# Effects of a Virtual Pointer on Trainees' Cognitive Load and Communication Efficiency in Surgical Training

**Azin Semsar, MS[1], Hannah McGowan[1], Yuanyuan Feng, MS[1], Hamid R. Zahiri, DO [2], Ivan M. George [2], Timothy Turner, PhD [2], Adrian Park, MD[2], Helena M. Mentis, PhD[1], Andrea Kleinsmith, PhD[1]**
**[1]University of Maryland, Baltimore County, Baltimore, Maryland, US; [2]Anne Arundel Medical Center, Annapolis, Maryland, US**

**Abstract**

*We investigated the cognitive load effect of a trainer providing surgical instruction by pointing/drawing over laparoscopic video to a trainee. Results showed that while cognitive load is higher overall with the use of the instructional system, there is a decrease by the second experience of being instructed by the Virtual Pointer. Further analysis showed that trainees were more likely to perform the surgical task and watch/listen to the trainer's instruction at the same time when the instructional system was used. This is thought to be an indication of more efficient communication when using the instructional system. Thus, although there is a small cognitive overload with the instructional system initially, the more efficient communication allows trainees to better integrate the knowledge and instructions being conveyed into the actions they must perform - indicating a better learning environment.*

**Introduction**

Evidence demonstrates that both cognitive underload and overload leads to decreased performance [1,2] and learning [3,4]. Greater learning outcomes are predicted when training tasks and technologies that reduce cognitive load are employed. The reasoning is simple – more working memory resources will be available for learning [5,6]. Understanding the effects of cognitive load in medical training is especially important given that many of the necessary activities require the learner to integrate different skills, knowledge and behaviors concurrently [7]. Studies have shown that trainees have fewer cognitive resources, i.e., less working memory, available for decision making when comprehension of the instruction is more demanding [8].

To evaluate the effects of instructional technologies on cognitive load, the multimedia dual-processing theory [9] may be a consideration. This theory posits that there are two systems for processing information, one for visual information and another for audio/verbal information [10,11]. Both systems have a finite amount of processing capacity, thus the presentation of two forms of visual information require the visual processing system to be divided in two, leaving the audio processing resources untouched. The claim is that new technologies that can make use of both processing systems allow for more availability of the total cognitive resources [1213]. Studies investigating this theory have demonstrated greater knowledge acquisition is achieved, and cognitive resources are not overloaded when both processing systems are used, i.e., the audio/verbal channel and the visual channel [14,1516].

We employed both subjective and objective measures of cognitive load of surgical trainees while performing tasks and being instructed by a trainer using an instructional technology called The Virtual Pointer. The Virtual Pointer enables a trainer to point or draw a free hand sketch over live laparoscopic video; adding the visual instruction channel to the existing traditional verbal instruction. Adding an additional channel of communication in a dyadic interaction can have positive implications for communication efficiency [17]. Cognitive load can be influenced by communication efficiency in a training interaction [18]. The more efficient the communication, the less cognitive processing is required for understanding the communicated content. Therefore, more cognitive resources are available for learning the content. Our hypothesis was that a Virtual Pointer can enhance communication efficiency, and consequently allow for further cognitive processing.

**Related Work**

*Cognitive Load And Information Processing Channels*

Cognitive load theory [19,20,21,22] describes working memory as having a limited capacity in processing information. When the number of informational units exceed the limited capacity of working memory, new information cannot be processed and thus learning will be hindered. In a learning environment, cognitive load may arise from the instructional tools and techniques rather than the complexity of the task itself. Learning, as an outcome of a training process, will be hindered if instructional materials and methods overwhelm a trainee's cognitive resources [20]. In multimedia learning (i.e., learning from words and pictures), cognitive load arising from instructional techniques is a central challenge faced by designers of multimedia instructions [15]. Mayer and Moreno (2003) suggest nine ways to reduce cognitive load in multimedia learning including a cognitive process involved in integrating verbal and visual cues. One other recommendation they make is to off-load one information receiving channel (e.g. verbal and visual) by splitting the information among multiple channels. This solution reduces the processing demand of one channel so the learner has further capacity for processing within other channels. They demonstrated that students understood a multimedia explanation better when the words were narrated rather than shown as an on-screen text.

Brunken et al.'s [14] study confirms the guideline of splitting the information among multiple channels using a dual-task methodology. In their study, participants were asked to perform two tasks at the same time: the primary task involved learning from a multimedia program on how the human cardiovascular system works, and the secondary task involved needing to press a button as soon as a letter color changed. Analyzing the reaction time for the secondary task as a cognitive load measure, they found that reaction times were lower for the audiovisual presentation compared to the visual only presentation. This showed that a combination of audio and visual presentation induced less cognitive load than a single modality presentation, i.e visual-only. The dual-task methodology is based on the assumption that cognitive capacity is limited, but can be flexibly allocated [14]. In the case of processing two tasks at the same time, the cognitive resources have to be split between the two. Therefore, when a person is able to better perform a second task simultaneously with the first task, the less cognitively loaded the person is on either task, thus allowing for better integration of cognitive processing.

To conclude, the way instructional techniques and tools are constructed may affect cognitive load differently. While it is suggested that designing an instructional method in a way that splits the information between auditory and visual channels reduces the cognitive load, a Virtual Pointer has not been studied in terms of how it affects the cognitive load even though it uses the same principle of splitting information between two channels. Because it has been reported that multiple modalities of processing channels allows for more cognitive processing space, we are interested in examining if the Virtual Pointer may affect cognitive load for this reason.

