Quantitative Assessment of Combat Casualty Skills

Christine Allen, Ph.D. & Mark Mazzeo
Army Research Laboratory (ARL) Human Research and Engineering Directorate (HRED)
SFC Paul Ray Smith Simulation and Training Technology Center (STTC)
christine.m.allen52.civ@mail.mil & mark.v.mazzeo.civ@mail.mil

Brian Goldiez, Ph.D. & Amanda Romeu
University of Central Florida (UCF) Institute for Simulation and Training (IST)
bgoldiez@ist.ucf.edu & aromeu@ist.ucf.edu

Thomas Pingel
Joint Base Lewis McCord (JBLM) Medical Simulation Training Center (MSTC)
thomas.a.pingel.civ@mail.mil

ABSTRACT

Evaluating proficiency in combat casualty training includes the assessment of hands-on training with mannequins through instructor observation. The evaluation process can suffer due to the subjective nature of the assessment: differences between instructor rating schemas, student to instructor ratios, and time to observe individual student performance. Because combat casualty care requires timely and accurate assessment for medical interventions, evaluators can look at the trainees' physical actions (e.g., hand motion) to assess proficiency, as seen in suturing literature. The Lempel Ziv (LZ) complexity index is then used to assess proficiency. The LZ algorithm reduces complex strings of data (i.e., hand motion) to a string of 1's and 0's. The string is then broken into small "unique" strings that are grouped together. The pattern formed is a measure of performance with more complex patterns per unit of time indicating expertise.

Expanding the current state of the art, experimentation occurs using several different precision tracking devices that are unobtrusive and require limited setup. During this effort, student hand motion is tracked and digitally stored as participants complete multiple tasks part of a cricothyroidotomy (emergency airway procedure in the neck). Motion data is subsequently processed using an algorithm adapted for text compression (LZ algorithm).

Data has been gathered from nearly 100 military combat medic trainees at Joint Base Lewis McChord (JBLM) Medical Simulation Training Center (MSTC). Participant hand acceleration data from an emergency surgical cricothyroidotomy reveals a statistically significant difference in ability among different expertise levels. The higher the LZ score and self reported expertise level, the better the participant performed. The results show that when presented with demographic and video performance-based data, it is possible to gauge experience using LZ scores.

ABOUT THE AUTHORS

Christine Allen, Ph.D., is a Science and Technology Manager with ARL-HRED STTC. She oversees medical simulation research endeavors, transitioned multiple technologies, and authored numerous papers and presentations within the medical simulation field. She received her Doctorate in Modeling and Simulation from the University of Central Florida and is a Certified Modeling and Simulation Professional.

Brian Goldiez, Ph.D., is the Deputy Director and Research Associate Professor at the University of Central Florida's Institute for Simulation and Training (UCF IST). He has conducted a number of research projects on interoperability, computer graphics, and augmented reality. Recent work in the medical area involves data interchange protocols, return on investment methodologies, and live tissue versus simulation studies.

Amanda Romeu, is a Research Technician at the UCF IST. She has worked on multiple projects in kinesic cue and combat medic training. She received her Bachelor of Science in 2013 from UCF and begins medical school at the Philadelphia College of Osteopathic Medicine in the fall of 2014.

Thomas A. Pingel, is the Course Director for the Medical Simulation Training Center, Joint Base Lewis-McChord (JBLM), Washington. He completed 21 years of service in the U.S. Army as a Combat Medic, LVN, Practical Nurse Course Instructor, while concluding his career as the Non-Commissioned Officer in Charge of the Practical Nurse Course at JBLM. He has a B.S. in Healthcare Management from Franklin University, completed post graduate studies at the Naval Post Graduate School, and is a Certified Healthcare Modeling and Simulation Professional.

Mark Mazzeo, is an Engineering Technician at ARL-HRED STTC. He supports the medical branch through designing experiments and collecting and analyzing data. Mr. Mazzeo holds a B.S. in Industrial Engineering from the University of Central Florida, and is currently pursuing a M.S. in Industrial Engineering through UCF.

