Eye-Tracking Technology to Determine Procedural Proficiency in Ultrasound-Guided Regional Anesthesia

G. Andrew Wright, MS4
Rahool Patel, MS3
Koraly Perez-Edgar, PhD

Xiaoxue Fu, PhD
Kayla Brown, MS

Sanjib Adhikary, MD
Adrian Zurca, MD, MEd

Introduction

Simulation training improves skill acquisition for clinical procedures that carry a high risk of complications and small room for error,[1] with evidence that simulation improves skill acquisition for ultrasound-guided regional anesthesia (UGRA) procedures.[2] Training programs are challenged to not only effectively and efficiently teach procedures to their trainees but also measure and certify trainees' procedural competence.[3] Current methodologies to assess procedural competence are crude, often relying on checklists or subjective observer ratings.[4,5] Little is known about how to determine expertise or proficiency with a procedure, which goes beyond knowing the steps or successfully completing a procedure.

Vision is an active process in which viewers seek out important visual information relevant to the task at hand. Attention plays a central role in visual and cognitive processing, and eye movements are considered to be an overt behavioral manifestation of the allocation of attention that can provide objective measures of how people make decisions. Eye tracking is a novel technology that has been used for various aspects of marketing research, including designing internet pages and evaluating internet advertisement placement,[6-8] and can dynamically and objectively measure the accuracy, efficiency, and attention patterns of individual health care professionals performing medical procedures. Eye tracking holds promise as

an additional tool to help anesthesiologists both teach and learn procedural skills, providing objective data with which to understand differences between expert and novice proceduralists that may impact both teaching strategies and assessment of procedural competence, and is being increasingly studied in various areas of medical education.[9,10]

Previous studies have investigated the use of eye-tracking technology for UGRA-related tasks;[11,12] however, none have evaluated attention patterns between groups of individuals at 3 different levels of learning (novice, intermediate, and expert), and none have measured the frequency of changes in participants' attention. The purpose of this pilot study was to determine the feasibility of using eye-tracking technology to assess attention and fixation when performing an UGRA procedure and to assess differences in patterns between expert, intermediate, and novice practitioners. Quantifying expertise based on stage of training, we hypothesized that experienced practitioners would display less fixation time and fewer gaze shifts to and from the monitor compared to the other groups. The primary end point was dwell time (fixation time) on the needle, monitor, and ultrasound probe. Secondary outcomes included feasibility metrics, number of gaze fixations, and satisfaction survey results. Dwell time and gaze fixation were compared between successful and unsuccessful completion of the procedure. These initial analyses treated the mobile eye-tracking data in a similar manner to

reaction time or response data typically found in computer-based tasks. However, mobile eye tracking allows researchers to capture the dynamic pattern of eye gaze across trials. We also hypothesized that experienced participants would have more stable visual attention when compared with novices, and we tested this hypothesis by comparing participant gaze entropy, a measure of visual attention stability that takes into account participant attention shifts, pattern shifts, and amount of time per shift.

Methods

Study approval was obtained from the Institutional Review Board at Penn State College of Medicine. Resident and attending physicians from the Department of Anesthesiology and medical students were recruited via email. Three cohorts were formed based on training level. Group 1 consisted of attending physician anesthesiologists who perform UGRA on a regular basis; this group was designated as expert level. Group 2 was made up of third-year anesthesia residents with some experience performing UGRA; this group was designated as intermediate level. Group 3 consisted of first-year or second-year medical students with no experience performing UGRA; this group was designated as novice level.

The procedure was conducted on a single Thiel-embalmed cadaver to ensure uniformity in identifying the anatomical landmarks. Each participant was fitted

with eye-tracking glasses (Pupil Labs, Berlin, Germany) and then went through a series of calibration exercises with software compatible with the eyeglasses. The participants' point of view and eye movements were recorded using the eye-tracking glasses. The ultrasound image during the procedure was also recorded. The participants were instructed to perform 3 consecutive ultrasound-guided sciatic nerve blocks (SNBs) on the cadaver model while wearing eye-tracking glasses. Participants were instructed to find the sciatic nerve and inject 2 to 3 mL of air around the nerve in order to determine the success of needle placement. Participants were not allowed to inject any air until they felt that they were touching the nerve structure with the needle tip. Air was injected instead of fluid to minimize distortion associated with multiple liquid injections. Additionally, ultrasonically air is bright, which made it easy to distinguish. The end point of the procedure was defined as when participants touched the nerve or when they felt that they had touched the nerve. The length of each procedure was measured from time of probe placement on the skin to injection of air via the needle. After finishing the first injection, participants were asked to put down the probe and reset the needle, at which point the timer was stopped and then restarted again when the participant started the next SNB.

After completion of 3 attempts, participants were asked to complete a survey regarding ease of use of the eye-tracking glasses, comfort with performing UGRA both before and after the study, and prior experience with UGRA. Participants' procedural success was determined by a blinded study team member with expertise in UGRA who reviewed the ultrasound video recordings and determined if the air had been injected in the appropriate location.

The eye-tracking data were coded and analyzed using Datavyu (datavyu.org, New York, New York), an application designed for researchers to review and code video-based data for analysis. For the eye-tracking data, gaze fixation was defined as eye gaze resting on a location for 3 consecutive frames (about 100 milliseconds) or longer. The parameters of each fixation included the location fixated and the amount of time (dwell time) spent fixated on the location. In order to illustrate patterns of gaze, we created state space grids using GridWare (www.statespacegrids.org, Kingston, Canada). Eye-gaze paths were then used to derive entropy to quantify the frequency with which gaze fixations changed from one location to another.

