine total training trials.

However, an analysis of the open-ended prompt responses indicates that at least some participants did engage in metacognitive activity such as monitoring and reflecting on their progress, as well as connecting past, current and future events and strategizing. For example, "I correctly knew what types they were but didn't identify class correctly", "I switched and made sure to pay attention to the targets that were moving fastest because they were clearly the most important", and "I was thinking about getting all targets within view and thinking about shortcuts I could use to identify them faster. For example, only looking at one piece of information in each category can correctly tell you the intent and allow you to classify. This saves time." While this may seem to indicate that participants engaged in metacognitive activity in response to the prompts, it is also possible that participants whose responses contained characteristics of metacognition tend to naturally engage in more metacognitive activity than others and thus, they are simply writing down their usual metacognitive thought processes and their responses are not actually due to the prompts. Controlling for individual differences in a trait-like propensity for engaging in metacognitive activities could help to clarify the true effects of the metacognitive prompts.

It is important to note that although the metacognitive conditions did not significantly differ from each other, a comparison of trends in paths from the four conditions to cognitive load suggests that the most effective conditions for reducing cognitive load were the specific-specific condition ($\beta = -0.07$, p = ns) and the specific-generic condition ($\beta = -0.07$, p = ns), followed by the generic-specific condition and the generic-generic condition (β = 0.07, p = ns). This trend is in the opposite direction predicted by my hypotheses and may hint that learners actually benefit most from receiving specific prompts earlier in the learning process, as evidenced by the two conditions with the specific prompts given earlier exhibiting the lowest cognitive load. Coding of the open-ended prompt responses seem to partially support this assertion; responses in the specific-specific condition contained the greatest number of words and unique ideas compared to the other three conditions while the generic-generic condition contained the least number of words.

A potential explanation for this trend is that after learners engage with new aspects of a complex task where they have to consider several pieces of information simultaneously while making quick decisions, they do not then have the cognitive resources to discern what aspects of the task or their learning progress they should reflect on when prompted in a more ambiguous way. Receiving specific prompts initially may help learners to focus their attention more directly in areas that are likely to help them evaluate their progress and identify what to do in future training trials. In support of this, some research has shown that having access to more information and greater explanatory text is beneficial to novice learners early on in the learning process while sparse text and information is more beneficial for more experienced learners (Kalyuga, Chandler, & Sweller 1998; Yeung, Jin, & Sweller, 1998; McNamara, Kintsch, Songer, & Kintsch, 1996).

However, as the metacognitive condition effects were non-significant and the metacognitive intervention was unrelated to cognitive load or any of the learning outcomes, it is unclear as to what strategy for scaffolding prompts is most beneficial to learners throughout the learning process for tasks that produce high cognitive load. Overall, this set of findings is inconsistent with research showing that both specific and generic metacognitive prompts can be beneficial to learners (Moreno, 2009; Wu & Looi, 2012). Additional research that experiments with different scaffolding strategies for specific and generic prompts is needed.

Cognitive load as a mediating mechanism

Another focus of the present study was to directly test cognitive load as a mediating mechanism between the metacognitive intervention and the learning outcomes. Many studies have focused on cognitive load theory as an explanation for research findings that indicate that metacognitive interventions increase metacognitive activity and improve learning outcomes, suggesting that this occurs through a reduction in cognitive load. Despite this, almost no studies could be located that actually measured cognitive load in this context; the present study thus provides a direct test of cognitive load as the mechanism by which metacognitive interventions impact learning outcomes.

Cognitive load was not found to be a mediator between any of the relationships in this study. The pathways from the metacognitive intervention to cognitive load and from the metacognitive intervention to the learning outcomes were all non-significant, indicating that the metacognitive intervention did not have an impact on learners, and thus that there was no relationship for cognitive load to mediate.

One area of question is in regard to the measure of cognitive load that was used, which is a mental workload scale. A review of the item content revealed that only one of the six items on this scale specifically asks about the mental demands of the task whereas the other items generally ask about frustration level and how rushed participants felt, which may mean that some of the aspects of cognitive load (e.g., balancing and remembering information) were not captured effectively by this measure. Consistent with other research that has used a single-item indicator of cognitive load asking about participants' mental exertion (e.g., Fiorella et al., 2012), the hypothesized fully mediated path analysis model was tested again using the single mental demand item from the cognitive load measure utilized in this study. However, while model fit remained excellent, no relationships between cognitive load and the prompt dummy-coded variables or the outcome variables were significant, suggesting that the full cognitive load scale was a more complete measure of this construct.

Despite non-significant findings in this study, cognitive load could still be a likely mediator between a more effective metacognitive intervention and learning outcomes, as cognitive load was significantly related to self-efficacy and displayed relationships that were non-significantly trending in the expected direction with procedural knowledge and skill performance in this study. Cognitive load should be related to the ability of a metacognitive intervention to influence metacognitive activity, and other research provides indirect support for the notion that cognitive load is an important mechanism for understanding the impact of metacognitive interventions. Additional research utilizing a different metacognitive intervention should continue to investigate cognitive load as a mediator of the relationship between metacognitive interventions and learning outcomes while also measuring metacognitive activity to verify that reductions in cognitive load do directly impact the extent to which learners engage in metacognition.

Cognitive load and learning outcomes

Cognitive load can prevent learners from effectively engaging with large amounts of information present during the learning process in complex situations. It was predicted that cognitive load would be negatively related to self-efficacy, procedural knowledge and skill performance, as high levels of cognitive load should interfere with learners' confidence in their ability to do well on the task, the extent to which they understand the task, and their actual ability to perform well on the task. Although cognitive load was not found to be a mediating mechanism in the present study, one of the hypothesized relationships between cognitive load and the three learning outcomes was supported. Cognitive load was negatively and significantly related to self-efficacy, such that individuals who reported experiencing lower cognitive load also reported higher self-efficacy with regards to the training task. This finding is consistent with other research indicating that reducing cognitive load can help learners to experience greater mastery and self-efficacy (Zheng et al., 2009). This suggests that cognitive load directly impacts learner perceptions of their ability to perform well on a task, an important learning outcome.

