Developing Effective Adaptive Training Systems to Enhance Military Instruction

Fleet Davis, Sandro Scielzo, Jennifer M. Riley
SA Technologies, Inc.
Marietta, Georgia
{fleetcdavis, sandro.scielzo, jmageeriley}
@gmail.com

Heather A. Priest
Navy Air Warfare Center Training Systems Division[1]
Orlando, Florida
heather.priest@navy.mil

ABSTRACT

Today's military training environment poses many challenges to instructors, developers, and support personnel. One obstacle is the limited amount of time instructors have to interact with trainees during live training, leading to an increase in simulation in the school house with often high instructor-to-student ratios, dividing an instructor's efforts across many students. To address this issue, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) and SA Technologies, Inc. developed an adaptive training system prototype that provides automated support for instructor-led training through trainee self-guided learning. Using predefined instructor input, the system actively monitors trainees' activities within a simulation and automatically provides targeted feedback and coaching through metacognitive prompts. These prompts mimic the essential input a live instructor would normally provide, allowing instructors to provide consistent, valuable input to all students while lowering their workload. However, despite a wealth of evidence promoting the efficacy of feedback during practice, such system-based interventions are often regarded as intrusive. The purpose of this paper is to provide an overall review of the constraints and considerations associated with developing and implementing such training systems. This includes a summary of our system evaluation, conducted at the United States Military Academy (USMA) at West Point during incoming cadets' normal land navigation training, which included classroom, simulation, and live training exercises. Results demonstrated the efficacy of the system for enhancing training with improved task performance both in the simulation and the subsequent live exercise. Additionally, subjective measures yielded positive evaluations for perceived effectiveness of the training intervention, usability of the system, and subjective workload associated with trainees' interaction with the tool. Together, these findings suggest that the training system and intervention it provides may be a viable approach to enhancing instructor-led, classroom training and provide guidance for the development of future adaptive training tools.

ABOUT THE AUTHORS

Fleet Davis holds an MS in Modeling and Simulation (University of Central Florida) and an MA in Applied-Experimental Psychology (Catholic University of America). He is currently a researcher with SA Technologies, Inc. while also pursuing a PhD in Applied-Experimental Psychology, specializing in human performance and human factors, at the Catholic University of America. Mr. Davis has over 10 years of experience in human factors engineering including conducting cognitive task analyses, designing user interfaces, conducting experimentation, as well as designing, developing, and testing customized training systems. His work is primarily focused on training and assessment techniques aimed at improving cognitive skills in both real-world settings and virtual environments. Mr. Davis has participated in the design and development of several training systems and measurement applications built for a variety of customers including the Army, Navy, Marine Corps, NASA, Federal Law Enforcement Training Center (FLETC), and State Farm Insurance. Some of his most recent work has focused on the development of new and innovative automated methods for real-time performance assessment and cognitive skills training for small, dismounted infantry units conducting training using desktop simulators.

Heather A. Priest is a Senior Research Psychologist at the Naval Warfare Center Training Systems Division (NAWCTSD) after spending over 5 years at the US Army Research Institute (ARI), Technology-Based Training Research Unit in Orlando. She received her PhD in Human Factors Psychology from the University of Central Florida and her MS in Experimental Psychology from Mississippi State University. Dr. Priest's projects have covered applications in military and medical training targeting culture, technology, leadership, teams, intelligent tutoring systems and adaptive training, automated feedback in game-based training, instructional design principles for training technology. Currently, her work at NAWCTSD revolves around Live, Virtual, and Constructive Simulation (LVC). Along with numerous publications, Dr. Priest also served as a co-editor on the Adaptive Training Special Issue of Military Psychology (2012).

Sandro Scielzo has a PhD in Applied Experimental Human Factors Psychology and a MS in Simulation & Modeling from the University of Central Florida. Dr. Scielzo has worked more than 11 years in human-centered paradigms aimed at augmenting training interface systems to support both individual and team performance. Specifically, his area of expertise revolves around the design, training, and measurement issues related to the development of complex interfaces to promote trainees' expertise, decision-making abilities, and operational readiness while considering their individual aptitudes such as spatial and verbal skills. Dr. Scielzo has published numerous publications while working on grants from various DoD agencies, including DHS and military branches, as well as NASA and NSF.

Jennifer M. Riley has a PhD in Industrial Engineering from Mississippi State University, with specialization in Ergonomics and Human Factors. Dr. Riley has 15 years of experience in cognitive engineering research and development, with a focus on situation awareness measurement and training. Her specialties include human factors experimentation, knowledge elicitation, Situation Awareness (SA) and cognitive skills training in virtual reality and games, training needs assessment, human factors analysis, and SA fundamentals training. Dr. Riley has evaluated SA issues in number of domains in order to define potential training interventions for improved performance. She has conducted training research and led the design and development of a number of training programs for the Army (infantry operations and network/signal operations) and Navy (Marine fire support teams), for federal law enforcement first responders, and for NASA mission control operations.

