RAPID PREPARATION OF EYE TRACKING DATA FOR DEBRIEFING IN MEDICAL TRAINING: A FEASIBILITY STUDY

E. Wolf[1], R. Heinrich[1], A. Michalek[1], D. Schraudt[1], A. Hohm[1], R. Hein[1], T. Grundgeiger[1], & O. Happel[2]

[1]Institute Human-Computer-Media, Julius-Maximilians-Universität Würzburg, Germany
[2]Department of Anesthesia and Critical Care, University Hospital Würzburg, Germany

Simulation-based medical training is an increasingly used method to improve the technical and non-technical performance of clinical staff. An essential part of training is the debriefing of the participants, often using audio, video, or even eye tracking recordings. We conducted a practice-oriented feasibility study to test an eye tracking data preparation procedure, which automatically provided information about the gaze distribution on areas of interest such as the vital sign monitor or the patient simulator. We acquired eye tracking data during three simulation scenarios and provided gaze distribution data for debriefing within 30 minutes. Additionally, we qualitatively evaluated the usefulness of the generated eye tracking data for debriefings. Participating students and debriefers were mostly positive about the data provided; however, future research should improve the technical side of the procedure and investigate best practices regarding how to present and use the data in debriefings.

INTRODUCTION

Medical errors are estimated to be the third leading cause of death in the US (Makary & Daniel, 2016). In recent decades, simulation-based education has emerged as an effective tool in teaching strategies for error prevention (Weller, Nestel, Marshall, Brooks, & Conn, 2012). A simulation-based medical training session usually consists of a simulated training scenario followed by a debriefing where participants can reflect on and learn from their experiences during the scenario (Fanning & Gaba, 2007; Garden, Le Fevre, Waddington, & Weller, 2015; Sawyer, Eppich, Brett-Fleegler, Grant, & Cheng, 2016). Many debriefing methods include video recordings of the scenarios. Recently, researchers have also used eye tracking videos for debriefing (Henneman et al., 2014; Henneman, Marquard, Fisher, & Gawlinski, 2017; Kok & Jarodzka, 2017). In this study, we evaluated (1) the technical feasibility of providing not only the eye tracking video but also gaze distribution data, and (2) the usefulness of gaze distribution data for debriefings.

Eye tracking is a method that enables researchers to record objects in the environment that an individual is looking at. The eye-mind-hypothesis (Just & Carpenter, 1980) suggests that an individual pays attention to the object that one is looking at. Eye tracking data, therefore, enables to study the cognitive processes of an individual (Kok & Jarodzka, 2017; Land, Mennie, & Rusted, 1999).

Due to its potential, eye tracking has been used as a training method and feedback tool in the context of medical training, not only to improve teaching but also to investigate the expertise or instructional design and to optimize the latter (for a recent review, see Ashraf et al., 2018). Research has demonstrated that adding eye tracking to the debriefing process was perceived as beneficial by the participants (Browning et al., 2016; Marquard et al., 2011; O'Meara et al., 2015). Furthermore, Henneman et al. (2014) showed that safety practices improved most with eye tracking debriefing only compared to verbal debriefing or verbal plus eye tracking debriefing. However, the time needed to process and prepare the data is a problem when working with eye tracking data. Previous research was limited by time restrictions for analyzing the available data and therefore only used eye tracking videos and not quantified results such as gaze distributions (e.g., Henneman et al., 2014).

We aimed to enrich debriefings by incorporating not only eye tracking videos, but also automatically generated gaze distribution data. Specifically, we generated gaze distributions on different areas of interest (AOI), such as the vital sign monitor, during simulated medical scenarios. We evaluated (1) the technical feasibility of this data preparation procedure, and (2) the usefulness of gaze distribution data for debriefings.

PRACTICE INNOVATION

Data captured during a medical simulation scenario – such as video recordings from stationary cameras – are often used immediately after the simulation in a debriefing (Fanning & Gaba, 2007). The immediate preparation of the captured data in a presentable form is crucial in order to provide participants of the training with detailed insights into their training scenario (Browning et al., 2016).