*Cognitive Load and Communication*

Analyzing cognitive load in communication settings is additionally important to allow for effective group collaboration. For example in Computer Support for Collaborative Learning (CSCL), it has been discussed that the interaction among subjects in a learning environment may generate communication activities (e.g. explanation, disagreement, and mutual regulation), which may trigger extra cognitive mechanisms (e.g. knowledge elicitation, and internalisation) [23]. The primary reason for the importance of cognitive load in CSCL environments is that CSCL environments force learners to coordinate one or more external, instructional representations [24]. In such environments, the need for learners to integrate the textual, verbal, and visual information sources presented to them creates an additional cognitive load on top of the task itself [25]. The balance between the reduction of individual computing due to division of labor and the increase of individual computing necessary for interaction is the key to avoid detrimental cognitive overload in CSCL environments [23].

Effective instructional methods can enhance the communication efficiency between the learner and the trainer. Once the communication becomes more efficient, the cognitive resources will be directed toward activities that are relevant to learning [18]. Communication is a joint activity, which is coordinated based on common ground – shared knowledge, beliefs, and suppositions [26]. And in turn, common ground is incrementally built on the previous joint activities within a group [27]. Thus, as common ground accumulates, communication becomes more efficient and efficient communication reduces the costs for the development of common ground [28].

In conversation, one major coordination task is turn taking, i.e., all people conversing coordinate the time of entry and exit in the conversation [26]. The changes in the coordination of turn taking on one hand indicates the efficiency of communication [29], and on the other, relates to the grounding costs, such as the costs in language processes, i.e., the construal of meaning, and the costs in signaling and accepting [30]. For example, more turns, fewer words and more synchronicity manifest in teams with an increased amount of shared understanding, and thus more efficient communication [29,27]. The increased common ground and improved communication efficiency facilitates the construal of what would be presented, manifested by more quick and short turns [26,31]. For instance, Fussell et al. [32] studied the effects of shared visual context in a collaborative repair task on communication efficiency through turn taking analysis. The greater number of turns in the audio-video instruction compared to audio-only instruction shows that audio-video instruction enhances the communication efficiency between the trainee and trainer in performing a repair task. Because of this, we decided to not only investigate cognitive load, but communication efficiency elicited by the Virtual Pointing system because of its integrated audio-visual instruction basis.

*Cognitive Load in Surgical Training*

An investigation of the effect of haptic feedback on surgical residents' cognitive load when performing a laparoscopic task in a virtual reality setting while they were imposed by a secondary mental arithmetic task showed that residents performed 36% faster with haptics compared to without.[33] This suggests that the addition of haptic technologies allowed for a reduction in cognitive load and thus increased performance efficiency. In comparison, Andersen, et al. [34] found that adding technology to a surgical training condition neither increased or decreased cognitive load nor did it affect performance. Here, they compared trainees' cognitive load with and without a simulator-integrated tutor function in a virtual reality surgical setting by analyzing the reaction time of secondary observational task. They found the integrated tutoring did not influence reaction times and did not have an effect on cognitive load. They did, however, observe that novices gain proficiency in the Virtual Reality surgical simulation after relatively few practice sessions. This effect of experience, i.e., increase in proficiency after additional exposure to using a new technology, has been observed in other studies and could play a critical role when examining a technology's effect on cognitive load. For example, Theodoraki et al.[35] assessed surgeons' cognitive load of using an image-guided navigation system - a system that combines the information captured from cameras, ultrasonic, electromagnetic sensors and relay the patient's body view and the surgeon's movements in relation to the patient, to the surgeon's screen. Non-significant difference in heart rate and heart rate variability between the two conditions showed that cognitive load did not differ between use with and without the navigation system. Nevertheless, the heart rate variability was slightly higher in the navigation-supported condition. Moreover, they observed that residents who had more practice in performing the procedures showed a slight decrease in mental workload while using the navigation system. This suggests that with more experience, there may be more cognitive processing space open to allow for the addition of new technology to be beneficial.

## Material and Method

*System Design and Setup*

The Park Trainer (Stryker Corporation, USA) was used for the simulated laparoscopic tasks (Figure 1.c). It consists of a housing unit for physical anatomical models, a flexible shield with openings for the laparoscopic camera and instruments to be inserted, a standard laparoscopic camera with light source using the Stryker computer system, and a standard laparoscopic monitor on an adjustable arm at the top. The Virtual Pointer was designed to facilitate the conveyance of knowledge during surgery by enabling attending surgeons to point or draw on the laparoscopic video for a surgical resident to see (Figure 1). To this end, the Microsoft Kinect sensor version 2 (Microsoft Corporation, USA) was used as a mechanism of touchless interaction – enabling the system to be used in the sterile operating field [36]. Refer to Feng et al. [37] for a detailed description of the system and how it works.

*Experiment Design and Procedure*

The experimental design is a counterbalanced, within-subject design, with two mentoring approaches: the control is Standard condition, and the intervention is Virtual Pointer condition. In the Standard condition, trainer instruction was conducted as it would be normally, through verbal or hand gestures. In the Virtual Pointer condition, the Virtual Pointer application was used by the trainers as an addition to standard guidance to facilitate instruction.

The trainees worked on four simulated laparoscopic tasks under trainer guidance. The tasks were selected based on a hierarchical task analysis of the laparoscopic cholecystectomy procedure and confirmed by an attending surgeon that they were of similar difficulty levels and required both skills of anatomical structure identification and instrument manipulation. The tasks were performed on a validated laparoscopic training physical model [38], including (1) mobilizing the cystic duct and the cystic artery, (2) clipping the cystic duct, (3) clipping the cystic artery, and (4) cutting the cystic artery and the cystic duct. Task order and condition were counterbalanced for each trainee yielding a total of 14 runs in the Virtual Pointer condition and 14 runs performed in the Standard condition.