Developing Effective Adaptive Training Systems to Enhance Military Instruction

Fleet Davis, Sandro Scielzo,
Jennifer M. Riley
SA Technologies, Inc.
Marietta, Georgia
{fleetcdavis, sandro.scielzo, jmageeriley}
@gmail.com

Heather A. Priest
Navy Air Warfare Center Training Systems
Division[2]
Orlando, Florida
heather.priest@navy.mil

BACKGROUND

Today's military training environment poses many challenges to instructors, developers, and support personnel. With increasingly complex and dynamic challenges in theatre, training requirements continue to increase while time and opportunities to train diminish. As a solution forward, Army leadership has mapped out a plan to invest in technology and methodology that can lessen that burden, while improving the quality of training. Based largely on guidance from the Army Learning Model (ALM, TRADOC, 2011), research has identified a number of technology-based solutions for the time and resource demands of current and future training needs. Published in 2011, the ALM laid out a plan, based on science and technology, which describes the learning environment envisioned for 2015. The vision, now rapidly approaching the ideal implementation time, "seeks to improve our learning model by leveraging technology without sacrificing standards so we can provide credible, rigorous, and relevant training and education for our force of combat seasoned Soldiers and leaders" (TRADOC, 2011, p. i).

One consideration addressed by the ALM is the directive to "dramatically reduce or eliminate instructor-led slide presentation lectures and begin using a blended learning approach that incorporates virtual and constructive simulations, gaming technology, or other technology-delivered instruction" (TRADOC, 2011, p. 36). Not only does this encourage instructors to move away from the static, boring "death by PowerPoint" approach, but it also helps alleviate another serious challenge for the modern training environment. Specifically, the reduced time to train and strain on resources have led to limitations on the amount of time instructors have to interact with trainees during live training (i.e., problem-centered, experiential learning). This, along with the directive to move away from lecture-based classroom interactions, has led to an increasing need for simulation in the school house and wherever Soldiers train. Ideally, this shift would enable instructors to interact with larger numbers of Soldiers more often as they practice leading to richer, more interactive learning. In reality, this technological push has enabled higher instructor-to-student ratios, allowing more trainees opportunities for instructor led experiential learning, but also dramatically increasing instructor workload and possibly reducing opportunities for learning, interventions, and feedback for trainees due to the instructor being forced to split his attention across more trainees. These unintended consequences are being addressed in several research projects by behavioral scientists working alongside the Army in an effort to provide insight into how, when and where to implement the principles under the ALM.

One possible solution is to use the technology to not only enable experiential learning (e.g., simulation-based practice), but to intervene within the system during game play (much like an instructor would) with feedback, prompts, and interventions. This type of intervention has potential to not only enhance the learning experience of the trainee but also take some of the burden of providing that feedback off of the instructor by allowing adaptive systems to promote self-guided learning. The vision would be to use a blended learning approach to allow an adaptive system to intervene for more well defined skills (self-guided learning), leaving the instructors more time to identify and focus on the harder to diagnose skills and knowledge (instructor guided). However, despite a wealth of evidence promoting the benefit of feedback during practice (Clark, 2003; Merrill, 2002), such system-based interventions are often regarded as intrusive and at risk of "destroying immersion, flow, and engagement in the game" (Kickmeier-Rust, Marte, Linek, Lalonde, & Albert, 2008, p. 11). This pervasive idea of the need for stealth or non-invasive methods of interventions in games has led to concern that the learning part of serious games is not

---

being fully realized. As a result, the purpose of this paper is to review and document empirically the constraints and considerations associated with developing and implementing such adaptive training systems on performance, added cognitive workload, and the perceived benefits of the training tool.

In the current effort, the US Army Research Institute partnered with SA Technologies, Inc. and the US Military Academy (USMA) at West Point to bring together government and industry technology to address these issues at their simulation center. Specifically, ARI, SA Tech, and USMA worked together to develop and test an adaptive training system prototype that provides automated support for instructor-led training through trainee self-guided learning. USMA's simulation center hosts over 1800 incoming Cadets every summer to learn and practice land navigation in a simulated gaming environment created by the US Army's Training Brain Operations Center (TBOC) Systems Integration, Modeling, and Simulation (SIMS) directorate using Virtual Battlespace 2 (VBS2 by Bohemia Interactive Simulations). While this technology created jointly by industry and government is an essential tool in the learning cycle for Cadets, the instructors were looking for ways to enhance the blended learning environment and lessen their workload in the simulation center based on their Cadet-to-instructor ratio, which regularly reaches a ratio of 40:2, while not detracting from the valuable experience offered by the gaming technology. The following study outlines these efforts and the effect of a within game intervention on trainee workload, experience, and learning, as well as identifies the usability of a within game intervention for the learner.