Analyzing eye tracking data, such as manually assigning fixations to AOIs, however, tends to be quite labor and time intensive. Therefore, the primary field of application for eye tracking in medical debriefings is the presentation of the captured videos (Kok & Jarodzka, 2017). To overcome this limitation, we decided to use an AOI tracking technique based on markers which can be analyzed automatically. The gaze distribution between multiple relevant AOIs in a simulation environment could be used by a debriefer to provide additional information about the participant's visual attention during the scenario.

Foremost, the necessary data analysis procedure needs to enable for fast preparation of eye tracking data. Considering the literature (e.g., Henneman et al., 2014), insights from an observation of a training session, and discussions with three subject matter experts, we identified the following requirements:

- The procedure should provide an eye tracking recording of the scenario and additional data about the gaze distributions of participants.
- The captured data should be prepared within a maximum of 30 minutes in order to be used in the debriefing.

 DOI 10.1177/1071181319631032

- The procedure should be tested within simulation scenarios with a maximum duration of 20 minutes.
- The procedure should be usable for various training scenarios.

From Training to Debriefing in 30 Minutes

Based on discussion with a medical simulation expert (author OH) and consideration of the literature (Grundgeiger, Klöffel, Mohme, Wurmb, & Happel, 2017; Schulz et al., 2011), we identified thirteen AOIs which needed to be tracked (Table 1). We used a cost-effective eye tracker from Pupil Labs (Kassner, Patera, & Bulling, 2014) and its associated software. The identified AOIs were tagged using ArUco surface tracking markers (Romero-Ramirez, Muñoz-Salinas, & Medina-Carnicer, 2018) which allow for fast and robust detection and can be tracked by the eye tracking software. Depending on the estimated gaze distance between participants and each AOI, we used makers of different sizes, between 3x3 cm and 15x15 cm (Figure 1). AOIs were tagged with at least two, and up to six, markers, depending on the occlusion probability and the size of the AOI. The eye tracking analysis software is able to automatically analyze the gaze distribution on AOIs after an eye tracking video is captured and allows the data to be exported into comma-separated values (CSV) files.

Because recording via Wi-Fi was only experimentally supported by the software and had issues such as frame drops or disconnections, we decided to record directly onto a capturing device. To avoid the time-consuming transfer of the captured eye tracking recording from a mobile device onto a computer for computationally intensive data analysis, we recorded the files on a notebook stored in a custom-made backpack carried by the participants. Because audio capture would have required an additional external microphone, we decided to rely on the audio recording of the stationary audio-/video-recording system of the simulation center.

The captured eye tracking recording was transferred to the Pupil Labs software directly after each scenario. Post-processing of the recording, including the marker and AOI detection, as well as gaze distribution calculation, took approximately 15-20 minutes for a recording with a length of 20 minutes.

Figure 1. Marker placement in a training scenario at the simulation center of the Department of Anesthesia and Critical Care of the University Hospital Würzburg.

The export of the files took approximately another two minutes. Based on the input of the debriefer, which was logged manually by one of the experimenters in the control room, different periods of the scenario (e.g., a participant enters the room, or the patient deteriorates) could be exported. The tight schedule between training and debriefing meant that we were not able to calculate fixations but used the recorded frame rate to analyze the gaze allocation. The scene camera recorded with a frame rate of 30 Hz. We used each frame in the analysis and therefore determined the gaze allocation every 33.3 ms.

The CSV files generated during the export contained the raw information about when a participant looked at an AOI. A prepared R (R Core Team, 2013) script processed this data and calculated the total time that each AOI was looked at. The script generated a pie chart that showed the proportion of gazes towards each AOI and a bar chart that showed the absolute time an AOI was looked at. Because the charts were intended to be displayed in a presentation, the same data was reformatted into charts more suitable for a visual presentation in the present paper (Figure 2 and 3).

FINDINGS

The introduced procedure was tested in three simulation scenarios at the simulation center of the Department of Anesthesia and Critical Care of the University Hospital Würzburg. Five medical students participated in the training with different roles in each scenario, such as the anesthesiologist, the surgeon, or a nurse. One participating student wore the eye tracker during each scenario. The debriefing focused on the individual and interpersonal social skills of the participants, like situation awareness, decision making, communication, and teamwork.