The study was approved by the University of Maryland, Baltimore County institutional review board (IRB) and informed consent was obtained from all participants before their participation. After consent, the trainees and trainers completed a demographics questionnaire, which included information on their surgical experience and familiarity with the Kinect system. After each task, the trainees and trainers completed a cognitive load questionnaire. The study was video recorded and the operative field was screen recorded.

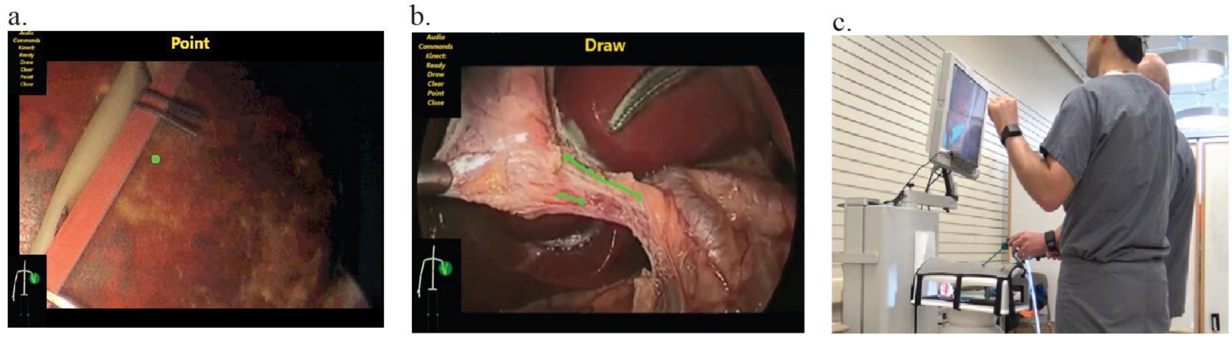


**Figure 1.** (a) Virtual Pointer user interface in pointing mode with list of verbal command (top left), current mode (top center), gesture recognition feedback window (bottom left), pointer (green dot on anatomy); (b) drawing mode (green lines trailing green dot); (c) Trainer using the Virtual Pointer - closed hand is used for drawing.

*Participants*

Participants were recruited from the Department of General Surgery, Anne Arundel Medical Center, Annapolis, Maryland - 7 surgical trainees (all male), including 1 surgical fellow, 1 research fellow, and 4 surgical residents (2 PGY-1 and 2 PGY-2) were recruited. One attending surgeon and one surgical fellow were recruited as the trainers. The attending surgeon guided the surgical fellow in performing the tasks and the surgical fellow guided the rest of the trainees. None of the participants had any previous interactions with the Virtual Pointer before the experiment.

*Data Collection*

*Subjective Measure of Cognitive Load.* Subjective measures are considered an unobtrusive, and inexpensive method of assessing workload [39] and are commonly employed to assess surgeons' cognitive load in training scenarios [40,41]. Commonly employed cognitive load questionnaires for training are the Paas scale [42], a short, single 9-point Likert scale; and the NASA-TLX [43], a considerably longer, multidimensional scale. Given that the NASA-TLX requires more time to complete and the fact that it aims to assess overall workload, i.e., not only cognitive load, we chose the Paas scale for our study. The trainees completed it with reference to the mental effort they invested in understanding the instructions, from 1 – very, very low mental effort was invested; to 9 – very, very high mental effort was invested. The trainee completed the questionnaire after performing each task. The Paas scale was also used by the trainers to examine their cognitive load with reference to the mental effort they invested in giving the instructions.

*Objective Measure of Cognitive Load.* We also included an objective, continuous physiological measure of cognitive load – EDA (electrodermal activity). This continuous measure allows us to assess temporal, dynamic aspects of cognitive load. Physiological measures have a high degree of sensitivity; allowing them to be used for measuring changes in cognitive load [44]. The Empatica E4 wristband was used to record the trainees' EDA used as the objective measure of cognitive load. The Empatica E4 wristband (Empatica Inc., US) was worn on the trainees' non-dominant wrist. According to the E4 user manual, the nondominant wrist is recommended to minimize motion artifacts. At the start and end of each task, an event was tagged on the wristband through its physical button. These tagged events were to be used in the data analysis process for time synchronization. The data was recorded directly onto the E4's flash memory, i.e., wirelessly. After recording each trainee, the E4 was connected via USB to a PC, and Empatica Manager was used to transfer the data to Empatica's secure cloud platform, Empatica Connect. Each trainee's data was then downloaded and renamed with her/his anonymized participant ID. The start and end time of each task was tagged directly on the data using the physical button on the wristband.

*Objective Measure of Trainee's Response to Trainer's Instruction.* Video recordings were annotated to determine how often the trainee stopped performing the task to listen to instruction (i.e., single-task instances) and how often they continued performing the task while also listening to instruction (i.e., dual-task instances). All non-study-related events that occurred between tasks, such as waiting for equipment or fixing of the model were not considered in the analysis. In both mentoring conditions, only the trainees' responses for each instruction given by the trainer were considered. In the Virtual Pointer condition, instructions using the Virtual Pointer, such as moving it to indicate a structure or using it to facilitate verbal instruction were annotated. In the Standard condition, verbal instructions such as "over here" and "move up higher" were also annotated.

*Objective Measure of Number of Turns and Turns Length.* To examine communication efficiency between the trainer and trainee, we analyzed the turn taking structure of communication. As communication becomes more efficient, turn-taking increases and turns become shorter [27]. The video recordings were annotated to count the number of turns the trainee and trainer took throughout each task. Both verbalizing and taking an action (e.g. the trainer giving an instruction and the trainee performing a laparoscopic action) were counted as turns. The length (i.e. amount of time) of each turn was also measured. As the number of turns throughout each task is associated with the length of the task, the fraction of number of turns to the sum of all turns' length was computed as a normalized measure for each task. The analysis was performed on the normalized number of turns and length of turns.