Although the relationships between cognitive load and procedural knowledge and skill performance were non-significant, they were in the expected direction. First, cognitive load was negatively associated with skill performance ($\beta = -0.05$, p = ns), indicating that although the degree of cognitive load experienced by learners did not significantly impact their actual skill performance on the training task, the directional trend suggests that experiencing greater cognitive load may have reduced skill performance.

There are a few potential reasons why this finding was non-significant, as it is not consistent with other research (Scheiter et al., 2009). It may be that cognitive load is unrelated to skill performance. On the other hand, perhaps all participants including those who performed relatively well on the performance trial still experienced relatively high levels of cognitive load, but were just better able to cope with that load. However, the average cognitive load score was close to the midpoint, suggesting that not all participants did experience high levels of cognitive load. As described by prompt responses, some participants reported feeling fatigued and also indicated that they experienced decreasing performance scores approaching the end of the study.

Because only the performance trial scores were included in analyses, the possibility that performance scores exhibited a curvilinear relationship whereby performance scores increased initially over time as learners became more familiar with the task but then decreased as they became fatigued was not tested.

Second, cognitive load was unrelated to procedural knowledge scores, although this trend was again in the expected direction and suggests that learners experiencing greater cognitive load may have acquired less procedural knowledge (β = -0.06, *p* = ns). To investigate possible explanations for this non-significant finding, which was also inconsistent with other research on cognitive load and metacognitive activity and both declarative and procedural knowledge (Fiorella et al., 2012; Schmidt & Ford, 2003), average item-level and scale-level procedural knowledge statistics were examined. This revealed that the majority of participants answered questions on the knowledge test correctly, and the average overall score was a 7.6 out of a possible score of 12. Additionally, the interquartile range showed the scale midpoint of six to be at the 25th percentile, while the 50th percentile and 75th percentile were scores of eight and nine respectively, indicating that scores were clustered at the high end of the distribution. It could be the case that after engaging in nine training trials and one performance trial, as well as reviewing the training manual multiple times, most participants were able to accumulate similar levels of procedural knowledge that resulted in their ability to score well on the knowledge test. However, a histogram of procedural knowledge scores does show that there was variability among participants, suggesting that this may not be the sole explanation for the non-significant result.

Alternatively, it is possible that the procedural knowledge test did not capture aspects of tacit knowledge that learners who were able to perform the training task well developed during the learning process. In other words, assessing procedural knowledge did not capture critical knowledge components that may have been more variable among participants and also more strongly tied to cognitive load. Other research using the TANDEM task (e.g., Baard, 2014) has utilized a longer form of a procedural knowledge test that incorporates questions involving adaptability and more complex application of knowledge that may better tap into the tacit knowledge that learners engaging with the task may need. Tacit knowledge, which tends to be harder to elucidate, may be more relevant to metacognitive activity operating at a more abstract and higher level of cognition. In support of this, other research on simulation-based training has found that metacognitive prompting enhanced the acquisition of higher-level conceptual knowledge, but not lower level procedural knowledge (Fiorellas & Vogel-Walcutt, 2011). It may thus be more appropriate to capture tacit knowledge as an indicator of the critical knowledge needed to perform well on the training task rather than procedural knowledge.

Limitations and future directions

There were several limitations associated with this study. First, the prompt manipulation may not have been structured effectively to influence cognitive load and learning outcomes. In order to keep the complexity of scaffolding prompts to a manageable level, only two types of prompts were administered in the present study. However, the repetitive nature of the prompts may not have assisted learners sufficiently as they progressed through the training simulation. Although the prompt sequences were designed to align with the introduction of a new training topic per training block, overall learning on the training task was cumulative, which was not incorporated into the scaffolding of the prompts. It is a possibility that learners would have benefited from receiving a greater variety of prompts throughout the training simulation. Future research on metacognitive prompting should experiment with scaffolding varying numbers of different types of prompts to determine the “minimum” amount of prompt variety needed to effectively scaffold metacognitive interventions.

Second, although an effort was made to take temporal elements into account in the design of the metacognitive intervention, the actual effects of the intervention were not examined dynamically over time within the training period as participants learned the training task. Only an overall effect was assessed during the performance trial, which was after the conclusion of the training period, and this effect was not found to be significant. However, it is possible that the prompts may have affected the rates at which participants were able to learn the task during the training period while not affecting their overall performance level at the end of the training, but this temporal effect was not examined. Within the TANDEM paradigm, investigating how participants’ performance scores change from trial-to-trial based on prompt condition and the type of prompt given could provide evidence as to which prompt conditions produce quicker performance improvement rates. More broadly, this could be examined within most training or learning studies by measuring changes over time in relevant performance indicators throughout the learning process.

Third, a few of the materials utilized in the study may not have functioned optimally. Participants indicated that the TANDEM task was tedious, repetitive and caused fatigue when coupled with the two-hour length of the study and the completion of ten trials. This could have obscured some of the effects that were targeted by impacting participants’ learning and performance in later trials towards the end of the study when outcome variables were measured. Designing a shorter training study in TANDEM or utilizing a less repetitive task altogether may help to clarify the results of this study and also increase their generalizability to other contexts.