Quantitative Assessment of Combat Casualty Skills

Christine Allen, Ph.D. & Mark Mazzeo
Army Research Laboratory (ARL) Human Research and Engineering Directorate (HRED)
SFC Paul Ray Smith Simulation and Training Technology Center (STTC)
christine.m.allen52.civ@mail.mil & mark.v.mazzeo.civ@mail.mil

Brian Goldiez, Ph.D. & Amanda Romeu
University of Central Florida (UCF)
Institute for Simulation and Training (IST)
bgoldiez@ist.ucf.edu & aromeu@ist.ucf.edu

Thomas Pingel
Joint Base Lewis McCord (JBLM)
Medical Simulation Training Center (MSTC)
thomas.a.pingel.civ@mail.mil

BACKGROUND

Medical simulators, such as mannequins, attempt to mimic the human body both physically and physiologically. They are used to practice interventions such as an emergency airway or surgical cricothyroidotomy. Military first responders typically perform the cricothyroidotomy while deployed in austere conditions with the threat of enemy fire, while civilian first responders typically have faster access to hospitals and emergency medical staff, a sterile environment, and additional treatment options (some less invasive in nature). Obtaining an airway quickly and accurately is extremely important, especially in the military setting, where the environment may not be ideal. Ensuring operational readiness is essential to military first responders (e.g., combat medics), who many times are the first to treat casualties. This study addresses technologies that quantify hand motion, speed, and help identify first responder levels of proficiency. Adding quantifiable methods of scoring may assist instructors in assessing trainee performance more effectively over modalities such as simulators and standardized patients (e.g., people role-playing as injured casualties). These quantitative methods may also aid learning and evaluation not only within the medical and military communities, but the research community at large.

OVERVIEW

The focus of this study is to observe time-sensitive emergency surgical procedures (such as the cricothyroidotomy) that combat medics and first responders perform and determine a trainee's level of expertise or skill mastery using quantitative methods. Following traditional didactic (i.e., classroom) training, instructors teach hands-on skills for combat casualty care. They must observe, evaluate, and provide input for students to reach proficiency. Although objective checklists are used to determine proficiency, there still remains an element of subjective evaluation.

This paper will demonstrate how objective measures can be combined with existing subjective observational methods that traditionally designate an expert, based on task completion. It is desired to establish how quantitative-based technologies can be more effective and precise while aiding evaluation. As a tool, quantitative scoring may assist instructors in the educational and evaluation processes, assist with efficiency as budgets decrease, and reduce instructor demand as instructor to student ratios increase. It may also allow instructors to discern differences between various simulators, while assisting with educational effectiveness, time management, and ensuring combat medic and first responder operational readiness.

Cricothyroidotomy Surgical Procedure

When a person is severely injured in combat, normal respiratory orifices (i.e., mouth and nose) can be blocked causing an inability to establish an airway via nasal pharyngeal intubation (nose), endotracheal intubation (mouth), or Bag Valve Mask (BVM) ventilation (nose and mouth). In these cases, an emergency cricothyroidotomy is performed to surgically establish an airway. A cricothyroidotomy is often a last resort to circumvent obstruction of a patient's airway when the ability to restore adequate oxygenation is not possible by any other means. The cricothyroidotomy involves the insertion of an endotracheal tube through the cricothryoid membrane (in the neck under the Adam's apple), creating a low-pressure ventilation system. The procedure can be very dangerous and

should only be performed by those who can rapidly and effectively locate the anatomical landmarks to successfully perform the associated task. Therefore, this procedure lends itself well for the speed (three minutes to complete) and accuracy (proper placement in the anatomy) measures.

The combat emergency surgical cricothyroidotomy procedure requires a variety of equipment contained within the combat medi-pack: scalpel, endotracheal tube, tracheal hook, tracheal dilator, inflation syringe, and gauze as seen in Figure 1, as well as alcohol wipes and gloves. For the purposes of this study, the cricothyroidotomy procedure was taken from the Combat Medic Card Game (Smith, 2010) playing cards envisioned by Dr. Roger Smith and developed by the University of Central Florida's Institute for Simulation and Training (UCF IST) in concert with the US Army Research Laboratory-Human Research and Engineering Directorate, Simulation and Training Technology Center (ARL-HRED STTC) who also funded the effort and are listed below.