*Data Analysis*

*EDA Data Preprocessing*. Observer XT version 14.1 (Noldus, Netherland) was used to synchronize the EDA data with the video recordings. Each trainee's EDA data was normalized to a range between 0 and 1 as is standard practice [45]. For each specific comparison, normalization was performed across tasks which enabled us to compare subsequent experiences of each mentoring condition for each trainee. We focused our analysis of the EDA data around the points of instruction provided by the trainer in both mentoring conditions. We identified the times that the trainer used the pointing and drawing mode to show a location on the laparoscopic display and the corresponding points of instruction in the Standard condition. A verbal instruction in the Standard condition was considered corresponding to a Virtual Pointer instruction if it was given to show the trainee a location (e.g. "right in the middle"). The instructions were considered as stimuli in the EDA signal. The EDA signal has two components: the slowly changing skin conductance level (SCL) component, and the rapidly changing skin conductance response (SCR) component. The SCR occurs as a peak in the signal, generally as a response to stimulus between 1 to 4 seconds after stimulus presentation [46]. Therefore, we extracted the EDA data for the first 5 seconds after the start of each instruction, Peak amplitude of the SCR has been used as an indicator of cognitive load [47]. The EDA signal was decomposed into SCL and SCR by performing continuous decomposition analysis using the Matlab-based software, Ledalab [48]. The amplitudes of the SCRs for the five-second intervals were computed and used in the statistical analysis.

*Statistical Analysis.* Statistical analysis was performed using a linear mixed model to compare the trainees' cognitive load between the two conditions for both subjective and objective measures. Because our intention was to focus on the effect of the Virtual Pointer on cognitive load over time, we modeled the mentoring conditions (Virtual Pointer or Standard) and the task order as fixed factors. Due to the similar level of difficulty of the tasks, we considered the task as a random factor. The trainees were also considered a random factor. All statistical analysis was performed using R version 3.2.0 (R foundation for Statistical Computing, Austria). The results are shown as mean and standard error of the mean. A p-value of less than 0.05 was considered statistically significant.

## Results

*Trainees' subjective cognitive load scores*

The trainees' average self-reported cognitive load scores were a bit higher in the Virtual Pointer condition (M=4.28, SD=1.72) compared to the Standard condition (M=3.78, SD=1.67) indicating that the trainees did not experience a significant increase of cognitive load when instructed using the Virtual Pointer (p=0.224) (Figure 2.a). We then investigated whether the trainees' self-reported cognitive load scores decreased over time, meaning the second experience of using the Virtual Pointer perceived as less cognitively demanding in comparison to a second experience with the Standard method. We analyzed the differences in cognitive load scores between the first and second experience of each condition (Figure 2.b). Although the average cognitive load scores decreased by the second experience of each condition, the decrease within each condition was not significant. However, because the scores did decrease, we also analyzed the decrease between the second experience of each condition. Somewhat surprisingly, these results were not significant either; suggesting that the trainees' perception of cognitive demand while using a Virtual Pointer is not significantly different from the Standard instruction.

Given that the trainers' cognitive load is also likely to be impacted by using the Virtual Pointer as an instruction mechanism, it was important to assess whether different levels of cognitive effort are required to provide the instruction. We compared the trainers' average self-reported cognitive load scores across all tasks using the Virtual Pointer with all tasks in the Standard condition (refer to Figure 2.c). The results indicate that the trainers' cognitive load actually decreased significantly with the Virtual Pointer (M=3.42, SD=1.45) in comparison to the Standard condition (M=5.64, SD=2.06) (p<0.001). This result suggests that the Virtual Pointer is more beneficial to use as an instructional aid than the traditional instruction method.

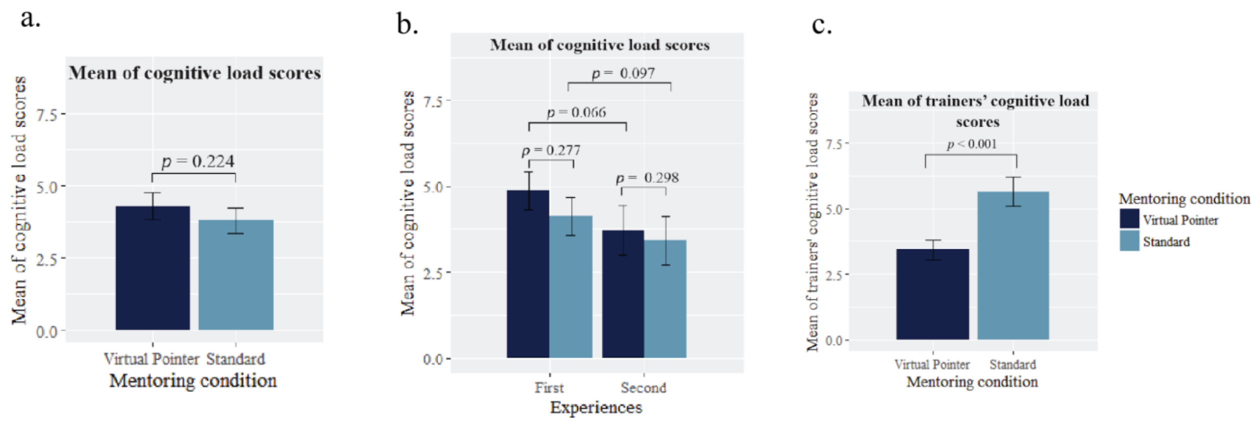


**Figure 2.** (a) Overall comparison of cognitive load scores of trainees between the Virtual Pointer and Standard conditions. (b) Comparison of cognitive load scores between the first and second experiences of each mentoring condition. (c) Comparison of trainers' cognitive load score between the Virtual Pointer and Standard conditions.