PURPOSE

This study sought to examine the effects of implementing an adaptive training system prototype that provides automated support for instructor-led training through trainee self-guided learning. This training system was designed to automatically deliver a structured metacognitive prompting training intervention while trainees perform a task practiced within a virtual environment. Metacognitive prompting involves delivering messages to trainees, sometimes in the form a query, designed to help trainees regulate their own thinking during task performance and learning. Specifically, metacognition, or simply "knowing about knowing" (Metcalfe & Shimamura, 1994) refers to an individual's awareness of their own knowledge including the ability to understand, control, and manipulate specific cognitive processes (Osman & Hannafin, 1992). Metacognitive skills, particularly as they are applied to self-regulation of learning, impact the acquisition, comprehension, retention and then application of what a trainee learns as well as how efficiently one learns (Hartman, 1998). In addition to their benefits to building knowledge, metacognitive skills allow a trainee to monitor and improve their own learning progress, thereby increasing their effectiveness as learners and the positive impacts of the training itself (Gourgey, 1998).

The goal of this effort was to develop a training intervention that enhances the effectiveness of existing training approaches over and above current outcomes while minimizing the cognitive load the system places on trainees and instructors. Specifically, we focused on developing prompts to support participants' metacognitive skill development in two ways. First, prompts were designed to improve the participants' knowledge of cognition (Schraw, 1998). This included helping the trainees understand themselves as a learner (i.e., declarative knowledge) including what factors could impact their performance (e.g., memory limitations), how to use a particular strategy (i.e., procedural knowledge), and when and why to use a particular strategy (i.e., conditional knowledge). Second, the prompts were created to improve participants' regulation of cognition (Schraw, 1998). This included helping the trainees control their own thought processes and learning including better use of attentional resources, better use of existing strategies, and a greater awareness of comprehension breakdowns.

We decided to explore the potential usefulness of embedding metacognitive prompts within an adaptive system that targets the administration of such prompts depending on, for example, trainee behavior. Therefore, the overarching research question for this study was: What are the impacts on trainees of using an automated metacognitive prompting training system in terms of performance, added cognitive workload, and the perceived benefits of the training tool itself?

METHODS

Participants

Participation was solicited from new cadets completing simulation-based land navigation training as part of a four-day land navigation exercise at USMA. The study was integrated with this exercise as part of the annual cadet summer training event. As this research involved the use of human subjects, it was fully reviewed and approved by the Institutional Review Board (IRB) at SA Technologies, Inc. as well as the IRB committees at ARI and USMA. The sample population (n = 160, 137 male, 23 female) included participants between the ages of 18 and 25 years (M = 18.9, SD = 1.5) who were first year cadets at USMA. All participants had normal or corrected to normal vision. The classroom data was gathered on July 23rd and July 24th, 2013, and the live exercise data was collected on July 25th, 2013.

Research Design

A between-subjects design was used in this experiment and participants were randomly assigned to one of two conditions: (1) the control condition which completed the training task without the aid of metacognitive prompts or (2) the "prompted" condition in which participants received automatically delivered, metacognitive prompts aimed at improving task performance. The task in this study was to complete a virtual land navigation course within the classroom portion of land navigation instruction.

Apparatus

VBS2 Land Navigation Trainer
For the training task, participants in both conditions were asked to complete a virtual land navigation course using the VBS2 Land Navigation Trainer developed by the Training Brain Operations Center (TBOC) Systems Integration, Modeling, and Simulation (SIMS) directorate. This consisted of using standard US Army land navigation techniques and procedures to navigate as efficiently and accurately as possible from a starting point, through a series of navigation points, and back to the starting point to complete the course. The desktop-based Land Navigation Trainer variant presents an accurate replica of a land navigation training area at USMA known as "Bull Pond". The only tools available to the participants for completing the task were a high-fidelity, virtual compass in VBS2 and a paper map of the terrain.

Virtual Environment Situation Awareness Rating System
Participants in the prompted condition received metacognitive prompts while completing the land navigation task via the Virtual Environment Situation Awareness Rating System (VESARS). The VESARS software was designed to monitor for specific entity metrics sent over a network by VBS2. In this study, VESARS was configured to listen for the participant avatar's location within the virtual environment. When an avatar reached a pre-designated location (e.g., an area near a specific set of navigation points), VESARS automatically delivered batches of prompting messages via the VESARS Trainee Interface. When the prompts were triggered to be sent out to a participant, the VESARS Trainee Interface appeared on top of the VBS2 Land Navigation Trainer interface and the task was automatically paused. The VESARS Trainee Interface displayed each prompt separately along with its own unique set of response options (Figure 1).

Figure 1. USMA Cadet responding to a VESARS prompts while navigating in VBS2

Metacognitive Prompts
The focus of the metacognitive prompts was to assess trainees' planning, monitoring, and evaluating cognitive activities. For this study, the adaptive training system prototype (i.e., VESARS) was configured to automatically deliver metacognitive prompts (i.e., messages in the form a query designed to elicit metacognitive processes in

trainees) depending on time spent, or the location of the trainee such as near, at, and after a land navigation point, in between two land navigation points, and around the map perimeter.