Additionally, the measure of cognitive load that was used was a mental workload scale, as few studies operationalize this construct. There is some question as to whether the content of this scale is deficient, although it was found to be a more effective measure than a single mental effort item. Future studies seeking to use an efficient and reliable measure of cognitive load should consider using a multiple-item self-report measure similar to the one in this study, but could experiment with adding additional items. For example, it may be useful to include an item that asks participants about their ability to remember information associated with a task, as cognitive load has a direct impact on the amount of information that can be contained within working memory. Alternative measurement strategies such as secondary task performance could also be explored as more objective indicators of cognitive load. In this methodology, a second task usually involving some sort of sustained attention, such as detecting a visual signal, is performed at the same time as the primary task and indicates through metrics like accuracy and response time the amount of cognitive load imposed by the primary task of interest (Paas, Tuovinen, Tabbers, & Van Gerven, 2003).

A somewhat different approach to measuring cognitive load effects is to compute learning efficiency scores that signify the amount of effort that is needed to achieve a particular level of performance (Fiorella et al., 2012). One example of this is based on Paas and Van Merrienboer's (1993) method of calculating learning efficiency, where P represents a standardized performance score and L represents a standardized cognitive load score:

$$E = \frac{(P - L)}{\sqrt{2}}$$

This approach combines performance and mental effort and could help to provide evidence as to why non-significant differences in performance outcomes across prompt conditions could still be meaningful if the mental effort required to achieve those similar levels of performance was different.

Another limitation of this study was that metacognitive activity was not measured directly because many self-report measures of metacognitive activity contained items that were similar to the specific metacognitive prompts and thus could have confounded the relationship between metacognition and the prompts. Open-ended metacognitive prompt responses were collected to determine if the prompt intervention affected metacognition in expected ways based on prompt condition, but this study did not actually utilize prompt responses as an indicator of metacognitive activity that could be incorporated into the hypothesized model. Future research should look into using this strategy as a way of measuring metacognitive activity by coding open-ended prompt responses for characteristics of metacognition such as monitoring and planning behaviors. Additionally, it was assumed in this study that metacognitive prompts would be beneficial to learners, but given the null findings, incorporating a control condition that could be used to determine if metacognitive activity is higher for participants who receive some type of metacognitive prompt compared to participants who do not would be beneficial.

Given the non-significant results of this study, another reasonable suggestion for future research might be to experiment with different scaffolding strategies for metacognitive prompts. A potential exploration of this could be to utilize dimensions other than generic to specific to form the basis for scaffolding prompts. For example, other research has utilized prompts based on the varying phases of metacognition that learners experience by aligning those phases with certain prompt types. In one study, learners received comprehension prompts (e.g., "What is the task about?") during the planning phase and before doing the task, strategy prompts (e.g., "What strategy is appropriated here?") during the action phase of the task and reflection prompts (e.g., “Does this solution make sense?”) during the evaluation phase after that portion of the task ended (Kramarski & Michalsky, 2009).

Reflection prompts given during the evaluation phase were most effective followed by comprehension prompts given during the planning phase, although strategic prompts during the action phase were not effective. Although metacognitive prompts were given to participants in between training trials in this study, there was minimal separation between the action phase of the task and the planning and reflection phases; one possibility is that this increased cognitive load associated with the task because learners did not have time to recover cognitive resources needed for the metacognitive prompts after completing the previous training trial. Future research using TANDEM could experiment with giving participants more time in between trials to engage with the prompts. Some studies do not allocate a particular amount of time for participants to use for engaging with metacognitive prompts, instead leaving this under participants’ control (e.g., Hofmann & Spatariu, 2008). Other studies give learners an overall several-minute-long reflection period (e.g., Ifenthaler, 2012); along those lines, one strategy within TANDEM might be to give prompts at the end of the training blocks to create more defined planning and reflection periods. Additionally, utilizing more than two types of prompts may be needed to effectively target different phases of the learning process.

Another potential scaffolding dimension investigated previously included individual-oriented and task-oriented prompts (Berthold et al., 2007). Individual-oriented prompts targeted the specific learning progress of the learner (e.g., “Which main points haven’t I understood yet?”) while task prompts targeted aspects of the learning task itself (e.g., “How can you best organized the structure of the learning content?”). Results of the study indicated that only task-oriented prompts improved learning outcomes in isolation, while a combination of individual-oriented and task-oriented prompts also improved learning outcomes over the control group (Berthold et al., 2007). Given that the metacognitive prompts included in this study were similar to the individual-oriented prompts, incorporating task-oriented prompts into the scaffolding design may be more effective in facilitating metacognition and reducing cognitive load. For example, in addition to asking learners to reflect on what they could improve or what they're struggling with, which are heavily performance-based, specific prompts could also ask learners about what information must be considered when making decisions or what tradeoffs have to be made within the task.

Another potential strategy for scaffolding is to adapt metacognitive prompts so that they are more responsive to an individual's specific learning progress. For example, different prompts could be employed based on whether an individual performed well on a recent task or performance indicator. It may be that individuals who are consistently struggling to learn even after the initial period of poor performance may benefit most from a more specific prompt that helps them to identify or target useful areas, such as things they are consistently not doing well, where they may need to focus their attention in order to better tackle the overall learning goals. Learners who are struggling may also benefit from hints that provide them with even more direct guidance as to how to approach the task with regard to specific aspects.

Conversely, individuals who are excelling may have greater cognitive resources available and may benefit more from higher level prompts that ask them to make connections between information about, for example, different strategies they've used, what they've learned so far, and how different task aspects are interrelated. Alternatively, prompts could also be selected based on more trait-like individual differences such as the inclination to adopt certain learning strategies, which has been shown to improve learning outcomes on some tasks (Schwonke, 2006). Other individual differences like tolerance for ambiguity could also potentially be used to scaffold prompts.