Prior to the scenario, the participant put on the eye tracker. An experimenter made sure that the eye cameras were recording the eyes and the world camera showed the correct excerpt from the field of view of the participant. The eye tracker was then calibrated using the Pupil Labs single marker calibration technique which always led to an accuracy better than two degrees.

As soon as the participant entered the scenario, the recording was started, and the laptop was stored in the backpack. Directly after the simulation scenario, the laptop and eye tracker were taken off the participant, and the eye tracking data were automatically analyzed. The prepared data were copied onto a USB stick and handed to the debriefer who had already started the debriefing by discussing the training and showing recordings from the stationary cameras. The eye tracking data was then integrated into the ongoing debriefing.

Validity of AOI Detection

To ensure the validity of the captured data, we analyzed one of the recordings after the training session. Using the offline fixation detection plugin of the eye tracking software, we calculated fixations for the recording and calculated the gaze distribution using an adapted version of our R script. We also manually determined the distribution between the AOIs by counting the individual fixations on the AOIs. A comparison of the data using Cohen's Kappa for each AOI revealed a good accuracy for most AOIs (Table 1). We calculated independent Kappa's for each AOI to identify poorly tracked AOIs. The average fixation count weighted Kappa was 0.71 which can be considered as good (Fleiss, Levin, & Paik, 2013). AOIs with a high risk of marker occlusion had insufficient accuracy which needs to be addressed in the future. Additionally, the AOI for the surgical team behind the sterile drape also had a high number of incorrectly detected fixations, since the AOI often overlapped with other AOIs.

Table 1
Overview of areas of interest with manually counted fixations per AOIs, percentage agreement, and Cohen's Kappa.

Area of Interest	Fixations on AOI	Percentage agreement	Cohen's Kappa
Anesthesia trolley	19	68.42	0.69
Drawer anesthesia Trolley	22	86.36	0.88
Monitor right	1	100	0.50
Monitor left	100	93.00	0.92
Anesthesia machine	84	86.90	0.87
Anesthetic vaporizer	7	28.57	0.39
Suction for secretion	0	-	-
Ventilation tubing	9	66.67	0.80
Syringe pumps	4	50.00	0.66
Mannequin's arm	16	25.00	0.37
Mannequin's chest	91	54.95	0.63
Surgical team behind drape	20	70.00	0.50
Telephone	0	-	-
Whitespace	104	79.81	0.53

We observed that the amount of whitespace in AOI detection differed between the single frame-based and the fixations-based analysis. Whitespace contains all gazes on objects that were not part of our AOI list. The single frame-based analysis produced more whitespace (Figure 2). The frame-based analysis is likely to include frames in which no information perception is possible because the gaze is not allocated to a single point in the environment long enough (Holmqvist et al., 2011). The fixations-based analysis, therefore, returned more valid results.

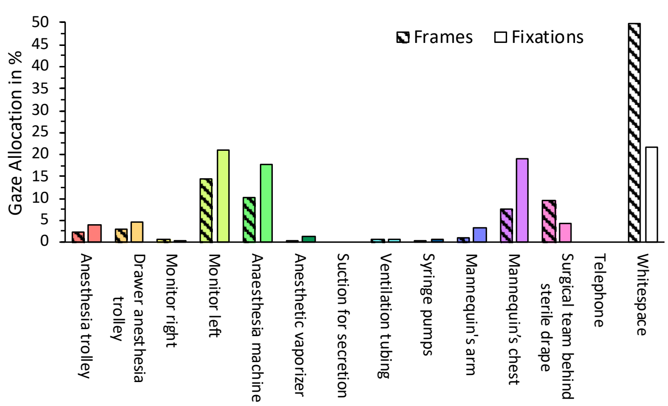

Figure 2. Bar chart with gaze distribution of one participant's eye tracking recoding between areas of interest based on frames and fixations. Whitespace indicates fixations on objects that were not part of our areas of interest.