The metacognitive prompts were expressly designed to serve as an enhancement to normal instructor-led training in situations where there are high cadet-to-instructor ratios. Thus, the prompts were formatted to mimic the essential input a live instructor would normally provide, allowing instructors to provide consistent, valuable input to all students while lowering their workload. Additionally, the contents of the prompts were specifically tailored to the land navigation task domain and vetted by USMA land navigation instructors before the final list of metacognitive prompts was complete. Following final subject-matter expert approval, the prompts were carefully mapped to specific locations on the virtual terrain in the VBS2 Land Navigation Trainer and phases (e.g., planning, monitoring) of the land navigation task.

Measures

VBS2 Land Navigation Trainer Performance Measures
Participant performance on the land navigation task was automatically captured by the Land Navigation Trainer. Main performance metrics are described in Table 1.

Table 1. Summary of Performance Metrics

Measure	Description	Hypothesis
Accuracy	Accuracy is a composite score expressed in percentage points that takes into consideration the number of land navigation points found and the score associated with each.	The prompted group will be more accurate than the control group
Distance off Trails	Total traveled distance in the virtual environment expressed in meters minus distanced covered on trails. This is an indirect measure of how effective trainees were at navigating based on an azimuth. Following trails is an element of land navigation strategy.	The prompted group will walk less distance off trails than the control group
Time on Course	Time from when the player moves 100 m away from the start point sign until mission end (time in seconds)	No a-priori hypothesis
Time Stationary	Time in seconds trainees were not moving in the virtual environment.	No a-priori hypothesis
Map Used	Tally of how many times the trainees access the in-game map of the Bull pond area. More use of the map is an indirect measure of disorientation.	The prompted group will access the map less often than the control group

Subjective Assessment of the Training Intervention
Participants in the prompted group were asked to provide feedback on the VESARS Trainee Interface used to deliver the metacognitive prompts as well as the prompts themselves in terms of how well they supported the participants' learning of land navigation skills.

Other Measures
All participants were asked to complete a subjective assessment of their workload experienced during the land navigation task. Subjective workload was assessed using the National Aeronautics and Space Administration Task Load Index (NASA-TLX) (Hart and Staveland, 1988). Additionally, all participants were administered a spatial abilities test. This was done to test for any differences that may have existed between the groups, potentially giving one group an advantage on the land navigation task. The participants' spatial ability was tested using the Manikin Test (Benson & Gedye, 1963).

Procedure

Participants entered the room in groups of 20 at a time. The 20 participants were randomly divided into two smaller groups of ten participants. During their performance of the land navigation tasks, the control group received instructor-based feedback and coaching only, as is common practice at USMA, while the prompted group received both the same instructor assistance as well as metacognitive prompts delivered automatically via the VESARS

system. Once the participants were seated at the computer station corresponding to their condition in the study, the experimenters guided them through the experiment.

First, all participants were provided with an introduction to the purpose and format of the study and then asked to read and sign a paper-based informed consent form. All participants were then asked to complete a computer-based Biographical Survey and spatial abilities test (i.e., Manikin Test).

Next, the participants completed the land navigation training task, using the VBS2 Land Navigation Trainer, following normal US Army guidelines and procedures. Performance data from the VBS2 were automatically collected by the simulation software. So as not to prevent the participants from receiving their standard land navigation training as USMA cadets, both groups were allowed to ask basic land navigation related questions from the USMA instructors present during the experimental trials. However, instructor involvement during task performance typically involved brief interactions with the trainee participants. There were no specific limitations on the content or duration of instructor-trainee interactions. Instructors were asked to respond to the participants' inquiries as they would during a normal training session.

Immediately following the land navigation training, all participants were asked to complete a computer-based NASA-TLX and those in the prompted group were asked to complete the VERSARS Trainee Feedback Questionnaire. Finally, all participants were thanked, debriefed, and released.

RESULTS

For the following analyses, an alpha value of .05 was used for significance; however, alpha levels close to .10 were considered indicative of a trend and included in the discussion. Prior to running statistical tests, all data were examined to determine whether or not they met the assumptions of normality. Data that significantly violated these assumptions were not transformed and non-parametric tests were used to account for these violations.

Land Navigation Performance Metrics

The results of the performance metrics collected by the VBS2 Land Navigation Trainer are summarized in Table 2. Overall, while land navigation accuracy, time on course, and map use was not improved for trainees in the prompted group, a significant improvement in performance was achieved for distance off trails and time stationary. Thus, metacognitive prompts either helped or did not affect performance on the land navigation.

Table 2. Performance Metrics Statistics

Measure	Prompted Group	Control Group	Statistic	Significance
Accuracy	$Mdn = 91.07$	$Mdn = 93.33$	$U = 2044; z = -0.86$	$p = .193$*
Distance off Trails	$Mdn = 2659.91$	$Mdn = 3201.71$	$U = 1808; z = -1.89$	$p = .029$*
Time on Course	$Mdn = 2921.26$	$Mdn = 2837.72$	$U = 2126; z = -0.48$	$p = .634$
Time Stationary	$Mdn = 1578.50$	$Mdn = 1392.00$	$U = 1670; z = -2.50$	$p = .012$
Map Used	$Mdn = 3.00$	$Mdn = 3.00$	$U = 2004; z = -1.04$	$p = .149$*

* indicates one-tail results

Furthermore, when correlating the performance measures with accuracy we found a significant inverse relationship between the number of times participants accessed their map and their accuracy, $r_s = -.27, p < .002$ (two-tailed). We also found a positive relationship between time stationary and accuracy, $r_s = -.20, p < .020$ (two-tailed). These significant relationships further support the notion that participants benefited from receiving metacognitive prompts.

