Human Factors and Simulation, Vol. 30, 2022, 101–108
https://doi.org/10.54941/ahfe1001497

Toward the Development of A Realistic, Low-Cost "Gender Retrofit Kit" For Use In Combat Medicine Training

Angela M. Alban[1], Cheryl Coiro[1], Trisha N. Patel[2], Jeffrey M. Beaubien[2], and Mark V. Mazzeo[3]

[1]SIMETRI, 7005 University Boulevard Winter Park, FL 32792, USA
[2]Aptima, Inc., 12 Gill Street, Suite 1400 Woburn, MA 01801, USA
[3]SFC Paul Ray Smith Simulation and Training Technology Center Soldier Effectiveness Directorate U.S. Army Combat Capabilities Development Command Soldier Center 12423 Research Parkway Orlando, FL 32826, USA

ABSTRACT

Bystanders often hesitate when rendering first aid to females, particularly when doing so requires disrobing the patient. The Army has invested heavily in realistic manikins for training combat medicine skills. However, logistical constraints prevent the rapid acquisition and deployment of a large number of female manikins. The purpose of this study was to assess the utility of a low-cost manikin "Gender Retrofit Kit" (GRK). We recruited a sample of 36 Combat Lifesavers and Medics who performed 3 medical procedures (tourniquet application, chest seal application, and needle chest decompression) on both a standard manikin and a GRK-equipped manikin. Linear mixed models were used to test the effects of job role, participant gender, manikin gender, and their interaction on task performance quality and speed. There was only one statistically significant main effect of manikin gender: the GRK-equipped manikin had lower mean task performance scores for the chest seal task, which required disrobing the patient.

Keywords: Medical simulation, Training, Human performance measurement

INTRODUCTION

Previous research suggests that bystanders often hesitate when rendering first aid to injured females, particularly when doing so requires touching the breast area or disrobing the patient (Becker et al., 2019; Leary et al., 2018). Although the effect tends to be magnified among male bystanders, female bystanders also hesitate (Leary et al., 2018). While the exact reasons are not fully understood, there is some evidence to suggest that it may be due to society's sexualization of the female body, concerns about being accused of sexual assault, and the mistaken belief that women's bodies are frail and therefore more prone to injury (Becker et al., 2019; Perman et al., 2020; Perman et al., 2019).

When rendering first aid, every second counts. Every second that cardiopulmonary resuscitation (CPR) is delayed is another second when the brain

is deprived of oxygen. Likewise, every second of uncontrolled hemorrhage is another second closer to death by exsanguination. In addition to the negative effects of treatment delays, research suggests that the quality of first aid rendered to female patients may be inferior to that which is rendered to males. This can occur, for example, when the lifesaver does not fully disrobe the patient and thereby misses one or more wounds that require treatment. Alternatively, the lifesaver may fail to remove the patient's bra, and apply the bandage such that it does not properly seal the wound (Bell, Thompson, Mazzeo, & Pike, 2020).

The Army has invested heavily in the acquisition of realistic patient manikins for training combat medicine skills. However, given logistical constraints associated with their production and purchase, it will be difficult for the Army to rapidly acquire an equal number of female patient manikins and then deploy them to every Army training facility. Therefore, the purpose of this study was to assess the utility a low-cost manikin "Gender Retrofit Kit" (GRK). The GRK included a breast "vest" that is affixed to the manikin's torso, a realistic vagina that is affixed to the manikin's groin, a wig, facial makeup, and step-by-step instructions for feminizing the manikin's facial appearance (Bell et al., 2020; Mazzeo, Sotomayor, Coulter, & Alban, 2018). Both the simulated breast and vagina were made from advanced polymers that were designed to simulate the look and feel of female human skin.