Initial analyses examined performance in the SNB procedure as a function of training level, examining success in finding the nerve using a χ^2 analysis, comparing number of successful attempts by training level. Turning to the data derived from the eye-tracking glasses, dwell time in task was examined using a 3 (trial number) by 3 (training level) repeated-measures ANOVA. Gaze patterns were further specified using a 3 (gaze location: Probe, Monitor, or Site) by 3 (trial number) by 3 (training level) repeated-measures ANOVA. Similar analyses were completed using the number and target of attention shifts as the variable of interest. A P value of <.05 was deemed to be statistically significant.

The participants were then subdivided based on having successfully completed the task. The analyses were repeated using a 3 (trial number) by 2 (success) ANOVA and a 3 (gaze location) by 3 (trial number) by 2 (success) ANOVA. Participant gaze entropy, a measure influenced by the number of attention shifts, pattern shifts, and amount of time per shift, was compared between novice versus experienced practitioners.

RESULTS

Five attending anesthesiologists, 5 third-year (categorical anesthesia year 2) anesthesiology residents, and 5 medical students (first-year and second-year) completed the study. Attendings had an average of 11.8 years of experience and reported having performed an average of 444 prior UGRA procedures. Residents had all completed at least 1 block of pain medicine rotation and reported having previously performed an average of 12.6 prior UGRA procedures. None of the medical students had any experience with use of ultrasound-guided procedures. Overall, the participants found wearing the eye-tracker easy (mean = 4.3/5) and reported little to no interference (mean = 1.5/5) from the headset. All of the participants completed the study, and none dropped out due to discomfort. Ten of the 15 participants (4/5 attendings, 2/5 residents, and 4/5 students) found the procedure realistic. Individuals with more experience were more likely to successfully perform the SNB (5/5 attendings, 5/5 residents, 0/5 students; χ^2 = 17.50, P = .002) and performed the procedure in less time (average time: attendings 62.6 ± 42.2 seconds, residents 106.4 ± 58.6 seconds, students 134.4 ± 96.8 seconds; P = .089).

When comparing the dwell time for the entire task, participants were progressively faster with practice (Trial 1: 41.8 seconds, Trial 2: 29.2 seconds, Trial 3: 28.9 seconds; F (2,26) = 5.28, P = .012, f = 0.64). In addition, participants spent more time with the ultrasound monitor (average 75.2 seconds per trial) relative to the needle insertion site (13.7 seconds per trial) and the ultrasound probe (11 seconds per trial) (F(2,26) = 37.07, P < .001, f = 1.69). When controlled for level of training, attending physicians spent significantly less time fixating on the ultrasound monitor compared to trainees (F(4,26) = 3.04, P = .035, f = 0.68) (Figure 1A).

When examining the number of fixations, there was also a trial effect, as the average number of eye shifts per trial decreased from 10.8 to 6.5 to 6 (F(2,26) = 5.59, P = .010, f = 0.66). In addition, more gaze shifts were directed to the ultrasound monitor (11.7) versus the needle insertion site (5.2) and the ultrasound probe (6.4) (F(2,26) = 22.93, P < .001, f = 1.33). Attendings physicians displayed fewer shifts in their gaze from one location to another during the procedure based on the average fixations per trial (attendings 5.1, residents 8.8, students 9.4; F(2,13) = 1.91, P = .19, f = 0.54). Attending physicians also deployed significantly fewer gaze shifts to the ultrasound monitor (7.6) versus students (14.3) and residents (13.1) (Figure 1B).

Comparisons Based on Successful SNB Completion

The analysis was repeated after differentiating groups based on successful completion of the SNB. Participants who completed the SNB successfully spent

significantly more dwell time to the ultrasound monitor compared to those who were unsuccessful ($F(2,26)=3.88$, $P = .034$, $f = 0.55$; Figure 2). There was no difference in average number of gaze shifts per trial for successful versus unsuccessful participants ($F(2,26) = 2.59$, $P = .095$, $f = 0.45$).

Analyses of Entropy

There was a progressive decrease in entropy among all participants from the Trial 1 to Trial 3 ($F(2,30) = 4.64$, $P = .03$, $f = 0.56$) (Figure 3A). Entropy also decreased with greater experience ($F(2,13) = 2.25$, $P = .15$, $f = 0.59$) (Figure 3B). There was a strong correlation between entropy and time on task ($r(16) = 0.826$, $P = .001$) (Figure 4). Students had an average entropy score of 3.35, while residents and attending physicians had average entropy scores of 2.51 and 1.36, respectively. Examples of individual stability/instability patterns across time and space for 3 participants are presented in Supplemental Online Material (Figures S1-S3) to illustrate the shifts in entropy that could be seen with training.

DISCUSSION

This pilot study confirms the hypothesis that the differences in experience and procedural success are associated with variations in visual attention patterns. As expected, providers with more experience were more likely to successfully complete the SNB procedure and do so more quickly. Providers performing SNB spend most of their time looking at the ultrasound monitor, regardless of experience level. However, attending physicians spend significantly less time looking at the monitor, which contributes to their ability to complete the procedure significantly faster. The number of eye shifts also decreased with practice, with most eye shifts being toward the ultrasound monitor. While attending physicians spent most of the procedure looking at the ultrasound monitor, they also had significantly fewer gaze shifts to the monitor, indicating that experts tend to maintain their gaze on the monitor and not look away from it. Finally, we demonstrated that visual entropy decreases with subsequent trials as well as with provider experience, and that there was a correlation between longer time spent

on the procedure with higher entropy. This indicates that decreased entropy may be an objective marker of experience with this type of procedure.