Spatial Ability

The results of the Manikin Test suggest that no between-group differences existed between the prompted group and control group participants in terms of their spatial ability. Participant reaction times on the test of spatial abilities

were not significantly different between the prompted group (*M* = 3.03, *SD* = 1.32) and the control group (*M* = 3.03, *SD* = 1.38), *t*(129) = –.03, *p* = .98 (two-tailed). Additionally, participant accuracy scores on the test of spatial abilities were not significantly different between the prompted group (*Mdn* = .91) and the control group (*Mdn* = .87), *U* = 1864.5, *z* = –1.19, *p* = .23 (two-tailed).

Subjective Workload

The results of the NASA-TLX suggest that the additional task of receiving and responding to metacognitive prompts did not lead to the perception of higher workload by the prompted group. To the contrary, the weighted workload scores for the participants in the prompted group (*M* = 50.09, *SD* = 12.21) were lower than scores for the control group (*M* = 54.08, *SD* = 11.54). The difference was not significant, *t*(124) = 1.88, *p* = .06 (two-tailed); however, the results of this test indicated a trend with the prompted group experiencing less workload overall. Furthermore, none of the sub-scales indicated a significant difference.

Trainee Feedback Questionnaire

All 60 participants in the prompted group responded to the Trainee Feedback Questionnaire. Overall, the results show that the metacognitive prompts were perceived by the participants as easy to use, helpful when performing the navigation task, and supportive of their metacognitive processes. The results are summarized by survey section below, and are presented in Figure 2, which illustrates stacked responses for each item (each item is abbreviated in the Figure).

Section 1 – General Reaction and Usability
Item 1, which asked whether "The prompts were helpful and supported my learning of land navigation skills") and item 2, asking if "The VESARS interface was easy to use to respond to the prompts," were designed to elicit general feedback on the metacognitive prompting training intervention as well as the usability of the VESARS system. The overall reaction to the prompts (Figure 2) was fairly positive with 38.33% of the participants either agreeing or strongly agreeing that the prompts supported their land navigation training. In regard to the usability of VESARS (item 2), the majority of the cadets (66.77%) either agreed or strongly agreed that the interface was easy to use. However, there was a potential confound with this question. Several participants asked for clarification on the name of each interface. Thus, some of the participants may have confused VBS2 as part of the VESARS system.

Section 2 – Prompting Utility by Navigation Phase
Item 3, which asked if "Prompts received while planning my route or in the beginning of the land navigation task were useful," item 4, asking if "Prompts received when I reached a navigation point were useful," and item 5, asking whether "Prompts I received periodically throughout the navigation task were useful" were designed to elicit feedback on the metacognitive prompting training intervention's utility for supporting each phase of land navigation (i.e., during planning, throughout the task between navigation points, and upon reaching a navigation point). Overall, the prompts were not viewed as useful in regard to the phase of navigation during which they were received. In particular, prompts received during the planning phase and throughout the navigation task between points were not viewed as helpful. In contrast, the prompts that were perceived as the most useful were received when a participant arrived at a navigation point (item 4). While 40% either agreed or strongly agreed that prompts received when they reached a navigation point were helpful, this distribution appears to be bimodal with 36.67% of participants either disagreeing or strongly disagreeing. The mixed results for these three survey items may reflect the fact that VESARS could not distinguish between participants arriving or leaving a certain location. Therefore, participants could have potentially received planning prompts while in an evaluative state and vice versa.

Section 3 – Prompts Supporting Knowledge of Cognition
Item 6, asking if "The prompts helped me to understand what factors could impact my performance of the land navigation task," item 7, which asked if "The prompts helped me understand how to utilize strategies and procedures for conducting land navigation," and item 8, asking whether "The prompts helped me understand when and why to use strategies and procedures for conducting land navigation" were designed to elicit feedback on the metacognitive prompting training intervention's utility for supporting the participants' knowledge of their own cognition. The results of these three items were positive. On average 42.78% of participants found that the prompts helped their understanding, while only 27.78% found they did not do so. The most compelling support for metacognitive prompting's potential to support knowledge of cognition comes from Item 6. Approximately 50% of the