Previous research using the GRK (Bell et al., 2020) has identified significant delays in wound exposure time when treating simulated female casualties. However, that study was based on a small number of participants (n=10), none of whom had either formal training or experience in combat medicine. In addition, that study focused on a single medical procedure – hemorrhage control of a gunshot wound (GSW) to the chest – that required fully disrobing the patient. In this conceptual replication of Bell and colleagues' study, we recruited a larger sample (n=36) of participants who were all trained and experienced in combat medicine. We also expanded the study design to include 3 different simulated wounds: treatment of penetrating trauma to the leg via tourniquet (TQ), treatment of GSW to the chest via the application of a chest seal, and treatment of tension pneumothorax via Needle Chest Decompression (NCD). Of the 3 medical procedures, only the last 2 require disrobing the patient. We hypothesized that if hesitation or other performance issues were to occur, the effects would be localized here.

METHOD

Design

The study used a within-subjects experimental design. Immediately after completing their annual recurrent combat medicine training, potential participants were invited to spend additional time practicing these 3 medical procedures. They were informed that their participation would not affect their already completed course grades.

After providing written informed consent, participants were outfitted with a helmet-mounted camera that was titled slightly downward to capture their first-person view while performing the 3 simulated medical procedures.

In addition, the participants' identities were purposely hidden by having them wear face masks and by removing their name tags.

The participants practiced all 3 procedures using two different manikin types: a standard male manikin and the GRK-outfitted manikin. The order of manikin presentation was randomized and counterbalanced to prevent order effects. To ensure that the study was ecologically valid, the participants were free to choose the specific sequence in which they performed the 3 medical procedures. The only stipulation was that they were required to perform all 3 procedures without assistance, and that they could use only the medical equipment contained in the wounded Soldier's Individual First Aid Kit (IFAK).

As the participants performed the experimental tasks, their performance was digitally recorded using the helmet-mounted video camera. After treating the first simulated patient, the participants were given a 10-minute rest prior to treating the second patient. After treating the second patient, the participants completed 2 post-training questionnaires (1 per manikin) as well as a brief demographic questionnaire. Finally, the participants were debriefed about the study purposes and thanked for their time.

Participants

We recruited a convenience sample of 36 Army Combat Lifesavers (CL) and Combat Medics (CM), all of whom had just completed their annual recurrent training. CLs are the first line of combat medical care providers. After quickly stabilizing an injured Soldier, CLs transfer them to the CM for more advanced care. Army CLs receive 40 hours of initial training in the principles of Tactical Combat Casualty Care (TC3), which focuses on medical equipment, Care Under Fire (CUF), Tactical Field Care (TFC), tactical trauma assessment, and hemorrhage control. As the second line of combat medical care providers, Army CMs provide more advanced care to injured Soldiers and coordinate their evacuation off the battlefield. CMs receive 16 weeks of initial training in TC3, CUF, prolonged field care, and other specialized medical topics. Upon graduation, they are also certified as Emergency Medical Technicians (EMTs).

The sample included a roughly equal mix of CLs (41.6%) and CMs (44.4%)[1]. It was composed largely of males (77.8%), who were less than 29 years of age (75.0%), and were Active Duty Soldiers (97.2%). Among the CL subsample, 50.0% had less than 1 full year of experience in their current role. By comparison, among the CM subsample, 62.5% had greater than 5 years of experience in their current role. Because both CL and CM roles are responsible for preventing massive hemorrhage via tourniquet (TQ), we expected that they would have similar experience with its use. As expected, 75.0% of CLs had prior experience applying TQs, as did 100.0% of the CMs. Because treatment of GSWs via chest seal and treatment of tension pneumothorax via NCD are performed more often by CMs, we expected that the CM subsample would have substantially greater experience than would the CLs.

[1]Three participants listed their role as "Other." Two other participants did not list their role at all, and were coded as "Unknown." All participants were included in the Linear Mixed Model (LMM) analyses that follow.