Implications and conclusion

Despite the finding that the hypothesized mediation model was not supported in this study, there are some potential cautious implications of the results. From a research perspective, cognitive load was found to be significantly related to self-efficacy and exhibited relationships with skill performance and procedural knowledge that were trending in the expected direction, although they were non-significant. This points to the utility of cognitive load as a potential mechanism for understanding learning outcomes, and future research should incorporate and directly measure this construct in studies seeking to understand the relationship between metacognitive interventions and learning outcomes.

From an applied perspective, practitioners tasked with designing training programs for employees or educational programs for students should consider the cognitive load that may be introduced by the content and/or tasks that learners will be engaging with at different stages throughout the program. Based on the findings of this study, structuring training in ways that reduce cognitive load may help to improve the self-efficacy of trainees during training, which may help to facilitate conditions that allow them to engage in metacognitive activity and learn training material, as well as encourage transfer outside of the training context.

For example, lecturing on a large amount of content and then having learners practice the entire process or task associated the content, especially if it’s complex, may cause trainees to experience greater cognitive load as they try to utilize and manage multiple pieces of information while also trying to practice the task in its entirety.

Chunking content and practice into more manageable pieces by lecturing on one facet of the task and then having trainees practice that part, then returning to lecture on another aspect of the task may be a more effective strategy because it allows learners to engage with a smaller amount of information at one time while learning how to do a task.

Overall, the findings of this study do not provide much clarity as to which scaffolding strategies for specific and generic prompts are most effective for learners, although non-significant trends in the data coupled with open-ended prompt responses suggest that contrary to what was expected, specific prompts may be beneficial early on in the learning process. Future research in this area should continue to investigate different scaffolding strategies with regard to specific and generic prompts, as well as other prompt dimensions. Cognitive load and metacognitive activity should be incorporated together into future research as mechanisms for understanding the effects of metacognitive interventions on learning outcomes. Additionally, dynamic learner effects should be measured over time and related to metacognitive prompts in order to help tease apart which particular prompt types should be given for how long and at what points in the learning process.

APPENDICES

Appendix A
Training task materials

TANDEM Practice Objectives

Block 1

<u>Getting familiar with the simulation and making contact decisions:</u>

- Using the mouse and other equipment to operate the simulation. □
- Hooking contacts and accessing the pull down menus. □
- Making TYPE contact decisions. □
- Making CLASS contact decisions. □
- Making INTENT contact decisions. □
- Making FINAL ENGAGEMENT contact decisions. □
- Viewing right button feedback after making contact decisions. □

Block 2

<u>Preventing contacts from entering the penalty circles:</u>

- Using the zoom function to view the "big picture" and monitoring the inner and outer perimeters.
- Using marker contacts to locate the outer defensive perimeter. □
- Watching for pop-up contacts that appear suddenly on your screen.

Block 3

<u>Applying strategies that better prevent contacts from crossing defensive perimeters:</u>

- Prioritizing contacts located on the radar screen to determine high and low priority contacts and the order in which contacts should be prosecuted. □
- Making trade-offs between contacts that are approaching your inner and outer defensive perimeters. □

TANDEM FAMILIARIZATION PERIOD INSTRUCTIONS

During the familiarization period, there will be 22 contacts on the radar screen. When you start this period, you will see a number of these contacts on the screen. Since your focus in on learning basic features of the task and making contact decisions, you should focus on these contacts for now. After the experimenter instructs you to start the scenario and the timer begins to count down, you will focus on hooking contacts, making decisions about the contacts, and viewing feedback about your decisions. Below is a list of steps that you should follow:

1. Hook a contact of your choice.

a. Using the mouse, place the arrow on a contact and click the left mouse button.
b. When the contact is properly hooked, it will turn green and the Hooked Track # in the lower right corner of your radar screen changes to correspond to the contact number.
c. When you gather information from your chip's sensors, that information will be given for the contact you currently have hooked.

2. Make TYPE, CLASS, INTENT sub-decisions for hooked contacts.

d. After the contact is hooked, place the arrow on the TYPE menu button located in the top right of your radar screen. Click on the right mouse button to display the menu options.
e. Move your arrow to the button that says "Speed" and click and hold the right mouse button to view the contact's speed. Use the chart on the next page to see what type of contact is indicated by the speed information. After viewing contact speed, do the same for "Altitude/Depth" and "Communication Time." Once again, use the chart on the next page to see what type of contact is indicated by the cue values. Note that one value you gathered may be inconsistent with the other two; if this is the case choose the option indicated by the majority (2 out of 3) of the values.
f. After viewing the three pieces of information, you are ready to make the TYPE decision. Choose "ID_Air/Sub/Surface" from the bottom of the TYPE menu.
g. A list of choices appears in a menu on the lower right corner of your radar screen. Choose the option that was indicated by the majority of the cues you collected by clicking your right mouse button on the option.
h. Perform steps a through d for the CLASS decision and the INTENT decision.

3. Make FINAL ENGAGEMENT decision.

i. After you have made the TYPE, CLASS, and INTENT decisions for a contact, you can then make the FINAL ENGAGEMENT decision.
j. Move your arrow to the OPER menu and click the right mouse button.
k. Move your arrow to the menu option that says "Engage_Shoot/Clear" and click the right mouse button. A list of choice appears in the lower right corner of your radar screen.
l. If the INTENT of the contact was peaceful you should click your right mouse button on the "clear" option, but if the INTENT of the contact was Hostile you should click your right mouse button on the "shoot" option. See next steps before doing this.
m. When you click your right mouse button on either "clear" or "shoot" you can hold it down to receive information on whether you have engaged the contact correctly.