participants either agreed or strongly agreed that the prompts helped them understand factors that could impact their performance of the navigation task.
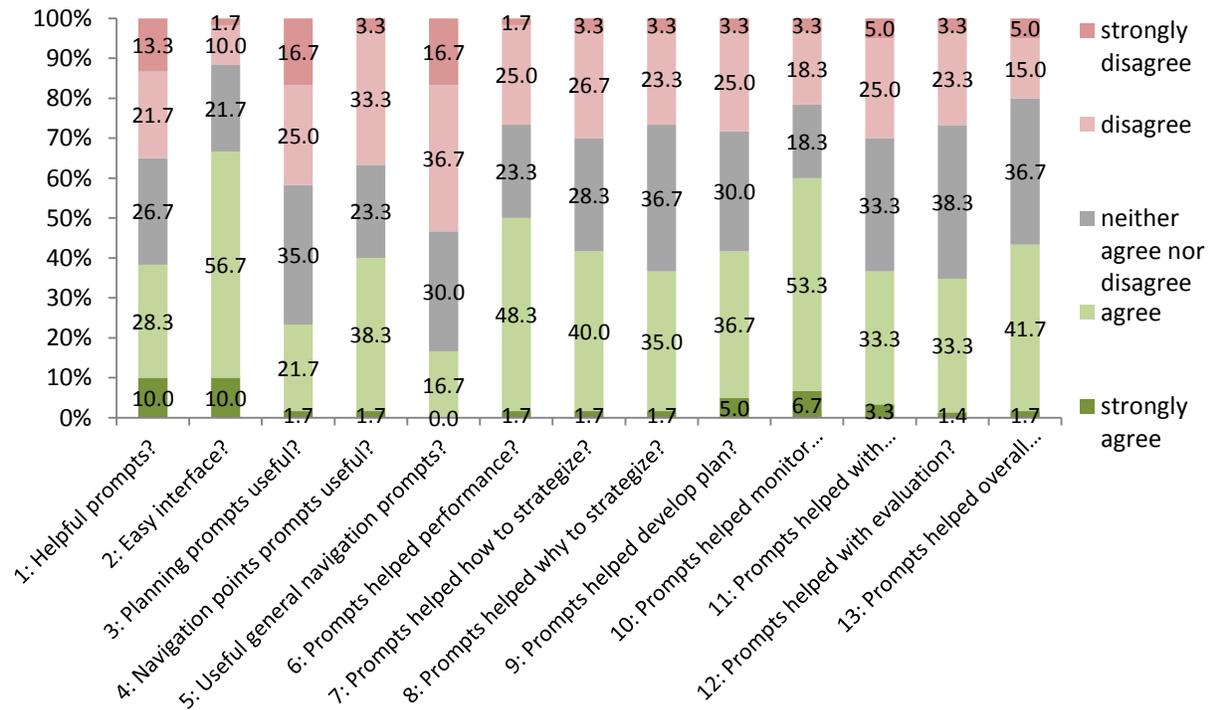

Figure 2. Results of Trainee Feedback Questionnaire showing stacked responses for each item

Section 4 – Prompts Supporting Regulation of Cognition
Item 9, asking if "The prompts helped me develop a plan for conducting the land navigation task," item 10, which asked if "The prompts helped me monitor my own thought processes including what I should be paying attention to during the land navigation task," item 11, asking whether "The prompts helped me monitor breakdowns in my comprehension of the tasks, steps, procedures, etc. used to perform the land navigation task," item 12, asking if "The prompts helped me evaluate and re-evaluate my goals, progress, successes, and failures throughout the land navigation task," and item 13, which asked whether "The prompts helped me understand what I did right or wrong so I could adjust my strategy and plan the next time I navigate" were designed to elicit feedback on the metacognitive prompting training intervention's utility for supporting the participants' regulation of their own cognition. The results of these four items were also positive, with, on average, 54.12% of participants finding that the prompts helped them regulate their own cognition, while only 31.67% finding that they did not do so. The results of Item 10 were particularly interesting. Approximately 60% of the participants either agreed or strongly agreed that the prompts helped them monitor their own thought processes including attention allocation.

The results of these questionnaire items are somewhat in conflict with the results of Items 3, 4, and 5. For example, rating of prompts received throughout the land navigation task (i.e., Item 5) should correspond to ratings of prompts designed to help the participants with monitoring activities (Items 10 and 11). However, Item 5 received a negative rating (i.e., 53.33% of participants either disagreed or strongly disagreed) while Items 10 and 11 each received a positive rating (i.e., at least a third or more of the participants either agreed or strongly agreed). It is unclear what factor is driving this discrepancy. Perhaps the participants viewed the prompts as helpful for understanding and regulating their own cognition, but that engaging in these activities was not useful for their performance of the navigation task.

Transfer of Training

Since the focus of this paper was to present evidence for positively enhancing instructor-led classroom training, this section offers some additional evidence in terms of transfer of training benefits. Performance was tracked for all

cadets on the subsequent live land navigation exercise around the actual Bull Pond terrain. Due to limitations outside of our control (e.g., technical problems with GPS tracking; schedule changes for individual Cadets), data were effectively collected on 131 participants, from an original 160 in the classroom environment. Overall, the land navigation exercise consisted of three trials: two trials involved a buddy system and one Cadets performed individually. For the buddy system trials, participants navigated in pairs, one being the navigator and the other being a shadow (they then switched for the second trial), to find seven land navigation points (for trials 1 and 2 only the navigators' performance was used in our analyses). The third trial also involved identifying seven navigation points. In this section we focus on two measures of performance in the live land navigation task: (1) an aggregate of performance based on number/accuracy of identified points across all trials and (2) an overall time on task aggregate across trials.