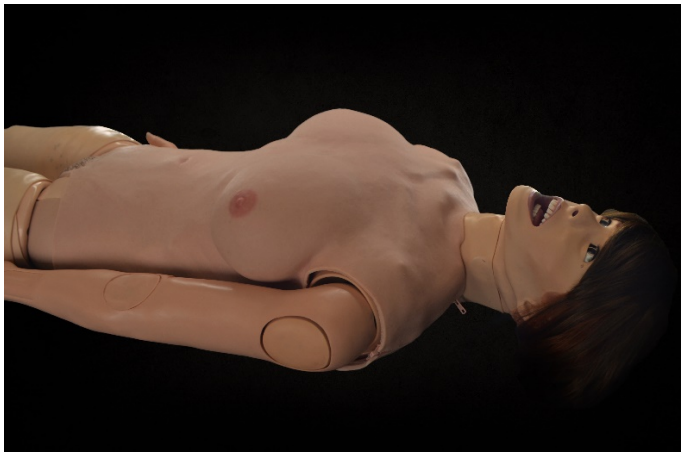

Figure 1: The female Gender Retrofit Kit (KIT) affixed to a standard male manikin. The image depicts the breast "vest," the simulated vagina, the wig, and the make-up that has been applied to feminize the manikin's appearance.

As expected, 87.5% of CMs had experience performing these procedures, while only 37.5% of CLs did.

Apparatus

The male patient manikin was the Laerdal SimMan 3G Trauma. It is a highly-realistic male manikin that contains articulated joints and replaceable limbs to simulate combat trauma. It is connected to a tablet-based computer that is used to modify the patient's physiology (e.g., heart rate, respiration) and responds to learner actions (e.g., CPR) that affect the patient's physiology. The GRK includes a breast "vest" that is affixed to the manikin's torso, a realistic vagina that is affixed to the manikin's groin, a wig, facial makeup, and step-by-step makeup application instructions for feminizing the manikin's appearance (Bell et al., 2020). Both the simulated breast and vagina are made from advanced polymers that were designed to simulate the look and feel of female human skin. A photo of the GRK is depicted in Figure 1.

Measures

Measures of task quality and completion times were generated after-the-fact by a recently-retired Army Delta Force Physician Assistant (PA) who was blind to the participants' identities. Using the video recordings, the PA rated the participants' performance using the Army's standard Individual Critical Task List (ICTL) checklists for TQ application (Army ICTL 081-COM-0048), treatment of GSW via chest seal (Army ICTL 081-COM-0069), and treatment of tension pneumothorax via NCD (Army ICTL 081-68W-0075). Each ICTL contains a checklist of critical task steps and the order in which they must be completed. If a task step was performed correctly, it was a recorded as a "Go" (1). If a task step was not performed or was performed incorrectly, it was recorded as a "No Go" (0). Each participant received a total score for each task by averaging the "Go" scores per manikin. The task

performance scores ranged from 0.00 to 1.00. In addition, task completion times were generated by marking the start and end times of each task, and then computing the difference between them.

The manikins' usability was assessed using a 17-item questionnaire that was developed specifically for this study. Most of the questionnaire items were sufficiently generic such that they could apply to either manikin. Example items include: *The simulated patient looked real*, and *The simulated patient was anatomically correct*. However, 6 items were specifically designed to differentiate the GRK-equipped manikin from the standard male manikin. Example items include: *The simulated patient's chest/breast was realistic*, and *The simulated patient's skin felt realistic*. All of the questionnaire items were rated on a 7-point scale. Many of the questions had anchors ranging from "Strongly Disagree" (1) to "Strongly Agree" (7). Other questions had anchors that ranged from "Not Masculine at All" (1) to "Very Masculine" (7), or "Not Rugged At All" (1) to "Very Rugged" (7), depending on the specific question wording.

RESULTS

Self-Report Usability Questionnaire

Each participant completed the usability questionnaire twice: once for the male manikin and once for the GRK-equipped female manikin. Mean differences between the paired sets of ratings were computed using the dependent *t*-test statistic. As expected, there were no statistically significant mean differences for the 9 manikin-generic questions that inquired about the manikins' design, ruggedness, and suitability as a training tool. Example questions that did not reveal statistically significant differences included: *The simulated patient looked real*, *The simulated patient was anatomically correct*, and *Training on the simulated patient provides meaningful practice in treating a trauma patient*.