*EDA Data and SCR Amplitudes*

We also examined the trainees' EDA as it is an involuntary, continuous objective measure of cognitive load. Statistical analysis was conducted using the SCR amplitudes for the first 5 seconds after the start of an instruction with the Virtual Pointer and equal instructions in the Standard condition. A total of 62 Virtual Pointer instructions and 35 equal instructions in the Standard condition were identified. Overall, no significant difference was found in SCR amplitudes between the Virtual Pointer (M=0.031, SD=0.057) and the Standard condition (M=0.025, SD=0.033) (p=0.576) (Figure 3.a). This result confirms the finding of the subjective cognitive load measure.

To have a better understanding of how additional familiarity with the system may impact cognitive load, we also looked at the SCR amplitudes between the first and second experience of each condition as illustrated in Figure 3.b For the Standard condition, similar to the subjective results, the decrease in SCR amplitude was not significant (first experience: M=0.030, SD=0.059; second experience: M=0.026, SD=0.021) (p=0.606). However, for the Virtual Pointer condition, there was a slightly significant decrease in SCR amplitudes from the first experience (M=0.037, SD=0.071) to the second experience (M=0.024, SD=0.034) (p=0.047). These results suggest that although there is an initial slight cognitive load on the trainees with the first experience with the Virtual Pointer system, this load appears to diminish by the second experience as the trainees become accustomed to the system. In fact, after the second experience of instruction with the Virtual Pointer, cognitive load in the Virtual Pointer condition is not significantly different from the traditional mentoring method of instruction.

*Single- and Dual-task Instances*

To investigate whether cognitive load may be attributed to how the trainees attended to the tasks at a fine-grained, instruction-by-instruction level, the occurrence of single- and dual-task instances was analyzed for each condition. Overall the percentage of single-task instances was higher than dual-task instances in both the Virtual Pointer condition (60.65%) and Standard condition (77.8%). To compare the two mentoring conditions, the percentage of single- and dual-task instances in each of the four laparoscopic tasks was normalized to a range between 0 and 1. Overall, there was no significant difference in the percentage of dual-task instances between the two mentoring conditions (p=0.146) even though the mean was higher in the Virtual Pointer condition (Virtual Pointer M=0.392, SD=0.152; Standard M=0.238, SD=0.373). To further examine the dual-task instances, the percentage of dual-task instances within and between the first and second experience of both conditions was analyzed (refer to Figure 4.a). The results show not only that the increase in dual-task instances within the Virtual Pointer condition was significantly higher in the second experience (p=0.01), but also that dual-task instances completely disappeared in the second experience of the Standard condition (p=0.004). Thus, the differences between the second experience of each condition are also significant (p<0.001) (Virtual Pointer: M=0.476, SD=0.115; Standard M=0, SD=0). These overwhelming differences in dual-task instances in the second experience of the Virtual Pointer suggest that as the trainees became more accustomed to the system, they were able to perform more of the laparoscopic task while watching and listening to the trainer's instructions without negatively impacting their cognitive load.

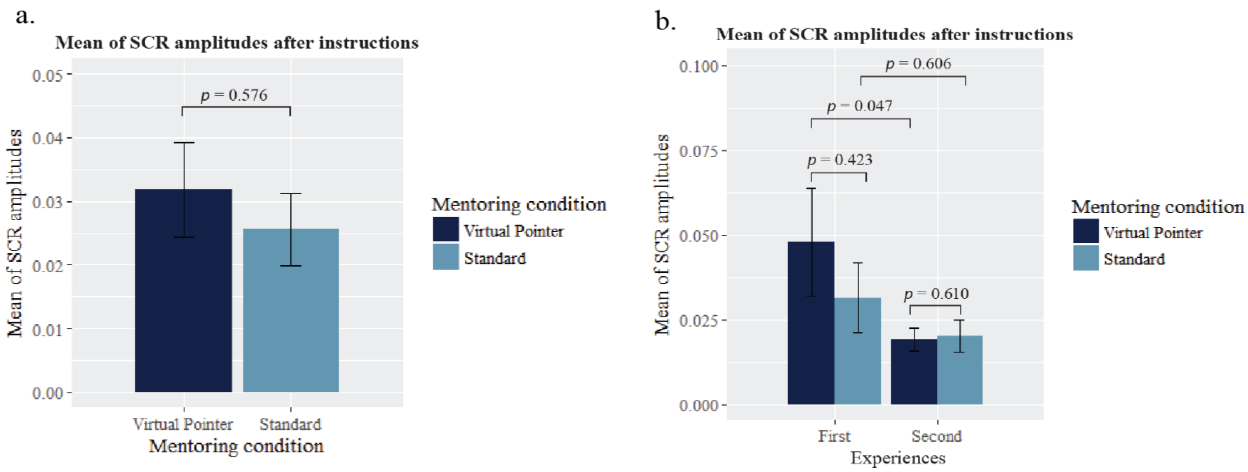


**Figure 3.** (a) Comparison of the SCR amplitudes between the Virtual Pointer and Standard condition. (b) Comparison of the SCR amplitudes between the first and second experiences of each mentoring condition.