The results of the transfer of training performance statistics are summarized in Table 3. Although it was not significant and merely suggests a trend, participants in the prompted group condition demonstrated slightly higher performance than those in the control group. However, a significant result was found for time on task indicating that participants in the prompted condition took significantly less time to complete the land navigation task when compared to participants in the control group. Finally, when correlating the two aggregates (i.e., performance and time across trials) we found a significant inverse relationship between performance and time on task, $r_s = -.44$, $p < .001$ (two-tailed), indicating that participants spending less time navigating were actually more successful. Together, these statistics provide some evidence that the skills acquired via the VESARS adaptive system transferred in a real-world setting.

Table 3. Transfer of Training Performance Statistics

Measure	Prompted Group	Control Group	Statistic	Significance
Performance	*Mdn* = 70.54	*Mdn* = 60.75	*U* = 1733.5; *z* = -1.48	*p* = .069*
Time on Task	*M* = 139.35, *SE* = 1.52	*M* = 143.63, *SE* = 1.36	*F*(1, 128) = 4.37	**p = .019***

* indicates one-tail results

DISCUSSION

The current research investigated the potential for automatic metacognitive prompting to enhance training of a military land navigation task in a virtual environment. The results showed that this type of training can be a valuable enhancement to existing training approaches and, more importantly, do not detract from the learning or add to the workload of the trainees. The following sections summarize the results of the objective and subjective measures collected during the study and conclude with a discussion of implications for future research.

Land Navigation Performance Metrics

Overall, the prompted group did not outperform the control group in terms of land navigation accuracy. While this is somewhat inconsistent with the findings of other similar research, which found metacognitive prompting to be an effective method for improving performance during training (Fiorella & Vogel-Walcutt. 2011; Fiorella et al., 2012; Wu & Looi, 2012), it may also be a function of the practical constraints of collecting data in a setting where the primary goal is learning and not experimental control. For example, our time on task was limited by the curriculum schedule. Therefore, the impact of training and the amount of variability in performance would be limited, indicated by the relatively small number of trainees who even finished the course. However, there were some indications that the intervention was effective, even given the limited exposure. As expected, the prompted group's ability to navigate by azimuth, as measured by distance off trails, was significantly better when compared to the control group. This indicates that the participants in the prompted group benefited from prompts about land navigation strategies (including when or when not to navigate by azimuth). Unfortunately, in part due to training time constraints, time on course showed no differences between the prompted and the control group. However, time stationary did show a significant difference, with trainees in the prompted group spending significantly more time stationary when compared to trainees in the control group. There was also a significant positive correlation between accuracy and time stationary. This would indicate that rather than this difference being related to trainees taking more time to read and answer prompts, trainees in the prompted group may have spent more time planning between navigation points and less time navigating. Thus, although the level of improvement for trainees in the prompted

group was not as strong as anticipated, again perhaps due to the practical time limitations, results depict a consistent image that the adaptive training via VESARS afforded at least equal, if not better, training benefits when compared to the traditional, instructor-led classroom training.

These findings are further reinforced when looking at transfer of training results. Specifically, the day after the land navigation training in the virtual environment, trainees performed an actual real-world land navigation exercise over the same terrain. Differences of performance in terms of accuracy in identifying land navigation points approached significance, with trainees in the prompted group performing better than those in the control group. Furthermore, trainees in the prompted group spent significantly less time navigating when compared to trainees in the control group. These results are further supported by the strong negative correlation found between performance and time on task. Together, these results do support the notion that skills acquired in adaptive training virtual environments do transfer to the real environment.

Subjective Measures and Feedback

The participants in the prompted group experienced a comparable level of cognitive workload to the control group based on standards for statistical significance. While these results only denote a lack of statistical significance (i.e., not that there were 'no differences' at all), there appeared to be a (nonsignificant) trend with the prompted group experiencing less workload overall. While this was only a trend, this finding was consistent with another similar study which found that participants who received prompts reported lower cognitive load after training (Fiorella & Vogel-Walcutt, 2011). Taken together, this is an important point as it supports a major goal of this work to provide evidence that the addition of metacognitive prompting has minimal if any negative impacts on trainee workload.

The results of the Trainee Feedback Questionnaire, while conflicting in some aspects, provided valuable insights into areas for potential improvement with regard to the development, content, and timing of delivery for the prompts. Many of the participants found the VESARS interface to be easy to use, but feedback on the prompts showed that there are many aspects of the prompting that could be improved. This is an important contribution for future development with regards of not just IF interventions should be automated in adaptive systems to promote self-guided learning, but WHEN and HOW.