Eye tracking may offer the "missing link" in quantifying the acquisition of expertise in UGRA and other procedural specialties in medicine.[2] Our study demonstrates that experienced practitioners and novice practitioners differ in their approach to UGRA procedures and offers insight into the differences in gaze patterns based on level of experience. Providers of all levels spend most of their time looking at the ultrasound monitor; however, maturation in procedural skills appears to manifest itself by having fewer glances away from and back to the monitor. It is important to note that although residents were equally successful at performing the procedure as attendings, the eye-tracking data indicates differences in visual entropy that hints at additional nuances in performance. Thus, eye tracking could provide an objective method to identify more skilled or experienced learners, with the potential to not only evaluate success or competence but also quantify level of proficiency.

Entropy decreased across the 3 trials, regardless of provider expertise, providing evidence that providers are able to develop greater stability in eye-gaze with practice. As would be expected, entropy also decreased with training. Instability in visual attention was correlated to more time devoted to the procedure, and generally less likelihood of success. However, care must be taken when evaluating the importance of this correlation because, while the strong correlation between entropy and time on task weakened slightly with training, it was generally because the attending participants were much quicker. Essentially, the fact that attendings spend less time performing the procedure means less time for gaze instability. Nonetheless, entropy could be another marker to measure and define procedural competence and proficiency.

One prior study demonstrated eye tracking can identify expert patterns of ultrasound image interpretation; however, that study did not evaluate participants' attention patterns when performing an actual UGRA procedure. While significant differences were found between residents and attendings when it came to time looking

at areas of interest on static ultrasound images, the study did not seek to measure dynamic patterns in the gaze fixations.[11] A previous study compared eye-tracking data with only 1 expert and 1 novice learner performing an UGRA procedure, however, the lack of additional participants in each group meant that it was not powered to demonstrate significant differences between experts and novices. Additionally, the lack of an intermediate group meant that there was no way of evaluating for different levels of proficiency. Finally, the heat maps presented in that study do not take full advantage of the dynamic data obtained via eye tracking; whereas our study quantifies entropy in a way that makes it possible to more objectively compare visual patterns between participants.[12] Given the moderate sample size for this initial proof-of-concept study, the significance and effect size values should be treated as suggestive for future more robustly powered studies.

Studies in other medical specialties have demonstrated the power of eye-tracking in identifying *expert* attention patterns that may help differentiate learners' skill levels even before differences are noted using other more traditional measures of competence.[13] Eye tracking in UGRA may also be used to provide real-time information to learners by helping them understand the process by which experts successfully carry out a task. In parallel, real-time feedback for the learner themselves may also potentially accelerate acquisition of procedural competence. While cost for eye-tracking devices varies, the eyeglasses we used for this study costs about $2000, plus about $700 for a laptop to record the video.

Limitations to this study include the small number of participants, the simulated environment, and the lack of patient feedback to determine if nerve block was successful. Additionally, since we unfortunately only had access to 1 cadaver for this study, the participants each performed all trials on the same "patient." Therefore, we were unable to control for how much of the trial-to-trial improvement was because of repeated attempts on the same simulated patient. We also had to rely on face validity in developing a novel survey tool to determine participants' experiences before and during the

continued from previous page

procedure. Despite these limitations, this initial proof of concept study points to a number of potential strengths for eye-tracking procedures. While the current study was not sufficiently powered to fully examine these patterns, this study provides a glimpse of the patterns highlighting the difference between novice and experienced practitioners. Larger follow-up studies will hopefully be able to further explore differences in visual attention patterns between expert and intermediate providers.

In conclusion, this study demonstrates that eye tracking is easy to use, well-tolerated, and provides little interference during UGRA procedures. Experienced providers make fewer back-and-forth visual fixations, spend less time in the procedure, and demonstrate less entropy during the procedure. Additional study is needed to continue to explore the role of eye tracking in procedural training to measure and demonstrate competence and levels of proficiency, as well as to use this data to potentially accelerate skill acquisition.

References

1. Al-Elq A. Simulation-based medical teaching and learning. *J Fam Community Med*. 2010;17(1):35-40.

2. Niazi AU, Haldipur N, Prasad AG, Chan VW. Ultrasound-guided regional anesthesia performance in the early learning period: effect of simulation training. *Reg Anesth Pain Med*. 2012;37(1):51-4.

3. Accreditation Council for Graduate Medical Education. ACGME program requirements for graduate medical education in anesthesiology. https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/040_Anesthesiology_2021.pdf?ver=2021-06-22-114014-463. Accessed July 7, 2021.

4. Ramlogan RR, Chuan A, Mariano ER. Contemporary training methods in regional anaesthesia: fundamentals and innovations. *Anaesthesia*. 2021;76(S1):53-64.

5. Mariano ER, Harrison TK, Kim TE, et al. Evaluation of a standardized program for training practicing anesthesiologists in ultrasound-guided regional anesthesia skills. *J Ultrasound Med*. 2015;34(10):1883-93.

6. Maughan L, Gutnikov S, Stevens R. Like more, look more. Look more, like more: the evidence from eye-tracking. *J Brand Manag*. 2007;14(4):335-42.

7. Liu HC, Lai ML, Chuang HH. Using eye-tracking technology to investigate the redundant effect of multimedia web pages on viewers' cognitive processes. *Comput Human Behav*. 2011;27(6):2410-7.

8. Chandon P, Hutchinson JW, Bradlow ET, Young SH. Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *J Mark*. 2009;73(6):1-17.