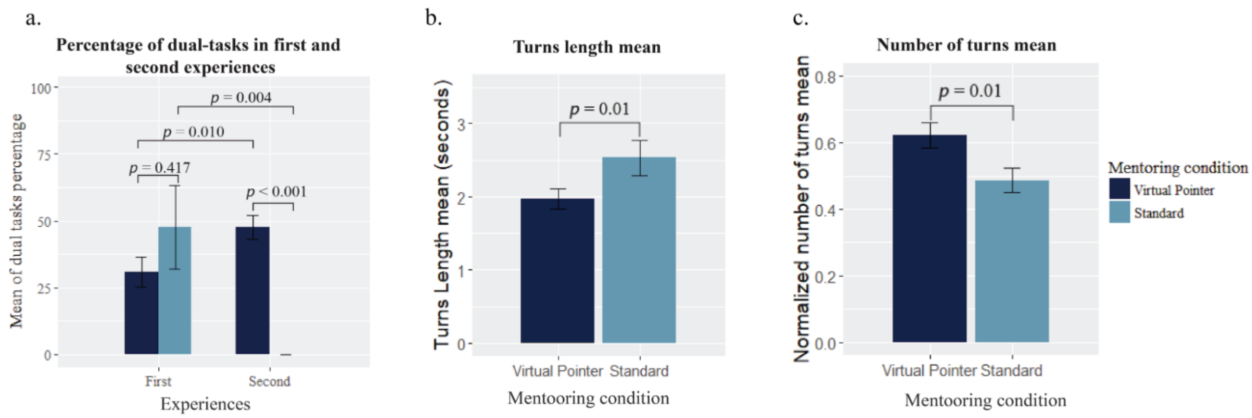


**Figure 4.** (a) Comparison of dual-task percentage between the 1st & 2nd experiences of conditions. (b) Comparison of the turns length mean between conditions. (c) Comparison of the number of turns mean between conditions.

*Turn Taking Results - Number of Turns and Turns Length Mean*

To investigate whether the higher number of instances of dual-tasks while experiencing the same level of cognitive load in the Virtual Pointer condition is due to more efficient communication, the turn-taking structure, specifically the normalized number of turns (Figure 4.c) and turns length (Figure 4.b), was analyzed. Overall, there was a significant difference (p=0.01) in the turns' length between the Virtual Pointer condition (M=1.975, SD=0.507) and Standard condition (M=2.535, SD=0.889). In addition, significantly higher normalized number of turns (p=0.01) were found in the Virtual Pointer condition (M=0.623, SD=0.142) compared to the Standard condition (M=0.487, SD=0.138). The shorter turns' lengths and higher number of turns in the Virtual Pointer condition supports our assumption that the Virtual Pointer improves the communication efficiency, thus allowing for further cognitive processing. Consequently, trainees are able to perform dual-tasks while experiencing no additive cognitive load.

**Discussion and Conclusion**

The goal of this study was to evaluate a Virtual Pointer system in terms of how the added visual information affects trainee's cognitive load, i.e. the effort needed for trainees to process instruction and task related information. Our hypothesis was that although the Virtual Pointer would add more visual information, it could improve communication efficiency between trainer and trainee, thereby allowing more space for cognitive processing.

Overall, we found no additional cognitive load was associated with the Virtual Pointer after trainees became accustomed to the system despite the indication of prior work [22]. There was a slight non-significant increase to the trainees' cognitive load in the first experience of the Virtual Pointer condition compared to Standard condition. However, despite the addition of a visual source of information, i.e., a component seen to have the potential of causing increased cognitive load on trainees in a CSCL environment [25], it in fact does not impose more cognitive load on trainees. In fact, once the trainees gained additional experience with the Virtual Pointer-based instruction, it was less cognitively demanding [20,21,50]. This result is also in line with our previous findings that the benefits of a virtual pointer in improving the trainees' performance become evident after initial knowledge is gained [37].

Our further analysis shows the Virtual Pointer could improve communication efficiency between trainer and trainee. Trainees were able to perform more dual tasks – the laparoscopic task itself and watching/listening to the trainer's instruction – with the Virtual Pointer without being cognitively overloaded. Moreover, an increase in dual task instances in the Virtual Pointer condition did not lead to higher cognitive load. This is especially interesting considering previous research has demonstrated that performing two tasks simultaneously is more cognitively demanding [14]. Our findings that increased dual-tasks did not increase cognitive load suggests that the Virtual Pointer actually improves the communication efficiency between the trainee and the trainer, allowing the trainees to have further cognitive processing as suggested by Lim [18]. According to the modality effect in Cognitive Load Theory, receiving both verbal and visual instruction leaves more cognitive capacity for trainees to perform more cognitive processing. Thereby, the combined visual and audio instruction provided with the Virtual Pointer may open more cognitive processing space enabling the trainees to listen to instruction and perform the surgical task simultaneously. This contrasts with the Standard condition in which less efficient communication, i.e only audio guided instruction, requires greater mental effort to visualize the instruction and perform the laparoscopic task simultaneously. Being able to perform two tasks simultaneously – the laparoscopic task in addition to watching and listening to the trainer's instruction - may contribute to better comprehension and thus better performance as was indicated in [37] due to benefits of embodied learning [51]. The embodied learning theory states that learning and consolidating mastery occurs better by physically practicing the task than by mentally simulating it. Therefore, the better performance of trainees while using a Virtual Pointer [37] might be the result of embodied cognition which is gained through performing and listening to the instruction at the same time.