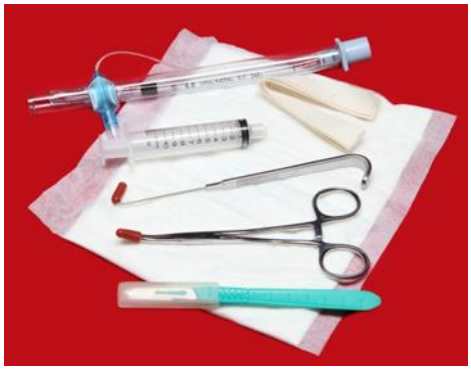

Figure 1 Combat medi-pack for a cricothyroidotomy (from top to bottom): endotracheal tube, inflation syringe, tracheal hook, tracheal dilator, scalpel on top of gauze

Combat Medic Cards Cricothyroidotomy Steps:

- Combat medic identifies if the casualty has a full airway obstruction
- If full obstruction is found, the medic obtains the necessary materials and equipment (endotracheal tube, inflation syringe, tracheal hook, tracheal dilator, scalpel, gauze, alcohol wipe, and gloves)
- Gloves are worn to avoid any further contamination to the injured body
- The medic locates the cricothryoid membrane, cleans the area, and then stabilizes the membrane in preparation for the incision
- A 1-1/2" vertical incision is made over the cricothryoid membrane
- After the initial incision, a secondary incision is made to cut through the cricothryoid membrane in the horizontal direction
- The cricoid cartilage should be visible and moved away from the trachea by the use of a tracheal hook or similar instrument
- Once the trachea is exposed, the endotracheal tube can be carefully inserted into the trachea
- Following endotracheal tube insertion, the inflation syringe is placed onto a small tube which comes off of the endotracheal tube
- Using the inflation syringe, fill 10 cc of air in the endotracheal ballooning device to open the airway
- Validate air exchange
- The endotracheal tube is secured in place with gauze and medical tape
- The final step, attaching a BVM (e.g., Ambu Bag) was not part of this study

In a wartime environment, a combat medic must be able to successfully complete all of the steps listed above within three minutes (Smith, 2011; Katos & Goldenberg, 2007). The full procedure is complex with respect to hand-eye coordination, delicate anatomy of a human casualty's neck, and military environment with a variety of sensory distracters and stressors. This type of situation creates a tradeoff between speed and accuracy as a medic must save

the casualty while also being mindful of safety in a potentially dangerous environment. This creates a need for objective measures of speed and accuracy making the cricothyroidotomy procedure far more significant as instructors try to maximize both speed and accuracy measures.

Speed Accuracy Studies

Many medical procedures are time sensitive and require precision and accuracy. Simulators should reinforce these general skills to ensure effective and quality training. Whether in the military environment or in the civilian trauma setting, medical interventions often require a tradeoff between speed and accuracy components. Fitts (1954) illustrates the tradeoff between precision in time sensitive situations, revealing that with increased accuracy more time must be taken, but if rushed, accuracy decreases.

More recently, Watson (2013) discovered a similar relationship between speed and accuracy and known levels of expertise in suturing. Students who completed these tasks developed a balance between speed and accuracy as they trained over time. Watson also addressed the tradeoff between speed and accuracy in medical training situations for hand motion and the resulting efficiency of the motion. The eventual efficiency of motion was introduced as an important feature in the measurement of medical tasks and further supported by a suturing and knot tying task study using the Imperial College Surgical Assessment Device (ICSAD), (Bann, Khanm & Darzi, 2003).

In the past, expert evaluators could only measure mastery of surgical or other emergency procedures through qualitative observation. This method has been, and still is, effective in most medical training settings. However, course growth and diminishing resources place additional demands upon instructors to teach and assess students efficiently and effectively. In conjunction with simulation training, hand tracking can play an important role in discerning mastery of a skill while potentially decreasing instructor workload.

The Lempel-Ziv (LZ) Complexity Algorithm

This study measured the relationship between time, accuracy, and motion in a cricothyroidotomy procedure. The data output from hand motion was translated into a score related to expertise. The Lempel-Ziv complexity algorithm, specifically LZ76 (Ziv and Lempel, 1976) or LZ for this study, is mathematical and continuous when used as a predictor to compress a finite string of data into a binary alphabet. The main idea is to parse data into distinct phrases, which can be converted into scores based on the alphabet. At each time step, accelerations were recorded by the Q Sensor (discussed in the methodology section), converted into a Root Mean Square (RMS), and then into binary strings. The following method was applied: if the acceleration was higher or lower than one standard deviation of the mean, a score of 1 was assigned to the acceleration, otherwise it was assigned zero. In this manner, each participant's data set generated a string of binary digits used to compute the number of "unique" words. The uniqueness is synonymous with complexity; higher complexity would imply rapid, large accelerations and Watson (2013) hypothesized that the pattern previously described is indicative of a higher level of expertise. Finally, the complexity was normalized into a number between [0, 1] through division by the equation $C = \frac{n}{log_2(n)}$ where n denotes the length of the string. The authors set out to show that application of LZ scores derived from the previously listed algorithm are an effective, objective method of assessing trainee expertise when used alongside traditional, qualitative methods.