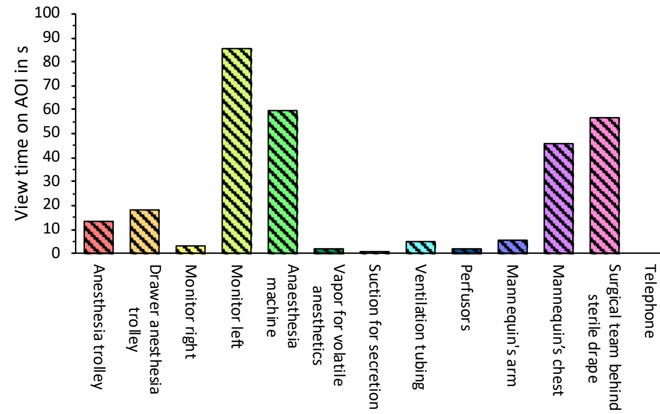

Figure 3. View times on areas of interest of one participant's eye tracking recoding based on frames.

Qualitative Evaluation of Debriefing

After the three trainings and debriefings, all five participating medical students and the three debriefers were interviewed using a standardized interview. The interviews were recorded, and three researchers summarized the statements in codes. An evaluation of the qualitative data is presented below.

Evaluation of student interviews. All five students had already participated in a simulation without eye tracking, three of them took part in previous simulation training in the context of anesthesia. Table 2 shows the expectations of the participants and the perceived distraction by the eye tracker and marker, as well as their perceived benefits of the graphs and eye tracking videos shown in the debriefing. Some participants were ambivalent and had both, positive and negative expectations.

Before the simulator training, the participants were relatively skeptical about the integration of the eye tracker and markers. One of the main concerns was that the attached markers could interfere with the training; however, this expectation was not confirmed during the simulation. One student stated that he did not pay attention to the markers during the simulation because he was so focused on the medical tasks. Despite the eye tracker, all participants perceived the simulation to be realistic or even very realistic.

The eye tracking data were seen as a good extension to the usual debriefing technique. The first-person perspective video offered different new visual aspects which were perceived as very enriching, and the generated gaze distribution charts were seen as a useful additional source of information.

Evaluation of debriefer interviews. The three experienced debriefers for simulation scenarios were also interviewed. Before the study, the debriefers were somewhat skeptical or had minimal expectations of the study. Because of their experience with eye tracking in other experiments, two of the three debriefers were curious about whether eye tacking would provide any added value at all.

Despite initial skepticism, the debriefers general impression of the training was very positive. In their opinion, the participants accepted the study very openly and were even more focused than usual.

Table 2
Results of the training participant interviews.

#	Code	Statement
	Expectations	
3	No expectations	-
2	Positive expectations	Additional information; better traceability of visual attention
2	Negative expectations	Overload; worries that markers might be interfering
	Distraction	
5	No distraction or manipulation through the eye tracker	During high activity the eye tracker was faded out; otherwise, it was perceived but not as disturbing
3	No distraction through markers	-
2	Distraction through markers	Distraction at the beginning of the simulation, but quickly got used to it
	Usefulness	
4	First person perspective video with gaze positions	Very useful, thoughts can be reconstructed easily; valuable addition to the static video recording
4	Awareness of one's own perception, reflection and learning from mistakes	Increased awareness of one's own gaze behavior (e.g. focusing, quick looking back and forth)
2	Graphs with fixation duration on AOI	Useful to get a general idea
1	No additional benefit from the eye tracking data	-
	Improvements	
3	Additional features for the video	For a better traceability of the situation the video requires sound
2	Comparison between experts and novices	Additional informational benefit, better traceability of visual attention
2	Graphs with fixation duration on AOI	Expandable in terms of color selection; useful for supporting video

There was no interference with the usual simulation procedure – neither in the actual simulation nor in the debriefing. Despite the use of an eye tracker and markers, the scenario was not perceived as different to other training scenarios. One debriefer stated that the communication patterns of the participants in the scenario did not differ from previous trainings and thus were apparently not influenced by the eye tracker, the eye tracker was "faded out", and the participants moved naturally.

After the simulation training and in contrast to their expectations, the debriefers considered the use of the eye tracker during simulations and the use of the data gained to have great potential for simulation-based training. The instructors found the following questions to be highly interesting:

- Where did the participants actually look?
- How fast did the participants switch between the AOIs?
- Was the debriefer able to estimate the participants' focus of attention during the simulation?