9. Kim E. Potential of eye tracking technology for assessment of performance and medical education in the field of anesthesia. *Korean J Anesthesiol*. 2018;71(4):253-4.

10. Tien T, Pucher PH, Sodergren MH, et al. Eye tracking for skills assessment and training: a systematic review. *J Surg Res*. 2014;191(1):169-78.

11. Borg LK, Harrison TK, Kou A, et al. Preliminary experience using eye-tracking technology to differentiate novice and expert image interpretation for ultrasound-guided regional anesthesia. *J Ultrasound Med*. 2018;37(2):329-36.

12. Harrison TK, Kim TE, Kou A, et al. Feasibility of eye-tracking technology to quantify expertise in ultrasound-guided regional anesthesia. *J Anesth*. 2016;30(3):530-3.

13. Kelly BS, Rainford LA, Darcy SP, Kavanagh EC, Toomey RJ. The development of expertise in radiology: in chest radiograph interpretation, "expert" search pattern may predate "expert" levels of diagnostic accuracy for pneumothorax identification. *Radiology*. 2016;280(1):252-60.

The following authors are at Penn State College of Medicine, Milton S. Hershey Medical Center, Hershey, PA: **G. Andrew Wright** *is a medical student;* **Rahool Patel** *is a medical student;* **Sanjib Adhikary** *is a Professor of Anesthesiology and Perioperative Medicine;* **Adrian Zurca** *is an Associate Professor of Pediatrics. The following authors are at The Pennsylvania State University, Department of Psychology, University Park, PA:* **Koraly Perez-Edgar** *is a Professor of Child Studies and Professor of Psychology;* **Kayla Brown** *is a graduate student in the Developmental Psychology program.* **Xiaoxue Fu** *is a graduate student in the Developmental Psychology programat The Ohio State University, Nationwide Children's Hospital, Columbus, OH.*

Corresponding author: *Adrian D. Zurca, MD, MEd, Penn State College of Medicine, Department of Pediatrics, 500 University Drive, Mail Code H085, P.O. Box 850, Hershey, PA 17033. Telephone (717) 531-5337, Fax: (717) 531-0809.*

Email address: *Adrian D. Zurca: azurca@pennstatehealth.psu.edu*

Abstract

Background: Eye-tracking measures attention patterns, which may offer insight into evaluating procedural expertise. The purpose of this study was to determine the feasibility of using eye tracking to assess visual fixation patterns when performing an ultrasound-guided regional anesthesia procedure and to assess for differences between experienced, intermediate, and novice practitioners.

Methods: Participants performed an ultrasound-guided sciatic nerve block 3 times on a fresh cadaver model while wearing eye-tracking glasses. Gaze fixation and dwell time on each location were compared between participants. Eye-gaze paths were used to derive a measure of entropy, or how often participants switched gaze fixations between locations.

Results: Five attending anesthesiologists, 5 third-year anesthesiology residents with prior ultrasound-guided regional anesthesia experience, and 5 medical students completed the study. Individuals with more experience were more likely to successfully perform the sciatic nerve block (5/5 attendings, 5/5 residents, 0/5 students; $P = .002$) and performed the procedure faster (average: attendings 62.6 seconds, residents 106.4 seconds, students 134.4 seconds; $P = .089$). Participants were progressively faster with practice (Trial 1: 41.8 seconds, Trial 2: 29.2 seconds, Trial 3: 28.9 seconds; $P = .012$), and the average number of eye shifts per trial decreased from 10.8 to 6.5 to 6 ($P = .010$). Attending physicians spent significantly less time fixating on the ultrasound monitor compared to trainees ($P = .035$). Average visual entropy progressively decreased from Trial 1 to Trial 3 ($P = .03$) and with greater experience ($P = .15$). There was a strong correlation between entropy and time on task ($r(16) = 0.826$, $P = .001$).

Conclusions: Experienced providers make fewer back-and-forth visual fixations, spend less time in the procedure, and demonstrate less entropy during ultrasound-guided regional anesthesia procedures. Mobile eye-tracking has the potential to provide additional objective measures of performance that may help not only determine procedural competence but also distinguish between levels of proficiency.

Keywords: Regional anesthesia, medical education, clinical competency, interventional ultrasound