CRICOTHROTOMY SPEED ACCURACY METHODOLOGY

Quantitative data were collected using several types of commercial off-the-shelf, computer-connected devices that were adapted for this study. These devices measured acceleration and motion data of the hand. Video capture was also used to correlate the hand actions during the procedure with the motion data to define specific tasks and determine procedure precision.

Participants and Setting

This study occurred at the JBLM MSTC. A total of 91 combat medics who were obtaining combat medic certification for the first time or recertifying, completed the study across six combat medic training courses between June and December 2013. The 91 medics included 76 men and 15 women between the ages of 20 to 52, with 32 as the average age. All participants received didactic training on combat casualty care with the same instructors, followed by participation in the study. While 30 participants had never performed the cricothyroidotomy procedure in the field or military environments, 61 had some level of experience. Participation in the study was voluntary and all data were anonymous and not linked to performance in the combat medic course.

The study was conducted in a lab training room using two medical simulators. The Multiple Amputation Trauma Trainer (MATT)® upper torso and the Medical Education Technologies, Inc. (METI, now CAE Healthcare) Human Patient Simulator were each placed on standard military litters (a portable stretcher on a stand), as shown in Figure 2.The necks of each simulator received 2" 3M Red Vinyl tape to simulate the cricothyroid membrane, followed by surgical self-adhesive tape to simulate the skin over the neck. This setup is the standard training modality used and recommended by the experts at JBLM.

A laptop, camera, and timer (placed within the camera's view) were set up on the left side of the simulator, while the SoftKinetic DepthSense® motion tracking device was set above the mannequin's head, (see Figure 2 and equipment description below). The DepthSense® was mounted on a flexible arm to overlook the simulator's neck and participant's hands and was directly connected to the computer via USB cable. Each participant also wore two Q Sensors (another type of motion sensing device), one on each hand, as seen in Figure 3. The experimenter stood behind the laptop to observe both the participant and the quantitative data collection.

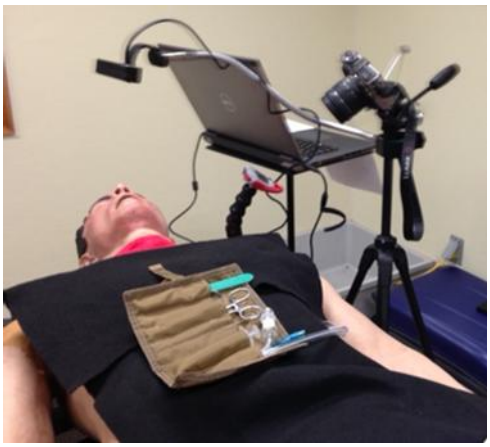

Figure 2 Lab room setup

Equipment

Hand motion tracking and other monitoring devices in simulations generally tend to be obtrusive or distracting to the user. The devices used for this study were chosen because they do not encumber the motion of the trainee nor do they add bulk or distraction.

Stopwatch

A commercially available stopwatch recorded the time to perform the cricothyroidotomy. The stopwatch was synchronized with the time on the computer and the Q Sensor. The timer allowed the cricothyroidotomy procedure to be divided into tasks to determine if a participant exceeded the total survival time, three minutes, which is the lower bound limit to establish the flow of oxygen (Smith, 2011; Katos & Goldenberg, 2007).

Affectiva Q Sensor 2.0
The Affectiva Q Sensor 2.0 or Q Sensor measures galvanic skin response and acceleration of a moving hand. However, only the acceleration function was used in this study. The Q Sensor is worn like a wrist watch. As seen in Figure 3, it is extremely small (only 2.16" by 1.5") and unobtrusive, so it does not interfere with the hand motions (Q Sensor, 2014) used in the cricothyroidotomy. The Q Sensor continuously (32Hz) collected data via an internal 3-axis accelerometer and stored the data on a memory card. These data were downloaded to the computer following conclusion of each procedure.