*Limitations*

This study was conducted in a simulated training environment. However, the Virtual Pointer is intended for instructing trainees while they are performing laparoscopic surgical tasks on real patients in the operating room - a more complex, stress-inducing, and cognitively demanding condition. In the operating room, even a slightly higher cognitive load may result in poorer performance which could be detrimental to real patients. Thus, further studies need to be conducted to evaluate the use of the Virtual Pointer and possible unintended consequences to workflow in a real life operating room environment. In addition, while EDA has been recognized as the most precise physiological signal for measuring cognitive load [52], it is still affected by other factors, such as activity and movement. Although we removed the segments of EDA data identified to be influenced by factors other than cognitive load, there is still a risk that some of the observed changes in EDA may have other underlying origins not associated with cognitive load. One concern is distinguishing stress from cognitive load in EDA-based measurement methods. However, results from [49] on discriminating between stress and cognitive load demonstrate that SCR interval rate is a predictor of stress, while the number of SCRs present, used as our measure, is a predictor of cognitive load. Additionally, use of the non-dominant wrist for collecting EDA data, although recommended by the Empatica E4 user manual, might miss some relevant information related to cognitive load from the dominant hand.

**Acknowledgements**

The authors gratefully acknowledge the Anne Arundel Medical Center for the use of equipment and space in the Simulation to Advanced Innovation and Learning (SAIL) Center, and would like to thank Ms. Katie Li and Ms. Jacqueline Mun for their support in data collection and analysis, as well as the participants who devoted their time to this study. This work was supported by NSF Grants IIS #1422671 and #1552837.

## References

1. Stanton NA, Young MS. A proposed psychological model of driving automation. Theoretical Issues in Ergonomics Science. 2000;1(4):315–31.
2. Young MS, Brookhuis KA, Wickens CD, Hancock PA. State of science: mental workload in ergonomics. Ergonomics. 2015;58(1):1–17.
3. Fraser K, Ma I, Teteris E, Baxter H, Wright B, McLaughlin K. Emotion, cognitive load and learning outcomes during simulation training. Med Educ. 2012 Nov;46(11):1055–62.
4. Ingalsbe G. The Emotional and Cognitive Impact of Unexpected Simulated Patient Death: A Randomized Controlled Trial. J Emerg Med. 2014;47(3):382.
5. Sweller J, Ayres P, Kalyuga S. Cognitive Load Theory. 2011.

6. Ayres P. Subjective measures of cognitive load: What can they reliably measure? In: Zheng RZ, editor. Cognitive Load Measurement and Application. Routledge; 2017. p. 9–28.
7. van Merriënboer JJG, Sweller J. Cognitive load theory in health professional education: design principles and strategies. Med Educ. 2010 Jan;44(1):85–93.
8. Sweller J. Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. Educ Psychol Rev. 2010;22(2):123–38.
9. Mayer RE, Moreno R. A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory [Internet]. Vol. 90, Journal of Educational Psychology. 1998. p. 312–20. Available from: http://dx.doi.org/10.1037/0022-0663.90.2.312
10. Miyake A, Shah P. Models of Working Memory: Mechanisms of Active Maintenance and Executive Control. Cambridge University Press; 1999. 506 p.
11. Baddeley AD, Logie RH. Working Memory: The Multiple-Component Model [Internet]. Models of Working Memory. p. 28–61. Available from: http://dx.doi.org/10.1017/cbo9781139174909.005
12. Chandler P, Sweller J. Cognitive Load Theory and the Format of Instruction [Internet]. Vol. 8, Cognition and Instruction. 1991. p. 293–332. Available from: http://dx.doi.org/10.1207/s1532690xci0804_2
13. Baddeley A. Working memory. C R Acad Sci III. 1998 Feb;321(2-3):167–73.
14. Brünken R, Steinbacher S, Plass JL, Leutner D. Assessment of cognitive load in multimedia learning using dual-task methodology. Exp Psychol. 2002;49(2):109–19.
15. Mayer RE, Moreno R. Nine Ways to Reduce Cognitive Load in Multimedia Learning. Educ Psychol. 2003;38(1):43–52.
16. Mayer RE, Moreno R, Boire M, Vagge S. Maximizing constructivist learning from multimedia communications by minimizing cognitive load [Internet]. Vol. 91, Journal of Educational Psychology. 1999. p. 638–43. Available from: http://dx.doi.org/10.1037/0022-0663.91.4.638
17. Fussell SR, Setlock LD, Parker EM, Yang J. Assessing the value of a cursor pointing device for remote collaboration on physical tasks. In: CHI ’03 extended abstracts on Human factors in computing systems - CHI '03 [Internet]. 2003. Available from: http://dx.doi.org/10.1145/765891.765992
18. Lim CP. Engaging learners in online learning environments. TechTrends. 2004;48(4):16–23.
19. Sweller J. Cognitive Load During Problem Solving: Effects on Learning. Cogn Sci. 1988;12(2):257–85.
20. Artino AR Jr. Cognitive load theory and the role of learner experience: An abbreviated review for educational practitioners. Aace Journal. 2008;16(4):425–39.
21. Kalyuga S, Chandler P, Sweller J. Incorporating learner experience into the design of multimedia instruction. J Educ Psychol. 2000;92(1):126–36.
22. Moreno R. Decreasing Cognitive Load for Novice Students: Effects of Explanatory versus Corrective Feedback in Discovery-Based Multimedia. Instructional Science. 2004;32(1/2):99–113.
23. Dillenbourg P. Collaborative Learning: Cognitive and Computational Approaches. Earli; 1999. 246 p.
24. Beers PJ, Boshuizen HPA, Kirschner PA, Gijselaers W, Westendorp J. Cognitive load measurements and stimulated recall interviews for studying the effects of information and communications technology. Educ Technol Res Dev. 2006;56(3):309–28.
25. van Bruggen JM, Kirschner PA, Jochems W. External representation of argumentation in CSCL and the management of cognitive load. Learning and Instruction. 2002;12(1):121–38.
26. Clark HH. Communities, commonalities, and communication. Rethinking linguistic relativity. 1996;17:324–55.
27. Convertino G, Mentis HM, Slavkovic A, Rosson MB, Carroll JM. Supporting common ground and awareness in emergency management planning. ACM Trans Comput Hum Interact. 2011;18(4):1–34.
28. Convertino G, Mentis HM, Ting AYW, Rosson MB, Carroll JM. How does common ground increase? [Internet]. Proceedings of the 2007 international ACM conference on Conference on supporting group work - GROUP ’07. 2007. Available from: http://dx.doi.org/10.1145/1316624.1316657
29. Sanford A, Anderson AH, Mullin J. Audio channel constraints in video-mediated communication [Internet]. Vol. 16, Interacting with Computers. 2004. p. 1069–94. Available from: http://dx.doi.org/10.1016/j.intcom.2004.06.015
30. Hancock JT, Dunham PJ. Language Use in Computer-Mediated Communication: The Role of Coordination Devices [Internet]. Vol. 31, Discourse Processes. 2001. p. 91–110. Available from: http://dx.doi.org/10.1207/s15326950dp3101_4
31. Clark HH, Brennan SE, Others. Grounding in communication. Perspectives on socially shared cognition. 1991;13(1991):127–49.
32. Fussell SR, Kraut RE, Siegel J. Coordination of communication. In: Proceedings of the 2000 ACM conference