Figures

Figure 1A. *Average dwell time fixated on needle insertion site, ultrasound monitor, and ultrasound probe per trial, per level of training. * P < .05*
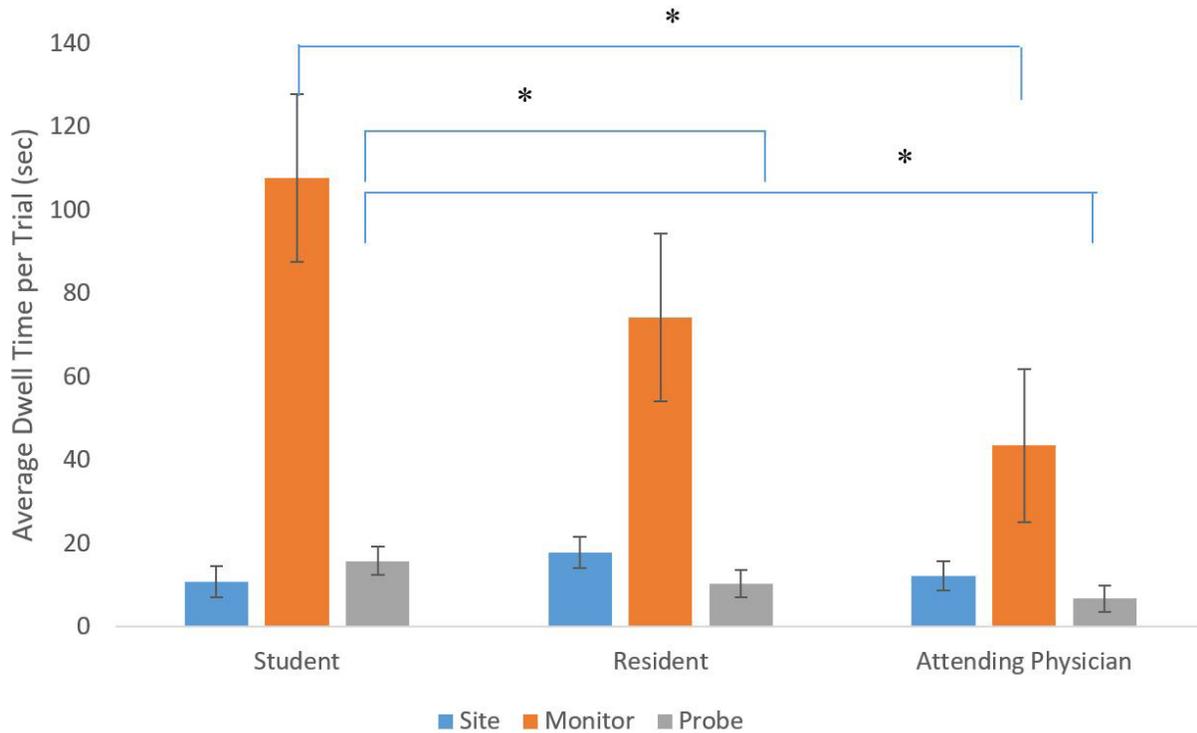

Figure 1B. *Average number of gaze shifts per trial to insertion site, ultrasound monitor, and ultrasound probe, by level of training. *P < .05*
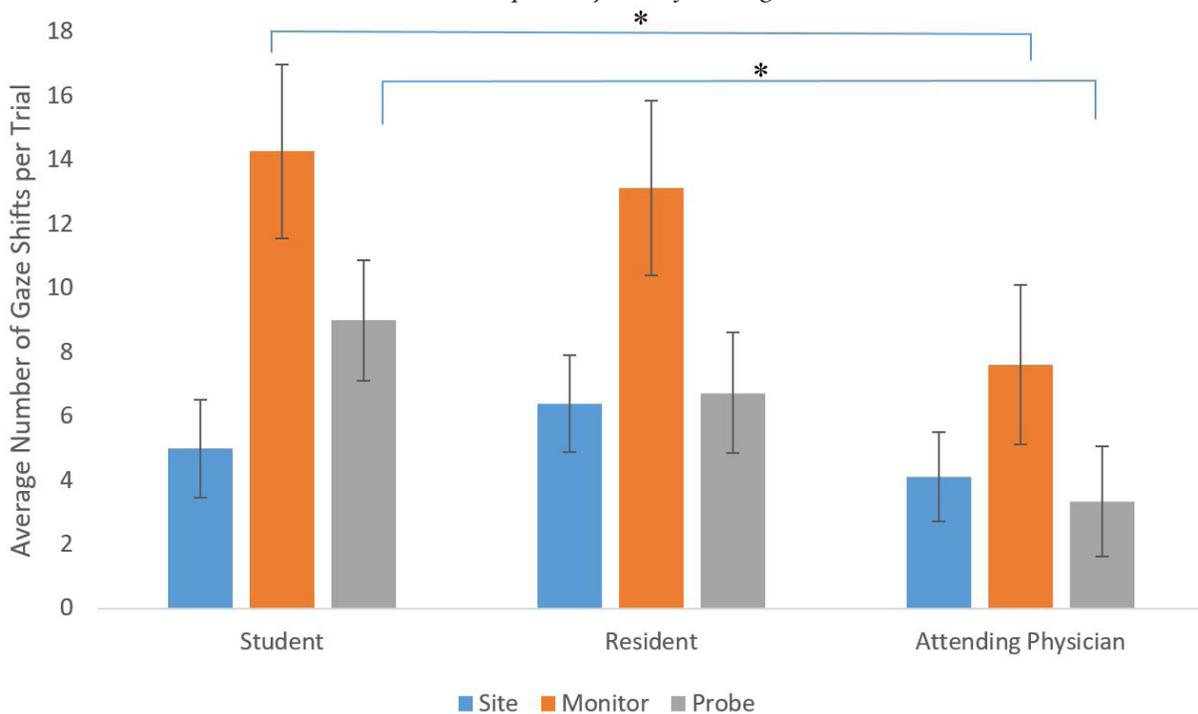
Figures *continued*