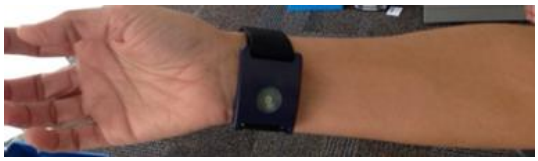

Figure 3 Affectiva Q Sensor 2.0

Equipment Not Used in the Scoring Process
The equipment described below was used to gather some participant data, but ultimately the data were not reliable and therefore not analyzed. The SoftKinetic DepthSense® (DS) 325 is a time-of-flight motion sensor technology that reads the shape and hand displacement (SoftKinetic, 2014). The SoftKinetic iisu® 3.6 Middleware is a software program that pairs with the SoftKinetic Depth Sense cameras supporting both full body and hand tracking (SoftKinetic, 2014). The LEAP Motion Controller, or LEAP, is a device that uses infrared light to track hand and finger motion.

Data Collection Materials

Procedure
Once participants completed didactic cricothyroidotomy training, a group entered the lab training room. The participants were briefed and instructed to read the informed consent and fill out a demographic questionnaire self-reporting expertise. These demographics were used to establish four parameters related to cricothyroidotomy performance that were utilized to establish LZ scores and were reported in the results section.

A small whiteboard containing an anonymous participant number (established by the participant) and simulator type was placed in front of the camera to signify the start of a new participant. Each cricothyroidotomy procedure was performed individually. Participants were given two Q Sensors one worn on each wrist. The Q sensors were manually initiated upon starting the cricothyroidotomy and stopped following procedure conclusion when the participant tapped the simulator's head to communicate completion. Following cricothyroidotomy completion, participants answered a post-questionnaire consisting of nine questions using a 5-point Likert scale with 1 = Strongly Agree, 2 = Agree, 3 = Don't Know, 4 = Disagree, and 5 = Strongly Disagree and three short answer questions. The results of the post-questionnaire are shown in Table 4 of the qualitative results section.

Video Data
Participant hand motion was recorded as they performed the cricothyroidotomy. The video recorded the neck area of the simulator and the hands of the participant; no personal identifiers were captured. Each video was assigned a unique number designated by the participant. The videos were used in conjunction with the Q Sensor to obtain cricothyroidotomy procedure data.

RESULTS

Data from 77 of the 91 participants were used in data analysis. There were 14 participants whose data were not used due to technological complications or missing demographic information. Some complications came from the Q Sensor, which did not pick up the data, or were not turned on for data collection. Other complications came from the video recordings, which did not have the stopwatch in the field of vision, thus not allowing video correlation with Q sensors.

Qualitative Data Collection Parameter Development

Using the self-reported demographic questionnaire, the authors selected four demographic parameters to best represent expertise. The four parameters (listed below) leveraged a study of anesthesiologists who performed a cricothyroidotomy on mannequins (John, Suri, Hillerman, & Mendonca, 2007) using amount of experience as a predictor of expertise.

These parameters were:

- Number of times a cricothyroidotomy was performed
- Medium (e.g., simulator, animal, or human) on which a cricothyroidotomy was performed
- Last time a cricothyroidotomy was performed
- Number of years as a combat medic

Furthermore, these demographic parameters were used to define experiential ranking levels in conjunction with the individual participant videos, creating three categories or expertise levels: novice, intermediate, and expert. It was hypothesized that some of the above parameters would have a significant effect on the expertise levels and LZ scores. To rank the participants, the videos were reviewed to determine if medics performed the procedure in the proper sequence within three minutes (Smith, 2011; Katos & Goldenberg, 2007) to "save" their patient. In the event that the medic did not complete the cricothyroidotomy within the three minute time frame, they were placed in a lower expertise level.

Response variables

In Table 1, the "Item" corresponds to the information on the demographic survey, while the "Value" was the numeric representation of the demographic answers used to determine expertise levels. For the purpose of this study, a value of 1 denotes a novice level, 2 intermediate, and 3 and 4 an expert level. Note that values of 3 and 4 were combined into one expert level due to the similarities between the two groups.