on Computer supported cooperative work - CSCW ’00 [Internet]. 2000. Available from: http://dx.doi.org/10.1145/358916.358947

33. Cao CGL, Zhou M, Jones DB, Schwaitzberg SD. Can surgeons think and operate with haptics at the same time? J Gastrointest Surg. 2007 Nov;11(11):1564–9.
34. Andersen SAW, Mikkelsen PT, Konge L, Cayé-Thomasen P, Sørensen MS. Cognitive load in distributed and massed practice in virtual reality mastoidectomy simulation. Laryngoscope. 2016;126(2):E74–9.
35. Theodoraki MN, Ledderose GJ, Becker S, Leunig A, Arpe S, Luz M, et al. Mental distress and effort to engage an image-guided navigation system in the surgical training of endoscopic sinus surgery: a prospective, randomised clinical trial. Eur Arch Otorhinolaryngol. 2015 Apr;272(4):905–13.
36. O’Hara K, Gonzalez G, Penney G, Sellen A, Corish R, Mentis H, et al. Interactional Order and Constructed Ways of Seeing with Touchless Imaging Systems in Surgery. Comput Support Coop Work. 2014;23(3):299–337.
37. Feng Y, McGowan H, Semsar A, Zahiri HR, George IM, Turner T, et al. A virtual pointer to support the adoption of professional vision in laparoscopic training. Int J Comput Assist Radiol Surg. 2018 Sep;13(9):1463–72.
38. Mentis HM, O’Hara K, Gonzalez G, Sellen A, Corish R, Criminisi A, et al. Voice or Gesture in the Operating Room. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA ’15 [Internet]. 2015. Available from: http://dx.doi.org/10.1145/2702613.2702963
39. Yeh Y-Y, Wickens CD. Dissociation of Performance and Subjective Measures of Workload. Human Factors: The Journal of the Human Factors and Ergonomics Society. 1988;30(1):111–20.
40. Haji FA, Rojas D, Childs R, de Ribaupierre S, Dubrowski A. Measuring cognitive load: performance, mental effort and simulation task complexity. Med Educ. 2015;49(8):815–27.
41. Carswell CM, Clarke D, Seales WB. Assessing mental workload during laparoscopic surgery. Surg Innov. 2005;12(1):80–90.
42. Paas F, Tuovinen JE, Tabbers H, Van Gerven PWM. Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. Educ Psychol. 2003;38(1):63–71.
43. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: Advances in Psychology. 1988. p. 139–83.
44. Zheng R, Greenberg K. The boundary of different approaches in cognitive load measurement. In: Zheng RZ, editor. Cognitive Load Measurement and Application. 2017.
45. Lykken DT, Venables PH. Direct measurement of skin conductance: a proposal for standardization. Psychophysiology. 1971;8(5):656–72.
46. Dawson ME, Schell AM, Filion DL. The Electrodermal System. In: Handbook of Psychophysiology. 2007. p. 217–43.
47. Nourbakhsh N, Wang Y, Chen F, Calvo RA. Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks. In: Proceedings of the 24th Australian Computer-Human Interaction Conference on - OzCHI ’12 [Internet]. 2012. Available from: http://dx.doi.org/10.1145/2414536.2414602
48. Benedek M, Kaernbach C. A continuous measure of phasic electrodermal activity. J Neurosci Methods. 2010 Jun 30;190(1):80–91.
49. Setz C, Arnrich B, Schumm J, La Marca R, Tröster G, Ehlert U. Discriminating stress from cognitive load using a wearable EDA device. IEEE Trans Inf Technol Biomed. 2010 Mar;14(2):410–7.
50. De Waard D. The measurement of drivers’ mental workload. Vol. 7. Groningen University, Traffic Research Center. Chicago; 1996.
51. Kirsh D. Embodied cognition and the magical future of interaction design. ACM Trans Comput Hum Interact. 2013;20(1):1–30.
52. Shi Y, Ruiz N, Taib R, Choi E, Chen F. Galvanic skin response (GSR) as an index of cognitive load. In: CHI ’07 extended abstracts on Human factors in computing systems - CHI '07 [Internet]. 2007. Available from: http://dx.doi.org/10.1145/1240866.1241057