Figure 2. Average dwell time per trial for each gaze location for participants successful in performing sciatic nerve block versus those who were unsuccessful. + P < .10
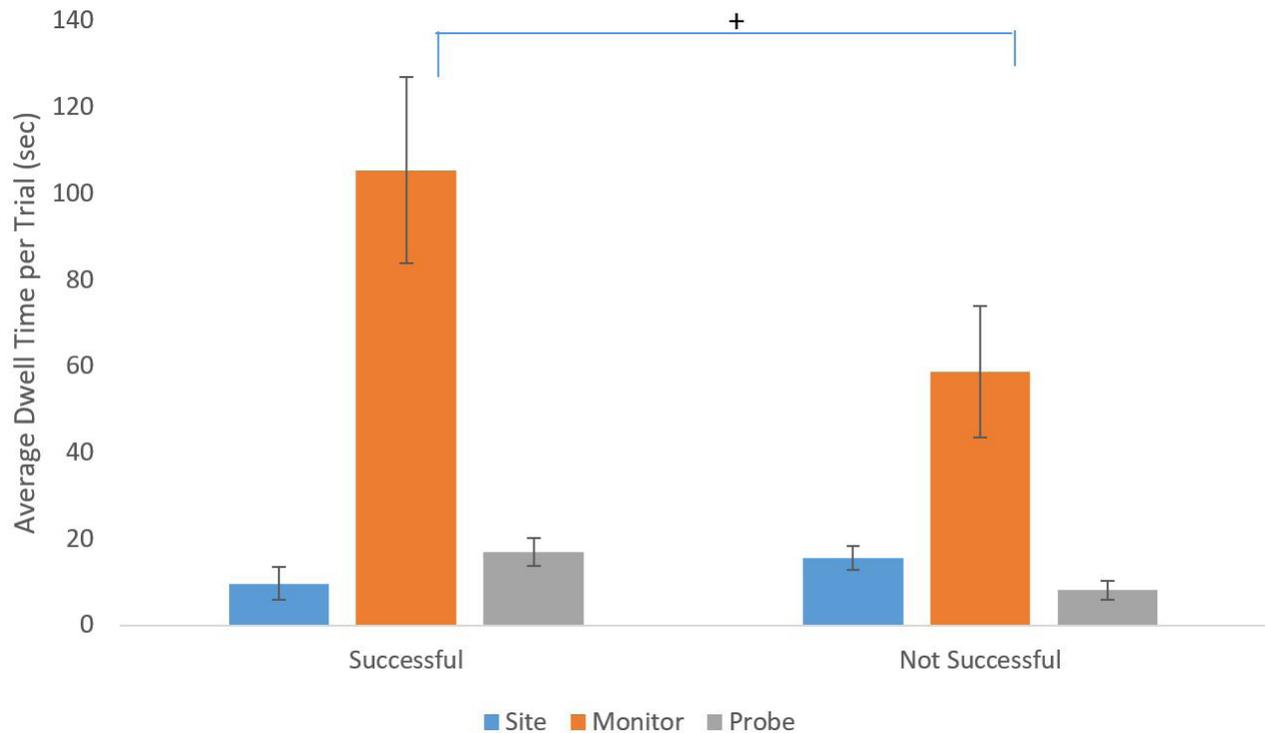
Figure 3A. Average total entropy score for all participants during each trial. Entropy decreases across the 3 trials $(F(2,30) = 4.64, P = .03, f = 0.56)$.
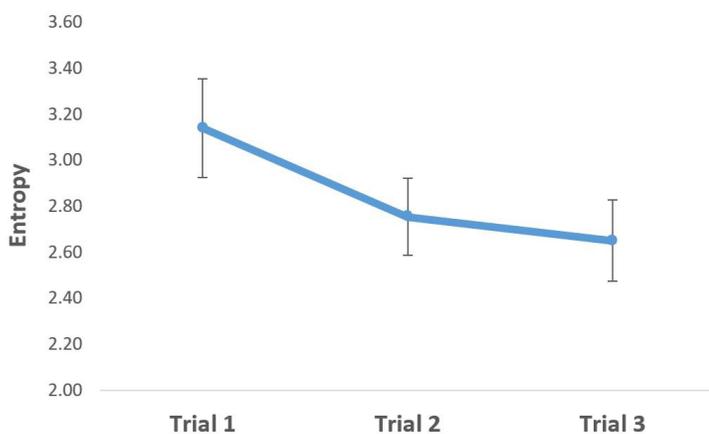
Figure 3B. Average total entropy score for each group, based on level of training. Entropy decreases with increasing level of training. Effect driven by difference between students and attendings $(F(2,13) = 2.25, P = .15, f = 0.59)$.
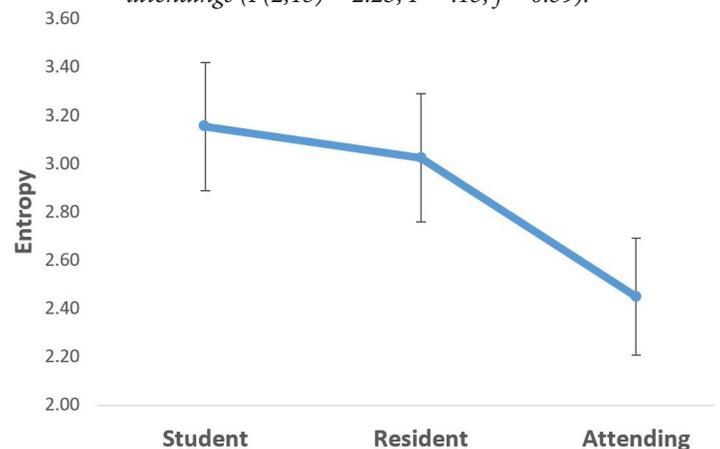
Figures *continued*