Table 1 Demographic parameter value assignments for cricothyroidotomy

Demographic Parameters							
Number of times the cricothyroidotomy was performed		**Type of medium the cricothyroidotomy was performed**		**Last time the cricothyroidotomy was performed**		**Length of time as a combat medic**	
Item	**Value**	**Item**	**Value**	**Item**	**Value**	**Item**	**Value**
None	1 (novice)	None	1 (novice)	Greater than 3 years	1 (novice)	None	1 (novice)
1 to 5 times	2 (intermediate)	Simulator	2 (intermediate)	1 to 3 years	2 (intermediate)	1 day to 6 months	2 (intermediate)
6 to 15 times	3 (expert)	Simulator + other	3 (expert)	3 to 12 months	3 (expert)	6 to 12 months	3 (expert)
Greater than 15 times	4 (expert)	Human and/or animal	4 (expert)	Less than 3 months	4 (expert)	Greater than 1 year	4 (expert)

Quantitative Data Collection

Lempel-Ziv score

The Q Sensor was used to gather acceleration data analyzing the individual tasks for each participant. Thousands of values were gathered for each participant's hands during the cricothyroidotomy procedure. Using MATLAB (assists with numerical calculation and programming) and GNU Octave (assists with numerical calculation) the RMS acceleration values were converted into LZ complexity numbers (Thai, 2012) for each hand and for a composite of both hands. The steps for computing LZ scores from the acceleration data are as follows:

1. Compute the RMS at each time step and each hand by combining the accelerations along each axis
2. Compute the mean and standard deviation over the entire time period for each hand
3. Compute a composite mean for both hands based on the RMS and standard deviation
4. For values within one standard deviation of the mean, LZ = 1
5. For values beyond one standard deviation of the mean, LZ = 0

Logistic Regression Models

The four parameters listed above under the qualitative data collection section were processed along with the LZ scores using a statistical program. This program placed the LZ scores into a logistic regression model giving the overall assessment of novice-expert designations. This analysis of motion data were similar to the scoring method used by Watson (2013). Equation 1 describes probabilistically whether each of the four demographic parameters is an indicator of expertise. Details on the logistic equation and methodology can be found in Chambers 2012.

$$P(EXPERTISE = i \mid EXPERIENCE, LZ) = \frac{\exp\left(\beta_{0_i} + \beta_{1_i} \times LZ + \beta_{2_i} \times EXPERIENCE\right)}{1 + \sum_{i_0}^{i_n} \exp\left(\beta_{0_i} + \beta_{1_i} \times LZ + \beta_{2_i} \times EXPERIENCE\right)} \quad (1)$$

Qualitative Output

The criterion used to initially designate an expert was more than five previous experiences performing the cricothyroidotomy, exposure to both the simulator and human in simulation-based scenarios or while deployed, and performing the cricothyroidotomy within the last three to six months (as displayed in Table 1). For the novice selection, those who had never performed the procedure and those with low levels of combat medic training were sufficient criteria, while the intermediate level fell in between both criteria. With these scores, participants were sorted as seen in Table 2.

Table 2 Expertise designation numbers

Explanation	Expertise Levels	Number of Study Participants
Novice	1	29
Intermediate	2	21
Expert	3	27

Quantitative Output

The LZ scores below in Table 3 illustrate the expert group with a mean LZ complexity score of 0.5969, the intermediate group with a score of 0.5741, and the novice group with a score of 0.5254. The standard deviations of these LZ scores for expert, intermediate and novice were 0.1132, 0.1084, and 0.1040, respectively. The expert group had a range of 0.4291 to 0.8852, the intermediate group with a range of 0.4071 to 0.6903, and the novice group with a range of 0.2102 to 0.6721. It is important to note that the LZ scores ranged from 0 to 1, and although the ranges overlap, the combination of demographic information was used to influence novice to expert rating levels. Again, all *p*-values were less than 0.05.

Table 3 Relationship between EXPERTISE and LZ Scores

Statistical Output	Novice	Intermediate	Expert
Range (lower and upper bounds)	0.2102 to 0.6721	0.4071 to 0.6903	0.4291 to 0.8852
Standard deviation	0.1040	0.1084	0.1132
Average	0.5254	0.5741	0.5969

Next, a multinomial logistic regression model (Chambers, 2012) was fitted to the data comprising LZ scores. Expertise (i.e., novice, intermediate, and expert) and experience level (number of times participant completed

procedure) and regression equations were obtained with the reference case being novice (EXPERIENCE level 1). Again, all *p*-values were less than 0.05.