However, 5 of the 6 items (83.3%) that were specifically written to differentiate the GRK-equipped manikin from the standard male manikin did produce statistically significant differences, and all were in the hypothesized direction. Example questions that did reveal statistically significant differences include *The simulated patient's chest/breast was realistic* ($t_{34} = 2.76$, $p. < .05$), and *The simulated patient's skin felt realistic* ($t_{34} = 2.14$, $p. < .05$). The only item that did not reveal a significant mean difference was *The simulated patient's face was realistic* ($t_{34} = 1.14$, *ns*).

Self-Report Usability Questionnaire

According to the Army's standard MARCH-E treatment algorithm, combat trauma treatments should proceed in the following order: extremity hemorrhage control via TQ, treatment of GSW via chest seal, and treatment of tension pneumothorax via NCD. Analysis of the task times (measured in minutes) confirmed this pattern. Dependent *t*-tests reveal that the start of TQ application occurred prior to the start of chest seal application ($t_{35}=10.40$, $p. < .05$). Similarly, the start of chest seal application occurred prior to start of the NCD

($t_{29} = 11.50$, $p. < .05$). This pattern of results suggests that participants were performing the 3 medical procedures in the doctrinally correct order.

Next, dependent *t*-tests were computed to compare mean task completion times between the male vs. GRK-equipped female manikins. Contrary to our hypotheses, no statistically significant mean differences were observed for the total completion time, TQ start times, chest seal start times, or NCD start times. Moreover, there were no statistically significant differences as a function of order of presentation (first vs. second). In summary, there was no evidence of hesitation with the female manikin.

Procedure Accuracy

We computed a Linear Mixed Model (LMM) to test for the effects of participant role, manikin gender, and their interaction on task performance accuracy. We observed statistically significant main effects of participant role for all 3 dependent variables: TQ application ($F_{3,32} = 3.95$, $p. < .05$), chest seal application ($F_{3,32} = 8.14$, $p. < .05$) and NCD tasks ($F_{3,32} = 6.79$, $p. < .05$), respectively. In all 3 cases, the differences were in the hypothesized direction, with CMs having higher mean performance than the CLs. We also observed a statistically significant main effect of manikin gender ($F_{1,32} = 5.56$, $p. < .05$) for treatment of the GSW via the chest seal. Moreover, the results were in the hypothesized direction, with the female manikin having lower mean scores than the male manikin. No main effects of manikin gender were observed for the TQ or NCD procedures, nor were any interaction effects observed.

Finally, we computed a LMM to test for the effects of participant gender, manikin gender, and their interaction on task performance accuracy. For all 3 procedures, there were no statistically significant main effects of participant gender, nor were there any significant interaction effects. However, the chest seal application again demonstrated a statistically significant main effect of manikin gender ($F_{1,34} = 5.05$, $p. < .05$). As before, the results were in the hypothesized direction, with the female manikin having a lower mean score than that of the male manikin. No other statistically significant main effects of manikin gender were observed for either the TQ or NCD procedures.

DISCUSSION

Previous research suggests that bystanders often hesitate when providing first aid to injured females, particularly when doing so requires first disrobing the patient and/or touching the chest area. Therefore, the purpose of this study was to evaluate the utility of a low-cost "Gender Retrofit Kit" (GRK) that can be applied to standard male patient manikins for training combat medicine skills on female Soldiers.

The current study was a conceptual replication of Bell and colleague's prior research with the GRK (Bell et al., 2020). The primary differences between the two studies were that the current study included a larger sample size; the inclusion of experienced combat medicine providers rather than inexperienced cadets, and; the requirement to perform medical procedures that required both disrobing the patient (application of the chest seal for treating

the GSW, and performing the NCD to treat the tension pneumothorax) and those that did not (application of the TQ to control extremity hemorrhage).