4. REPEAT

n. After you make the final engagement decision, the contact will disappear and you should repeat steps 1 – 3 for another contact.

o. If you eliminate all contacts in your viewing range, place your arrow on the OPER menu and click the right mouse button. Then place the arrow on the “Zoom-Out” option and click your right mouse button. When you do this, more contacts should appear and you can continue.

Appendix B
Demographics

Please provide the following demographics information below. If you cannot remember your ACT or SAT scores or if you only took one of the two tests, please write a “0”in the corresponding blank. This information will be kept confidential and will only be used for research purposes.

Ethnicity:

1	2	3	4	5
Caucasian	Hispanic	Asian	Black	Other

Gender:

1	2
Female	Male

Years in College:

1	2	3	4	More than 4

College GPA: ________________

ACT score: ________________

SAT score: ________________

Appendix C
Trait goal orientation

Please indicate how much you agree or disagree with each of the following statements.

Learning:

1. I am willing to take on challenges that I can learn a lot from. □
2. I often look for opportunities to develop new skills and knowledge. □
3. I enjoy challenging and difficult activities where I'll learn new skills. □
4. For me, development of my abilities is important enough to take risks. □
5. I prefer to do things that require a high level of ability and talent. □

Performance- Prove:

6. I'm concerned with showing that I can perform better than my peers.
□
7. I try to figure out what it takes to prove my ability to others. □
8. I enjoy it when others are aware of how well I am doing. □
9. I prefer to participate in things where I can prove my ability to others. □

Performance- Avoid:

1. I would avoid taking on a new task if there was a chance that I would appear rather incompetent to others. □
2. Avoiding a show of low ability is more important to me than learning a new skill. □
3. I'm concerned about taking on a task if my performance would reveal that I had low □ability. □
4. I prefer to avoid situations where I might perform poorly. □

Appendix D
Cognitive load

Please indicate your responses to each of the questions listed below using the response scales provided.

1. How mentally demanding was learning about and practicing the training task?

1	2	3	4	5
Not at all demanding	A little demanding	Somewhat demanding	demanding	Very demanding

2. How physically demanding was the pace of the coordination and movement required by the training task?

1	2	3	4	5
Not at all demanding	A little demanding	Somewhat demanding	demanding	Very demanding

3. How hurried or rushed did you feel while learning and practicing the training task?

1	2	3	4	5
Not at all rushed	A little rushed	Somewhat rushed	rushed	Very rushed

4. How successful were you in accomplishing the training objectives?

1	2	3	4	5
Not at all successful	A little successful	Somewhat successful	successful	Very successful

5. How hard did you have to work to accomplish your level of performance in the training task?

1	2	3	4	5
Not at all hard	A little hard	Somewhat hard	hard	Very hard

6. How frustrated were you when you were learning and practicing the training task?

1	2	3	4	5
Not at all frustrated	A little frustrated	Somewhat frustrated	frustrated	Very frustrated

Appendix E
Self-efficacy

Please indicate your level of agreement with the statements below asking about your confidence in your ability to perform on the TANDEM training simulation.

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neither Agree Nor Disagree**	**Agree**	**Strongly Agree**

1. I can meet the challenges of this training simulation.
2. I am certain that I can manage the requirements of this training simulation.
3. I believe I can develop methods to handle changing aspects of this training simulation.
4. I am certain I can cope with the different parts of the training simulation competing for my time and attention.
5. I can accomplish the objectives of this training simulation.
6. I am confident that I can perform well at the training simulation.

Appendix F
Procedural knowledge test

Please complete the following knowledge-based questions related to the TANDEM task by selecting the BEST response.

1. If a contact's characteristics are Communication Time = 20 seconds and Speed = 50 knots, which of the following actions should you take?

    a. Choose Intent is Peaceful □
    b. Choose Type is Surface □
    c. Get another piece of information □
    d. Choose Type is Air □

2. If a contact's characteristics are Intelligence is Private and Maneuvering Pattern is Code Foxtrot, which of the following actions should you take?

    a. Choose Class is Military □
    b. Choose Intent is Peaceful □
    c. Choose Class is Civilian □
    d. Choose Intent is Unknown □

3. If you've just noticed three contacts near your inner perimeter, which of the following should you do next?

    a. Engage the contact nearest the inner perimeter □
    b. Engage the faster contact near the inner perimeter □
    c. Zoom-Out to check the outer perimeter □
    d. Zoom-In to check how close the contacts are to the inner perimeter □

4. If you Zoom-Out to find three contacts around your Outer Perimeter, how would you determine which contact is the marker contact?

    a. Check to see which contact is closest to the outer perimeter □
    b. Check the speeds of the contacts □
    c. Check to see which contact is Civilian □
    d. Check to see which contact is Hostile □

5. What is the purpose of marker contacts?□

    a. To determine which Contacts are Hostile and which are Peaceful
    b. To locate your Inner Defensive Perimeter□
    c. To quickly determine the speeds of contacts near your perimeters
    d. To locate your Outer Defensive Perimeter □

6. Which of the following pieces of information is NOT useful for prioritizing contacts?

    a. The distance of contacts from the Outer Defensive Perimeter □
    b. Whether the contact is Peaceful or Hostile □
    c. The distance of contacts from the Inner Defensive Perimeter □
    d. The Speed of contacts near your Inner and Outer Defensive Perimeter □

7. Which of the following functions is most useful for identifying marker contacts?

    a. Zoom-In □
    b. Right-button feedback □
    c. Engage Shoot or Clear □
    d. Zoom-Out □

8. If three contacts are about 10 miles outside your Outer Defensive Perimeter, which of the following should you do to prioritize the contacts?

    a. Engage the fastest contact □
    b. Engage the hostile contact □
    c. Engage the closest contact □
    d. It makes no difference in what order you engage the contacts □