Based on the logistic regression, only the LZ scores and demographic category (number of times the participant performed the cricothyroidotomy) were statistically significant in the expert-novice determination supporting only a portion of the hypothesis. The other three parameters were not statistically significant. The authors postulate that this is because these variables were highly correlated with the LZ scores, and thus added no further information to explain the response. Of these three demographic parameter variables, the next closest to being statistically significant was years of experience.

As shown in Table 3, the average LZ score was greater than 0.5. Combined with the observation that higher LZ scores indicate higher experience levels, an LZ score below 0.5 should represent a low probability of high experience level. This implies that conditional probabilities computed with LZ scores below 0.5 (for a given experience level) should give a low probability of higher expertise. This was tested using the logistic regression equation below:

$$\begin{aligned} P(EXPERTISE = 2 | EXPERIENCE = 2, LZ = 0.2) &= 0.05836966 \\ P(EXPERTISE = 2 | EXPERIENCE = 2, LZ = 0.7) &= 0.5351121 \\ P(EXPERTISE = 0 | EXPERIENCE = 3, LZ = 0.7) &= 0.002143988 \end{aligned} \quad (2)$$

Participant Reaction

Following completion of the surgical cricothyroidotomy, medics completed a post-questionnaire survey. Nine questions were answered using a 5-point Likert scale with 1 = Strongly Agree, 2 = Agree, 3 = Don't Know, 4 = Disagree, and 5 = Strongly Disagree. The average responses are displayed below in Table 4. Most participants agreed that the cricothyroidotomy procedure was similar to the normal training procedure and the experimental setup did not restrict performance. This confirms that the study did not appear to differ or impede with normal training.

Table 4 Post-Questionnaire Survey

Questions	**Average Response based on 5-point Likert**
The hands-on procedure was similar to the procedure learned in the classroom.	1.67
The experimental set up restricted me while performing the task.	3.89
The instructions for using the mannequin were confusing.	4.26
My personal performance today was similar to my prior performance of this procedure.	2.70
Overall, I felt comfortable throughout the procedure.	2.02
I was motivated to perform the procedure to the best of my ability.	1.61
I was confident in my ability to perform this medical procedure **prior** to participating in this study.	2.02
I was confident in my ability to perform this medical procedure **during** my participation in this study.	2.05
I was nervous during my participation in this study.	3.16

DISCUSSION AND SUMMARY

The hand motion pattern data from an emergency surgical cricothyroidotomy requiring speed and accuracy reveals a statistically significant difference in ability among different expertise levels (i.e., novice, intermediate, and expert); the higher the LZ score and expertise level, the better the participant performed. Higher experience (number of times participant completed the cricothyroidotomy procedure) appears to indicate a higher total LZ score which is

consistent with Watson's (2013) findings. These findings add to the body of knowledge by introducing procedures with a number of tasks without a requirement for a test/retest. Additionally, the equipment in no way impedes participant performance. The results, as hypothesized, demonstrate that when presented with demographic and video performance-based data, it is possible to ascertain experience level using LZ scores.

Data were organized by both subjective and objective measures. The participant-subjective relationship was based on the combined categories of demographics and video, while the objective relationship was based on the LZ score. This allowed for a comparison between the expected attributes a novice or expert would depict versus what the data actually revealed based on accuracy and time constraints. The video, demographics, and LZ data were then used to organize the participants into an expert, intermediate or novice level.

Though the demographic information was used to make the expertise classifications, 3 of the 4 demographic parameters (Table 1) were not statistically significant. These parameters were the last time the cricothyroidotomy procedure was performed, the type of medium (e.g., animal, human, simulator) on which the procedure was performed, and length of time as a medic. However, the number of times a participant performed the procedure was significant. Possible reasons that the three parameters were not significant were that medium (e.g., simulator, human, or animal) may not matter for novice participants or that there was not enough data to truly determine the medium importance. There also appears to be a relationship between the number of times a cricothyroidotomy has been performed or gap in time the medic last performed the cricothyroidotomy. Some medics may have more experience, but there may have been a greater time period between the last time they performed the cricothyrotomy affecting significance. Although not significant, the three parameters should be retained and used when the LZ scores and experience levels are not aligned.