Figure 4. Correlation between average time per trial and average entropy score, by training level
*(r(16) = 0.826, P = .001).*
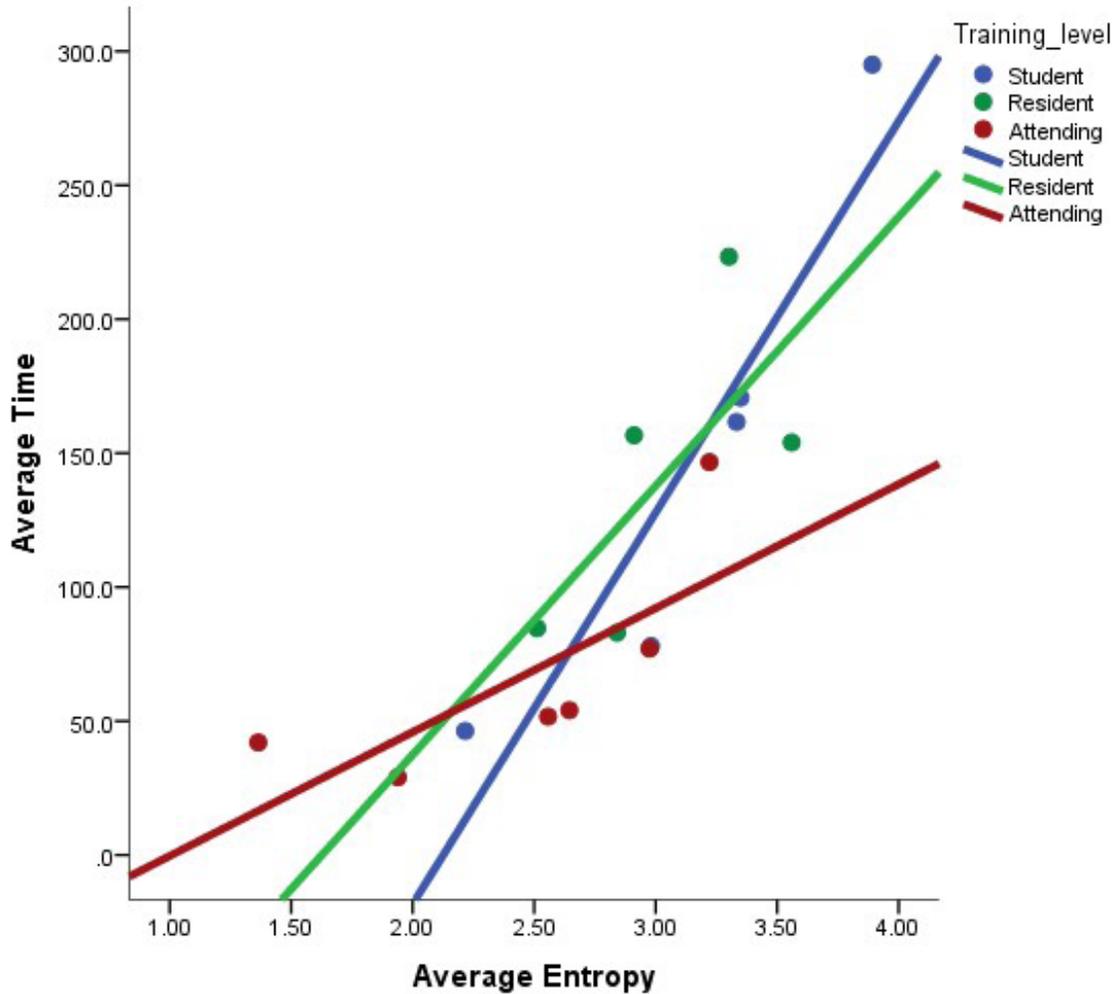
Supplemental Online Material

Figure S1. Entropy over state-space grid for student participant, trial 1.
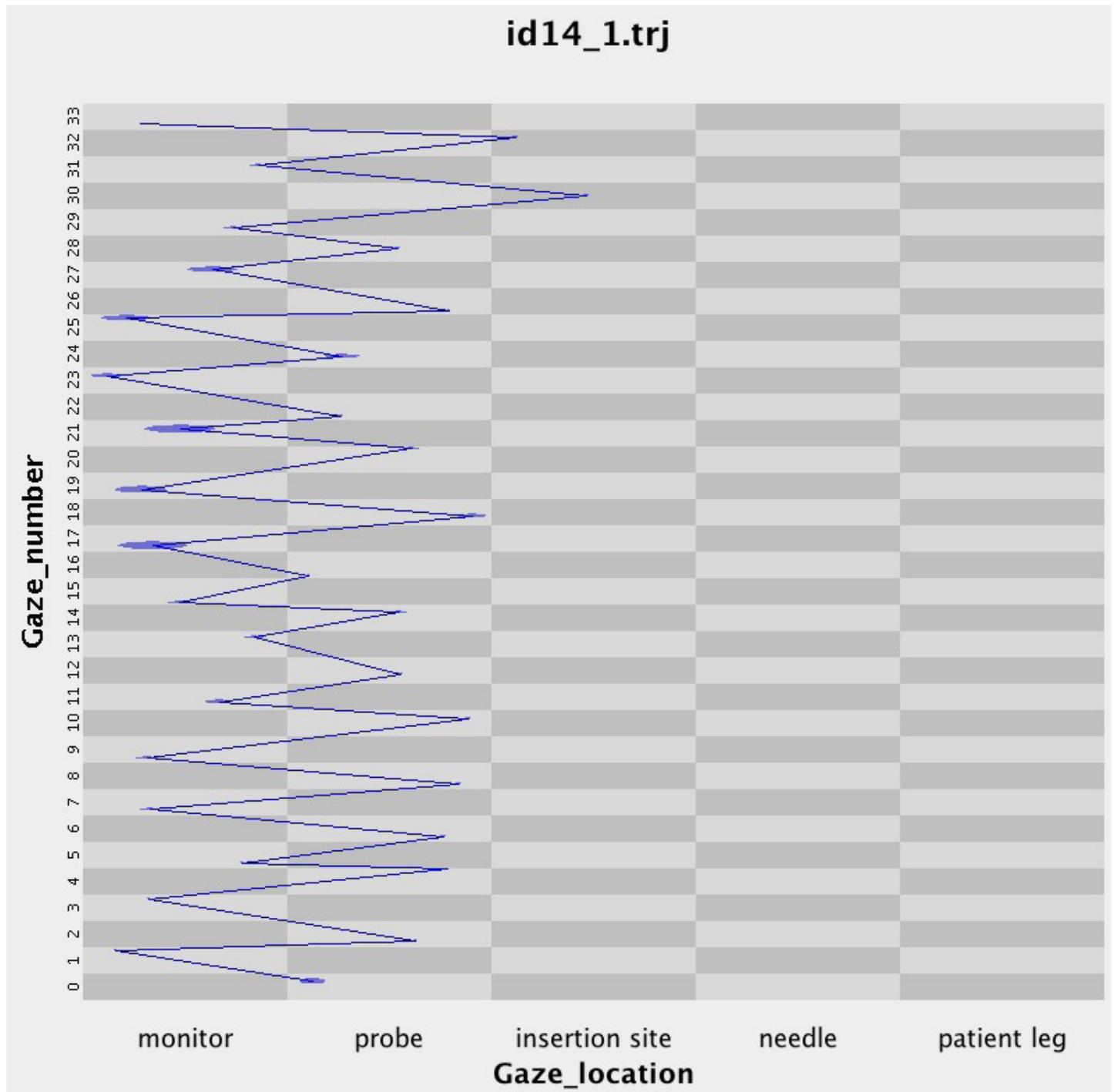
Supplemental Online Material *continued*