9. On the average, approximately how many contacts pop-up during each practice trial?

    a. 1 □
    b. 3

c. 6

d. 9 □

10. Which of the following would be the most effective strategy for defending your Outer Defensive Perimeter?

a. Zoom-Out to 128 nm, locate the Marker Contacts, and check the Speed of contacts near the Outer Perimeter □

b. Zoom-Out to 256 nm, locate the Marker Contacts, and check the Speed of contacts near the Outer Perimeter □

c. Zoom-Out to 128 nm, locate a Hostile Air Contact, and check the Speed of contacts near the Outer Perimeter □

d. Zoom-Out to 256 nm, locate a Hostile Air Contact, and check the Speed of contacts near the Outer Perimeter □

11. If all penalty intrusions cost -100 points, which would be the most effective strategy? □

a. Do not allow any contacts to enter your Inner Defensive Perimeter, even if it means allowing contacts to cross your Outer Defensive Perimeter

b. Do not allow any contacts to enter your Outer Defensive Perimeter, even if it means allowing contacts to cross your Inner Defensive Perimeter □

c. Defend both your Inner and Outer Defensive Perimeters □

d. None of these are effective strategies

□

12. It is important to make trade-offs between contacts:

a. That are Hostile and those that are Peaceful □

b. That are near your Inner and Outer Perimeters □

c. That are Civilian and those that are Military □

d. That have already crossed your Inner Defensive Perimeter and those that are □approaching your Outer Defensive Perimeter □

Appendix G
Metacognitive prompt coding instructions

Overall, you need to provide 3 things for an individual's response set. You will be coding **each of their individual responses** for two things: the total number of words and the total number of distinct ideas. You will also provide a rating for an individual's **overall** prompt complexity across **all** of their open-ended responses. Each of these things is explained in greater detail in later sections.

You should follow this general process when coding:
1. Read the individual's first response.
2. Enter the total number of words for that response.
3. Enter the total number of distinct ideas for that response.
4. Repeat this process for the rest of their responses.
5. Provide an overall rating based on all of the individual's responses for complexity.

Coding Practice

When you are ready to practice coding, open up the excel file called "Pilot_open-ended responses". Each individual will have around 4-6 rows of data; in the dataset, you can see that these rows are grouped together; spaces separate participants' responses. You should see four columns in the dataset: participant response, number of words, number of ideas, and complexity rating. You need to provide a rating for the number of words and the number of distinct ideas for **every individual response**, whereas you only need to provide **one overall rating** for complexity. Refer to the file called "EXAMPLE_open-ended responses" for an example of what this looks like.

Actual Coding

When you are ready to begin coding actual data, open up the excel file called "(your name)_open-ended responses" and follow the same process you did when practicing coding.

Coding Information

Total # of words:

Obtain the total number of words for the prompt.

Total # distinct ideas:

When coding the total number of distinct thoughts or ideas, you need to count things that represent **either different subjects or different actions**. A distinct idea could consist of a couple of words or a whole sentence or two. For example:

"Practice hooking and classifying targets": counts as two distinct ideas because hooking and classifying represent two actions

"Zooming out for marker targets": counts as one distinct idea because there is one action referring to one target

"*How to identify the targets and how to zoom in and out so you can protect your lines. I was thinking about code names, the description of the craft and its threat level.*": counts as five distinct ideas ("identify the targets", "zoom in and out so you can protect your lines", "thinking about code names", "description of the craft", "threat level")

"I was prioritizing, which ones were close and more threatening, which ones were not. Try to give the correct response.": two distinct ideas ("prioritizing which ones were close and more threatening", "try to give the correct response")

Overall complexity depth of processing/cognitive complexity/reflection/specificity:

When coding complexity, make a 1-5 rating (using the scale below) based on how well you feel that a participants' responses demonstrated the criteria listed below the rating scale.

1 = Not at all complex
2 = Slightly complex
3 = Somewhat complex
4 = Moderately complex
5 = Very complex

- Expands upon and explains basic ideas (depth & complexity)
- Provides specific details (specificity)
- Makes connections between previous, current and/or future events (evaluation of progress)
- Provides future strategies to try out (planning)
- Provides information about current performance (monitoring)
- Prioritizes tasks

Appendix H
RMSEA and CFI formulas

Root Mean Square Error of Approximation (RMSEA):

$$\frac{\sqrt{(\chi^2 - df)}}{\sqrt{[df(N - 1)]}}$$

Comparative Fit Index (CFI)

$$\frac{d(\text{Null Model}) - d(\text{Proposed Model})}{d(\text{Null Model})}$$

Where d = χ^2 - *df*

REFERENCES

REFERENCES

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102,* 3-27.

American College Testing Program. (1989). *Preliminary technical manual for the Enhanced ACT Assessment.* Iowa City, IA.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103(3),* 411-423.

Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology, 29(3),* 344-370.

Baard, S. K. (2014). An insight into adaptation: Self-regulatory mechanisms as a driver of adaptive performance over time (Unpublished masters thesis). Michigan State University, East Lansing, MI.

Bannert, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research, 35(4),* 359-375.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51,* 6.

Bell, B. S. & Kozlowski, S. W. J. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology, 55*, 267-306.

Bell, B. S., & Kozlowski, S. W. J. (2008). Active learning: Effects of core training design elements on self-regulatory processes, learning and adaptability. *Journal of Applied Psychology, 93(2),* 296-316.

Berk, L. E., & Winsler, A. (1995). *Scaffolding children's learning: Vygotsky and Early Childhood Education.* Washington, DC: National Association for the Education of Young Children.

Berthold, K., Nuckles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17,* 564-577.

Brett, J. F., & VandeWalle, D. (1999). Goal orientation and goal content as predictors of performance in a training program. *Journal of Applied Psychology, 84(6)*, 863-873.