When evaluating LZ scores and the number of times the procedure was performed, it is possible to place military combat medics in appropriate categories. Those with a higher LZ score, better than their pre-designated expertise level (e.g., a LZ score of expert and pre-designation of novice) upon class commencement could be moved into more difficult tasks earlier in the training, almost taking an adaptive learning approach.

However, if the LZ score is statistically lower than the pre-designated category (e.g., a novice LZ score and pre-designation of an expert) further investigation should be conducted, as this impacts the instructor and course outline. Two additional items should be investigated including other demographic factors that might impact results; such as 1) the number of years since the medic last performed the procedure and 2) the need for additional training to bring the participant back to their pre-designation of expert due to skill decay from previous evaluations. As more people are evaluated with this LZ method, the brackets for scoring should be adjusted or decomposed to compensate for parameter fluctuations.

LIMITATIONS AND LESSONS LEARNED

This study encountered some obstacles during execution. The first obstacle was to identify which technology was least intrusive and operated the best to obtain motion data. The DepthSense® 325 was intended to capture displacement of fingers and hands as a measure of accuracy. However, this could not be completed because the device had problems with various actions, such as differentiating human hands from mannequin skin, differentiating between both left and right hands, and occasionally dropping data values or collecting data at random times. This rendered the data unusable because of its inconsistencies. The LEAP technology was also not useful because it could not record above the participants hands or capture hand motion in any position unless the device was upright.

FUTURE RESEARCH

As was discovered in previous research (Watson, 2013) and the research presented in this paper, a lower LZ score suggests a lower expertise level while higher LZ scores suggests higher expertise, a relationship that should be statistically tested in future research with an ever increasing population building an ongoing database. Additional research that focuses on difficult tasks decomposed within the cricothyroidotomy procedure or other medical

procedures can be useful in determining expertise levels, thus assisting instructors in curriculum development and evaluation. LZ scores and accompanying demographics should also be expanded to include a repository of procedures across varying medical interventions while increasing the database of participant performance.

Additionally, since the SoftKinetic and LEAP technologies were not useful for this experiment, only the Q Sensors were employed. However, the Q Sensors are no longer being produced. This means that another technology should be researched (such as some of the commercially available fitness sensors) or developed that allows the capture of hand movement speed and acceleration to establish performance rating levels.

ACKNOWLEDGEMENTS

The authors would like to acknowledge ARL-HRED STTC for funding this experiment and thank the instructors and participants from JBLM MSTC. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Department of Defense or U.S. Army. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES

Bann, S., Khan, M., & Darzi, A. (2003). Measurement of surgical dexterity using motion analysis of simple bench tasks. World Journal of Surgery, 27(4), 390-394.

Chambers, J. (2012). The R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47(6), 381-391. doi: 10.1037/h0055392

Hosmer, D.W., & Lemeshow, S. (2000). Applied Logistic Regression. Canada: John Wiely and Son, Inc.

John, B., Suri, I., Hillermann, C., & Mendonca, C. (2007). Comparison of cricothyroidotomy on manikin vs. simulator: a randomized cross-over study. Anaesthesia, 62(10), 1029–1032. doi: 10.1111/j.1365-2044.2007.05219.x

Katos, M.G., & Goldenberg, D. (2007). Emergency cricothyrotomy. *Operative Techniques in Otolaryngology*, 18(2), 110-114. doi: 10.1016/j.otot.2007.05.002

Q Sensor. (2014). Q Sensor powered by Affectiva. Retrieved January, 2014 from Q Sensor Website: http://www.qsensortech.com/features/#measure_movement

Smith, N., (2011). Anoxic Brain Damage. NYU Langone Medical Center. Retrieved from http://www.med.nyu.edu/content?ChunkIID=96472

Smith, P.R. (2010) Combat Lifesaver Playing Cards. Institute from Simulation and Training, #CLS34 -2010 - #CLS52-2010. Orlando, FL

SoftKinetic. (2014). Depth Sense Cameras. Retrieved January, 2014 from SoftKinetic Website: http://www.softkinetic.com/en-us/products/depthsensecameras.aspx

Thai, Q. (2012). Calc_lz_completity. Retrieved Feburary, 2014 from MATLAB Central Website: http://www.mathworks.com/matlabcentral/fileexchange/38211-calclzcomplexity/content/calc_lz_complexity.m

Watson, R. (2013). Quantification of Surgical Technology Using An Inertial Measurement Unit. Society for Simulation and Healthcare, 8(3), 162-165.