Debriefing with eye tracking data seemed to have a greater value than debriefings without eye tracking data due to the recorded video. The graphs provided an additional visual representation of the gaze distribution and served as a good summary of the video analysis. The debriefers were enthusiastic about their experience with the method and indicated further ideas for future simulation training with eye tracking:

- The dwell times on single AOIs or AOI groups (i.e., monitoring equipment) should be better visualized using an additional timeline.
- The relevant phases of the training (i.e., start of scenario until patient deteriorates) should be documented more precisely. The person noting the phases should have sufficient knowledge of the training to identify the relevant parts. The gaze distributions should be separated by these phases.
- Eye tracking could be used to explore the gaze distribution of the debriefer while watching the video recordings of a scenario to train the trainers.
- Use of multiple eye trackers in one simulation to capture the perspective of different participants.
- The application should be able to track moving AOIs such as other participants.
- Repeated scan paths and gaze routines should be detected and visualized.
- The standard simulation video and the eye tracking video should be played simultaneously during the debriefing.
- The eye tracking video should contain audio recording.

DISCUSSION

We have presented a novel procedure to provide gaze distribution data as a feedback tool for debriefing in simulation-based education. We showed that the use of eye tracking in a medical simulation scenario could contribute to the debriefing in the form of video recordings with gaze visualization and enriching graphs regarding gaze distributions.

From the technical perspective, we were able to prepare gaze distribution data almost automatically within the self-imposed time limit and with acceptable accuracy. We analyzed the provided eye tracking data and compared it to manually analyzed data with good results. Although some AOIs were hard to track, the majority of AOIs were tracked sufficiently. However, we were not able to address all our goals due to technical limitations. We used a bulky laptop for the recording which the participant had to carry in a backpack, and we did not record any audio. Instead of fixation-based gaze distribution analysis, we had to use a frame-based approach. Fixations can be considered as the times when participants gather information from their surrounding (Holmqvist et al., 2011), a fixation based analysis should be preferred; however, we were not able to calculate fixations within the time constraints.

The qualitative evaluation showed that the eye tracking video was perceived as a positive contribution to the debriefing by the participants and the debriefers. The gaze distribution graphs of the eye tracking data, although initially seen as a good addition, could not provide the expected helpful indications for the participants without information about how to interpret the data. The debriefers were only able to tell the participants that they were looking at specific areas for a certain amount of time within a predefined period but strategies on how to interpret the data were missing. This issue could be addressed through comparison with expert data (Ashraf et al., 2018) or visualizations along a time axis. Because this kind of information was missing for our visualization, it was difficult to use the graphs.

This work has several limitations. First, the gaze distribution calculations were based on frames and not on fixations due to technical limitations. Further data analysis after the study found that there was a difference between the two calculation methods. Future studies should preferentially use data analysis based on fixations.

Second, the used capture method had limitations: participants had to wear a backpack with a recording device on their back, sound recording was not enabled, and the resulting files were uncompressed and therefore very large, and only partly preprocessed. Further studies should investigate the use of a smaller recording device that enables streaming of the compressed captured data directly via Wi-Fi to the analysis device which could then perform further preprocessing in real time.

Third, AOIs were not always tracked correctly due to occlusions and poor marker recognition. Small AOIs with few markers were particularly difficult to detect and were less accurately recognized. Future experiments should work on improved marker distribution and detection. In order to minimize automatic evaluation, the optimal distribution of markers on AOIs remains to be found.

Fourth, we only provided data about the general gaze distribution for the whole scenario or particular events in the scenario. However, it would be interesting to obtain further information on how the gaze distribution changed between uneventful and stressful parts (e.g., patient deterioration) during the training (cf. Schulz et al., 2011).

Fifth, the current method did not add other trainees as trackable AOIs. Further studies should consider marker placement on the surgical caps and masks of participants, or even more complex techniques such as picture-based automatic face detection.

In conclusion, we demonstrated that it is technically feasible to provide gaze distribution data within an acceptable time frame for debriefings of medical simulation training. The evaluation of the debriefing, however, showed that more research is needed to incorporate the data in the debriefing process. Overall, the results of this paper contribute to the use of eye-tracking videos and gaze distribution data for the ongoing improvement of medical simulation training and have identified future research directions.

PRACTITIONER TAKE-AWAYS

- Debriefers are able to automatically generate gaze distribution data within a short period.
- Debriefers are able to provide additional information in the form of figures about a participant's gaze distribution.
- Debriefers and participants are able to obtain further insights into the participants' visual attention distribution during a simulation scenario.