Cannon-Bowers, J. A., & Bowers, C. A. (2009). Synthetic learning environments: On developing a science of simulation, games and virtual worlds for training. In S. W. J. Kozlowski & E. Salas (Eds.), *Learning, Training and Development in Organizations.* New Jersey: Erlbaum.

Caputi, P., Chan, A., & Jayasuriya, R. (2011). How helpful are error management and counterfactual thinking instructions to inexperienced spreadsheet users' training task performance? *British Journal of Educational Technology, 42(4)*, 592-597.

Cera, R., Mancini, M., & Antonietti, A. (2013). Relationships between metacognition, self-efficacy and self-regulation in learning. *Journal of Educational, Cultural and Psychological Studies, 7,* 115-141.

Combs, G. M., & Luthans, F. (2007). Diversity training: Analysis of the impact of self-efficacy. *Human Resource Development Quarterly, 18(1),* 91-120.

Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences, 12(1),* 91-142.

Davis, E. A., & Linn, M. C. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education, 22,* 819-837.

De Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist, 31(2),* 105-113.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95(2)*, 256-273.

Dwyer, D. J., Hall, J. K., Volpe, C., Cannon-Bowers, J. A., & Salas, E. (1992*). A performance assessment task for examining tactical decision making under stress.* Orlando, FL: Naval Training Systems Center.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13(4),* 277-287.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211-245.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Fiorella, L., Vogel-Walcutt, J. J., & Fiore, S. (2012). Differential impact of two types of metacognitive prompting provided during simulation-based training. *Computers in Human Behavior, 28(2),* 696-702.

Fleiss, J. (1986). *The Design and Analysis of Clinical Experiments.* New York: John Wiley & Sons.

Ford, J. K., Smith, E. M., Weissbein, D. A., Gully, S. A., & Salas, E. (1998). Relationships of goal orientation, metacognitive activity, and practice strategies with learning outcomes and transfer. *Journal of Applied Psychology, 83(2),* 218-233.

Furnham, A., Swami, V., Arteche, A., & Chamorro-Premuzic, T. (2008). Cognitive ability, learning approaches and personality correlates of general knowledge. *Educational Psychology, 28(4),* 427-437.

Garner, R., & Alexander, P.A. (1989). Metacognition: Answered and unanswered questions. *Educational Psychologists, 24,* 143-158.

Gully, S. M., Payne, S. C., Kiechel, K. L., Whiteman, J. K. (1999). *Affective reactions and performance outcomes of error-based training.* Paper presented at the 14th Annual Conference of the Society for Industrial and Organizational Psychology, Atlanta, GA.

Hannafin, M., Land, S., & Oliver, K. (1999). Open learning environments: Foundations, methods, and models. In C. Reigeluth, (ed), *Instructional Design Theories and Models,* (Volume III), pp. 115-140. Erlbaum: Mahway, NJ.

Hoffman, B., & Spatariu, A. (2008). The influence of self-efficacy and metacognitive prompting on math problem-solving efficiency. *Contemporary Educational Psychology, 33(4),* 875-893.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

Ifenthaler, D. (2012). Determining the effectiveness of prompts for self-regulated learning in problem-solving scenarios. *Educational Technology & Society, 15(1),* 38-52.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40,* 1-17.

Kauffman, D. F. (2004). Self-regulated learning in web-based environments: Instructional tools designed to facilitate cognitive strategy use, metacognitive processing and motivational beliefs. *Journal of Educational Computing Research, 30,* 139-161.

Keith, N., & Frese, M. (2005). Self-regulation in error management training: Emotion control and metacognition as mediators of performance effects. *Journal of Applied Psychology, 90(4)*, 677-691.

Kenny, D. A. (2011, September 11). Respecification of latent variable models. Retrieved from http://davidakenny.net/cm/respec.htm

Kester, L., Kirschner, P. A., van Merrienboer, J. J. G., and Baumer, A. (2001). Just-in-time information presentation: Improving learning a complex troubleshooting skill. *Contemporary Educational Psychology, 17(4),* 373-391.

Kozlowski, S. W. J., Gully, S. M., Brown, K. G., Salas, E., Smith, E. M., & Nason, E. R. (2001). Effects of training goals and goal orientation traits on multidimensional training outcomes and performance adaptability. *Organizational Behavior and Human Decision Processes, 85(1),* 1-31.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78(2),* 311-328.

Kramarski, B., & Gutman, M. (2006). How can self-regulated learning be supported in mathematical E-learning environments? *Journal of Computer Assisted Learning, 22(1),* 24-33.

Kramarski, B., & Michalsky, T. (2009). Three metacognitive approaches to training pre-service teachers in different learning phases of technological pedagogical content knowledge. *Educational Research and Evaluation, 15(5),* 465-485.

Landine, J., & Stewart, J. (1998). Relationship between metacognition, motivation, locus of control, self-efficacy, and academic achievement. *Canadian Journal of Counseling, 32(3),* 200-212.

Leonard, N., & Insch, G. S. (2005). Tacit knowledge in academia: A proposed model and measurement scale. *The Journal of Psychology, 139(6),* 495-512.

Lin, X., & Lehman, J. D. (1997). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching, 36,* 837-858.

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1(3),* 86-92.

MacInnes, J., Santosa, S., Kronenfeld, N., McCuaig, J., & Wright, W. (2008). nAble adaptive scaffolding agent: Intelligent support for novices. Proceedings from IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology.

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

Meloth, M. S. (1990). Changes in poor readers' knowledge of cognition and the association of knowledge of cognition with regulation of cognition and reading comprehension. *Journal of Educational Psychology, 82,* 792-798.