In regard to navigation phase, several participants indicated in the comments of the Trainee Feedback Questionnaire that they found the prompts to be annoying as the prompts often appeared during the wrong phase of the land navigation task (e.g., planning related prompts delivered during evaluation of the previous movement) or even interrupted task performance (e.g., prompts appearing on the computer screen blocking the participants' view). One major limitation of the current system was that VESARS could not distinguish between arrival at and departure from a navigation point, which led to an occasional mismatch between the prompting content and the participants' activities. Future iterations of these adaptive systems must ensure that trainee activity is adequately tracked and monitored such that distinctions can be made between specific task states used to enhance the timing of prompt.

In spite of the manner in which the prompts were delivered, many of the participants still viewed them as useful overall. Generally, participants reported the prompts to be useful for helping them both better understand and regulate their own cognition. Furthermore, when looking at the written comments and demographics of the participants that either did not view the prompts as useful or were neutral in their response, their answers appear to be influenced by the prompt delivery timing issue discussed previously or their land navigation expertise. Again, these are important issues to inform future development.

Implications for Training in Virtual Environments

The results of the current study provided evidence for the efficacy of metacognitive prompting as a training enhancement when embedded within adaptive training tools such as VESARS. More importantly, we have provided concrete evidence that skills acquired in virtual environments have the potential to transfer to real, live environments. While not definitive, this study lends support that metacognitive prompts do not negatively impact trainees' cognitive load and may potentially alleviate it while being perceived by the trainees as a helpful learning tool, overall. Thus, it is our conclusion that adaptive training systems that provide automated support for instructor-led training not only benefits the trainee, but also benefits the instructors by effectively supporting their instruction.

However, some limitations of this study and approach were identified which should be addressed in future investigations. For example, the results demonstrated that the timing of prompt delivery must be improved. By tracking specific task states, future systems will be able to deliver metacognitive prompts at an appropriate phase of the task such that it helps shape the cognitive activity of the trainee.

In sum, future research should further explore metacognitive prompting as a training approach within adaptive training systems. For example, in the current study the time allotted was most likely too brief for the intervention to take full effect. This type of training may also hit a point of diminishing returns where it is no longer an effective training enhancement. Many important factors remain to be uncovered in order to determine the extent to which metacognitive prompting in virtual environments not only can support core skills acquisition, such as land navigation, but also support more complex and higher-order skills (e.g., leadership and decision-making) including how well they transfer to operational settings. Regardless of the limitations that exist, adaptive training tools, and the training enhancements they can provide clearly have a critical role to play in training in virtual environments.

ACKNOWLEDGEMENTS

REFERENCES

Benson, A. J., & Gedye, J. L. (1963). Logical processes in the resolution of orientational conflict. *RAF Institute of Aviation Medicine, Ministry of Defence (Air), London*.

Carter, R., & Woldstad, J. (1985). Repeated measurements of spatial ability with the manikin test. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *27*(2), 209-219.

Clark, R. E. (2003). What Works in Distance Learning: Instructional Strategies. In H. F. O'Neil (Ed.), What Works In Distance Learning (pp. 13-31). Los Angeles: Center for the Study of Evaluation.

Fiorella, L., & Vogel-Walcutt, J. J. (2011, September). Metacognitive Prompting as a Generalizable Instructional Tool in Simulation-Based Training. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 55, No. 1, pp. 565-569). SAGE Publications.

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.

Gourgey, A. F. (1998). Metacognition in basic skills instruction. *Instructional science*, *26*(1-2), 81-96.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload*, *1*(3), 139-183.

Hartman, H. J. (1998). Metacognition in teaching and learning: An introduction. *Instructional Science*, *26*(1), 1-3.

Kickmeier-Rust, M. D., Marte, B., Linek, S. B., Lalonde, T., & Albert, D. (2008). Learning with computer games: Micro level feedback and interventions. In M. E. Auer (Ed.), Proceedings of the International Conference on Interactive Computer Aided Learning (ICL), September 24-27, 2008, Villach, Austria.

Merrill, M. D. (2002). First principles of instruction. Educational Technology Research and Development, 50(3), 43-59.

Metcalfe, J. E., & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. The MIT Press.

Osman, M. E., & Hannafin, M. J. (1992). Metacognition research and theory: Analysis and implications for instructional design. *Educational Technology Research and Development*, *40*(2), 83-99.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional science*, *26*(1-2), 113-125.

Stanton, N. A., Salmon, P. M., Walker, G. H., Baber, C., & Jenkins, D. P. (2012). *Human factors methods: a practical guide for engineering and design*. Ashgate Publishing.

TRADOC (U.S. Army Training and Doctrine Command) (2011). The U.S. Army Learning Concept for 2015 Fort Monroe, VA: Dept. of the Army Training and Doctrine Command.

Wu, L., & Looi, C. K. (2012). Agent Prompts: Scaffolding for Productive Reflection in an Intelligent Learning Environment. *Educational Technology & Society*, *15*(1), 339-353.