Figure S2. *Entropy over state-space grid for resident participant, trial 2.*
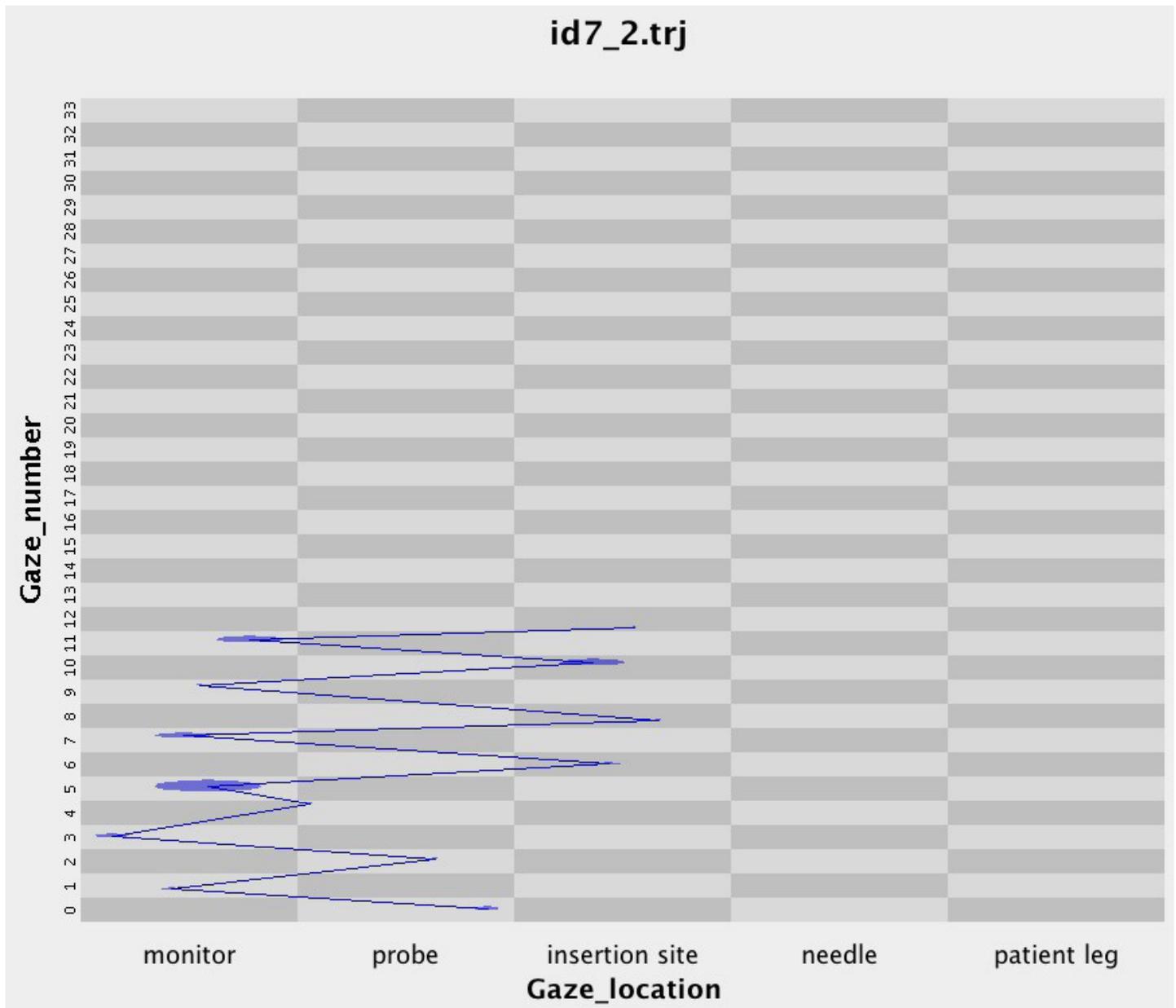
Supplemental Online Material *continued*

Figure S3. Entropy over state-space grid for attending physician participant, trial 3.