Miller, T. M., & Geraci, L. (2011). Training metacognition in the classroom: The influence of incentives and feedback on exam predictions. *Metacognition and Learning, 6(3),* 303-314.

Moreno, R. (2009). Learning from animated classroom exemplars: The case for guiding student teachers' observations with metacognitive prompts. *Educational Research and Evaluation: An International Journal on Theory and Practice, 15(5),* 487-501.

Muthen, L. K., & Muthen B. O. (1998-2015). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthen & Muthen.

Muthen, L. K. (2010, October 13). Mplus discussion: Exploratory factor analysis [Msg 6]. Message posted to http://www.statmodel.com/discussion/messages/8/5908.html?1367889553

Nadolski, R. J., Kirschner, P. A., and van Merrienboer, J. J. G. (2005). Optimizing the number of steps in learning tasks for complex skills. *British Journal of Educational Psychology, 75(2),* 223-237.

Najjar, M. (2008). On scaffolding adaptive teaching prompts within virtual labs. *International Journal of Distance Education Technologies, 6(2),* 35-54.

Paas, F. , Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38(1),* 63-71.

Paas, F., & van Merrienboer, J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35(4),* 737-743.

Pennequin, V., Sorel, O., Nanty, I., & Fontaine, R. (2010). Metacognition and low achievement in mathematics: The effect of training in the use of metacognitive skills to solve mathematical word problems. *Thinking & Reasoning, 16(3)*, 198-220.

Pintrich, P., Wolters, C., & Baxter, G. (2000). Assessing metacognition and self-regulated learning. In G. Schraw (Ed.), *Metacognitive assessment.* Lincoln, NE; University of Nebraska Press.

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

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA: Sage Publications, Inc.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118(3),* 219-235.

Renninger, K. A., & Granott, N. (2005). The process of scaffolding in learning and development. *New Ideas in Psychology, 23(3),* 111-114.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66,* 181-221.

Roehler, L., & Cantlon, D. (1997). Scaffolding: A powerful tool in social constructivist classrooms. In K. Hogan & M. Pressley (Eds.), *Scaffolding student learning: Instructional approaches and issues* (pp. 27-37). Cambridge, MA: Brookline.

Saye, J. W., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research and Development, 50(3),* 77-96.

Scheiter, K., Gerjets, P., Vollmann, B., & Catrambone, R. (2009). The impact of learner characteristics on information utilization strategies, cognitive load experienced, and performance in hypermedia learning. *Learning and Instruction, 19(5),* 387-401.

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods, 12(2*), 347-367.

Schmidt, A. M., & Ford, J. K. (2003). Learning within a learner control training environment: The interactive effects of goal orientation and metacognitive instruction on learning outcomes. *Personnel Psychology, 56(2)*, 405-429.

Schwoerer, C. E., May, D. R., Hollensbe, E. C., & Mencl, J. (2005). General and specific self-efficacy in the context of a training intervention to enhance performance expectancy. *Human Resource Development Quarterly, 16(1),* 111-129.

Schwonke, R., Hauser, S., Nuckles, M., & Renkl, A. (2006). Enhancing computer-supported writing of learning protocols by adaptive prompts. *Computers in Human Behavior, 22(1),* 77-92.

Simons, K. D., & Klein, J. D. (2007). The impact of scaffolding and student achievement levels in a problem-based learning environment. *Instructional Science, 35,* 41-72.

Stadtler, M., & Bromme, R. (2008). Effects of the metacognitive computer-tool *met.a.ware* on the web search of laypersons. *Computers in Human Behavior, 24(3),* 716-737.

VandeWalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement, 57,* 995-1015.

Van Gog, T., Kester, L., & Paas, F. (2011). Effects of concurrent monitoring on cognitive load and performance as a function of task complexity. *Applied Cognitive Psychology, 25(4),* 584-587.

van Merrienboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17(2),* 147-177.

Weaver, J. L., Bowers, C. A., Salas, E., & Cannon-Bowers, J. A. (1995). Networked simulations: New paradigms for team performance research. *Behavioral Research Methods, Instruments and Computers, 27(1),* 12-24.

Wenden, A. L. (1998). Metacognitive knowledge and language learning. *Applied Linguistics, 19,* 515-537.

Wesiak, G., Steiner, C. M., Moore, A., Dagger, D., & Power, G. (2014). Iterative augmentation of a medical training simulator: Effects of affective metacognitive scaffolding. *Computers & Education, 76,* 13-29.

Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *The Academy of Management Review, 14(3)*, 361-384.

Wu, L., & Looi, C. K. (2012). Agent prompts: Scaffolding for productive reflection in an intelligent learning environment. *Journal of Educational Technology & Society, 15(1),* 339-353.

Xiangyan, C., & Ji, D. (2007). An experimental study on the effects of metacognition reading strategy training for junior high school students. *Psychological Science (China), 30(5)*, 1099-1103.

Yarmohammadian, A., & Asli-Azad, M. (2012). Effects of metacognition training on the improvement of mathematical function in children with mathematic learning disability. *Advances in Cognitive Science, 14(1)*, 41.

Yeung, A., Jin, P., & Sweller, J. (1998). Cognitive load and learner expertise: Split-attention and redundancy effects in reading with explanatory notes. *Contemporary Educational Psychology, 23,* 1-21.

Zheng, R., McAlack, M., Wilmes, B., Kohler-Evans, P., & Williamson, J. Effects of multimedia on cognitive load, self-efficacy, and multiple rule-based problem solving. *British Journal of Educational Technology, 40(5),* 790-803.

Zimmerman, B. J. (2000). Attaining of self-regulation: A social cognitive perspective. In P. Boekaerts, M. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). Orlando, FL: Academic Press.