REFERENCES

Ashraf, H., Sodergren, M. H., Merali, N., Mylonas, G., Singh, H., & Darzi, A. (2018). Eye-tracking technology in medical education: A systematic review. *Medical Teacher, 40*(1), 62-69.

Browning, M., Cooper, S., Cant, R., Sparkes, L., Bogossian, F., Williams, B., . . . Black, B. (2016). The use and limits of eye-tracking in high-fidelity clinical scenarios: A pilot study. *International Emergency Nursing, 25*, 43-47.

Fanning, R. M., & Gaba, D. M. (2007). The role of debriefing in simulation-based learning. *Simulation in Healthcare, 2*(2), 115-125.

Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical methods for rates and proportions*: John Wiley & Sons.

Garden, A. L., Le Fevre, D. M., Waddington, H. L., & Weller, J. M. (2015). Debriefing after simulation-based non-technical skill training in healthcare: a systematic review of effective practice. *Anaesth Intensive Care, 43*(3), 300-308.

Grundgeiger, T., Klöffel, C., Mohme, S., Wurmb, T., & Happel, O. (2017). An investigation into the effects of real vs. stimulated cases and level of experience on the distribution of visual attention during induction of general anaesthesia. *Anaesthesia, 72*(5), 624-632.

Henneman, E. A., Cunningham, H., Fisher, D. L., Plotkin, K., Nathanson, B. H., Roche, J. P., . . . Henneman, P. L. (2014). Eye tracking as a debriefing mechanism in the simulated setting improves patient safety practices. *Dimensions of Critical Care Nursing, 33*(3), 129-135.

Henneman, E. A., Marquard, J. L., Fisher, D. L., & Gawlinski, A. R. (2017). Eye Tracking: A Novel Approach for Evaluating and Improving the Safety of Healthcare Processes in the Simulated Setting. *Simulation in Healthcare, 12*(1), 51-56.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*: OUP Oxford.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Kassner, M., Patera, W., & Bulling, A. (2014). *Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction.* Paper presented at the Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication.

Kok, E. M., & Jarodzka, H. (2017). Before your very eyes: the value and limitations of eye tracking in medical education. *Medical Education, 51*(1), 114-122.

Land, M., Mennie, N., & Rusted, J. J. P. (1999). The roles of vision and eye movements in the control of activities of daily living. *28*(11), 1311-1328.

Makary, M. A., & Daniel, M. (2016). Medical error-the third leading cause of death in the US. *BMJ, 353*, i2139.

Marquard, J. L., Henneman, P. L., He, Z., Jo, J., Fisher, D. L., & Henneman, E. A. (2011). Nurses' behaviors and visual scanning patterns may reduce patient identification errors. *Journal of Experimental Psychology: Applied, 17*(3), 247.

O'Meara, P., Munro, G., Williams, B., Cooper, S., Bogossian, F., Ross, L., . . . McClounan, M. (2015). Developing situation awareness amongst nursing and paramedicine students utilizing eye tracking technology and video debriefing techniques: a proof of concept paper. *International Emergency Nursing, 23*(2), 94-99.

R Core Team. (2013). R: A language and environment for statistical computing.

Romero-Ramirez, F. J., Muñoz-Salinas, R., & Medina-Carnicer, R. (2018). Speeded Up Detection of Squared Fiducial Markers. *Image Vision Computing, 78*(August 2018), 38-47.

Sawyer, T., Eppich, W., Brett-Fleegler, M., Grant, V., & Cheng, A. (2016). More than one way to debrief: a critical review of healthcare simulation debriefing methods. *Simulation in Healthcare, 11*(3), 209-217.

Schulz, C. M., Schneider, E., Fritz, L., Vockeroth, J., Hapfelmeier, A., Brandt, T., . . . Schneider, G. (2011). Visual attention of anaesthetists during simulated critical incidents. *British Journal of Anaesthesia, 106*(6), 807-813.

Weller, J. M., Nestel, D., Marshall, S. D., Brooks, P. M., & Conn, J. J. (2012). Simulation in clinical teaching and learning. *The Medical Journal of Australia, 196*(9), 594.