The results partially, but do not completely, support prior research. Specifically, the usability questionnaire results demonstrate that participants perceived the GRK as realistically representing a female patient. With the exception of the face, other aspects of the GRK-equipped manikin's appearance were significantly rated as more feminine than that of the standard male manikin. Moreover, expert observer ratings revealed a consistent main effect for the chest seal task, which was the first task that required disrobing the patient. As hypothesized, the GRK-equipped female manikin had lower mean performance scores than did the standard male manikin. No other significant main effects were observed with regard to the manikin gender, nor were there any significant interactions with either participant role (CL vs. CM) or participant gender (M vs. F).

Contrary to our hypotheses, we did not find any evidence to support initial hesitation when providing first aid. Not only were the mean task completion times consistent across both manikins, but their order of completion was fully consistent with the Army's MARCH-E combat medicine treatment algorithm.

As with every study, this one has its limitations. Even with the larger size (n=36), the sample size was relatively small in an absolute sense. Moreover, the participants were all trained professionals who had recently completed their recurrent combat medicine training. As a result, they were unlikely to react like an untrained bystander with zero medical experience. This may help explain not only the lack of hesitation, but also the high level of compliance with the standard MARCH-E treatment algorithm.

Future studies are currently being planned to further explore the GRK's utility. These will include a larger and more diverse sample of participants such as Army cadets (who have zero medical experience) as well as combat medicine practitioners who are 6 months away from completing their annual recurrent training. Finally, the planned studies will involve the use of external cameras that will allow us to better decompose treatment time into its component parts: reaction time (the time from when the participant crossed the door threshold until the first touch of the patient simulator) and exposure time (the time from when the participant first touched the patient simulator to when the patient's chest was completely exposed), as was done by Bell and colleagues (2020).

ACKNOWLEDGMENT

This work was supported by the U.S. Army Combat Capabilities Development Command Soldier Center (DEVCOM SC). The authors would also like to thank Mr. Thomas Pingel and Mr. Gerald Bickett at I Corps Medical Range: Joint Base Lewis McChord (JBLM) Medical Simulation and Training Center (MSTC) for their support and making this experiment possible. The views, opinions and/or findings are those of the authors and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REFERENCES

Becker, T. K., Gul, S. S., Cohen, S. A., Maciel, C. B., Baron-Lee, J., Murphy, T. W., . . . Hart, L. (2019). Public perception towards bystander cardiopulmonary resuscitation. *Emergency Medicine Journal, 36*(11), 660–665.

Bell, J. L., Thompson, R., Mazzeo, M., & Pike, W. Y. (2020). Same injury, different outcome? Investigating hesitation while treating female casualties. Paper No. 20410. In *Proceedings of the 2020 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*. Alexandria, VA: National Training and Simulation Association.

Leary, M., Almodovar, A., Buckler, D., Patel, J., Chaudhary, Z. A., Karwat, A., . . . Blewer, A. L. (2018). Comparing bystander response to an unannounced cardiac arrest based on victim gender using an immersive virtual reality system. *Circulation, 138*(Suppl_2), A196-A196.

Mazzeo, M., Sotomayor, T. M., Coulter, J. N., & Alban, A. M. (2018). Development & assessment of a human patient simulator gender retrofit kit. Paper No. 18266. In *Proceedings of the 2018 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*. Alexandria, VA: National Training and Simulation Association.

Perman, S. M., Shelton, S., Rice, J. D., Knoepke, C., Rappaport, K., Adelgais, K. M., . . . Daugherty, S. L. (2020). Understanding why differences in the provision of bystander CPR exist for women versus men: Does the sex of the rescuer matter? *Circulation, 142*(Suppl_4), A139-A139.

Perman, S. M., Shelton, S. K., Knoepke, C., Rappaport, K., Matlock, D. D., Adelgais, K., . . . Daugherty, S. L. (2019). Public perceptions on why women receive less bystander cardiopulmonary resuscitation than men in out-of-hospital cardiac arrest. *Circulation, 139*(8), 1060-1068.