

2013

Differences between trained and naive participants on triage classification using gaze tracking

<https://hdl.handle.net/2144/12154>

"Downloaded from OpenBU. Boston University's institutional repository."

BOSTON UNIVERSITY
SCHOOL OF MEDICINE

Thesis

**DIFFERENCES BETWEEN TRAINED AND NAÏVE PARTICIPANTS ON
TRIAGE CLASSIFICATION USING GAZE TRACKING**

by

SARAH ERICA LOTH

B.A., Saint Anselm College 2011

Submitted in partial fulfillment of the
requirements for the degree of
Master of Arts

2013

Approved by

First Reader

Kevin "Kip" Thomas, Ph.D., M.B.A.
Director, Healthcare Emergency Management Program

Second Reader

Ann Zumwalt, Ph.D.
Assistant Professor, Gross Anatomy

ACKNOWLEDGEMENTS

First, I would like to thank Kip for guiding me throughout this entire process. You always believed in me even when I doubted myself—Thank you for standing by me and guiding me when I needed the extra support. I would also like to thank everybody who supported me in the lab—Dr. Zumwalt, Dr. Romney, Amanda, Lev, Lee, Kitzia and many more!

I also would like to thank my parents. Mom, Dad, you both have put an inordinate amount of effort into ensuring I succeed in all avenues of life, for that I am utterly grateful. I would not be in this position without either of you...through late nights, early mornings, snowstorms, bumper-to-bumper traffic, you name it, we've conquered it.

Lastly, I want to thank my friends for standing by me and supporting me every step of the way. I am proud of each and every one of you for all of your accomplishments.

DIFFERENCES BETWEEN TRAINED AND NAÏVE PARTICIPANTS ON TRIAGE CLASSIFICATION USING GAZE TRACKING

SARAH ERICA LOTH

Boston University School of Medicine, 2013

Major Professor: Dr. Kevin Thomas, Ph.D., M.B.A., Director, Healthcare
Emergency Management Program

ABSTRACT

The Simple Triage And Rapid Treatment (START) program is a well-known triage-training program that helps first responders correctly sort victims into proper triage categories. START recognizes four categories: 1) Green/Minor: aka walking wounded (able to walk about with minor injuries that are not immediate or life threatening); 2) Yellow/Delayed: victim presents with injuries that needs medical attention after the red/immediate victims are treated; 3) Red/Immediate: severely injured, needs immediate help and gets priority; 4) Black/Expectant: the victim is either deceased or there is little that can be done to save them. Using the START triage program as a model, images of injured victims were used to simulate a typical field triage experience. While viewing the images, participant's gaze patterns were recorded. Two subject groups were used: one group of subjects received triage training (experimental group) while the other group received unrelated, yet comparable, training (control group). The purpose of this experiment was to test two hypotheses. First, it is expected that participants in the experimental training group will demonstrate an increase in accuracy of triaging victim images as compared to the non-trained participants in the control group. Secondly, the

experimental group will demonstrate a change in fixation patterns in regards to salient features versus non-salient features, compared to those without training in the control group. Both experimental and control groups were given images to properly categorize into the triage categories. Experimental and control groups viewed the same images twice (baseline and second view/post-training). Each image ranged in severity, from minor to severe, and contained text boxes with vital sign information necessary to properly triage the victim images; presented within each image was a separate text box for each of the following: airway, breathing rate and circulation (presence of radial pulse). Data was collected using the BeGaze gaze eye tracker during the experiment. This recorded the participants' eye movements and was used to analyze their gaze patterns during the visual stimulus. Data collected for resulting triage accuracy and gaze patterns were compared using Fischer Exact tests as well as 1-tailed t-tests. The Fischer Exact tests failed to reach significance to support the alternative hypothesis that participants with the triage training would increase in accuracy when triaging the victim images. The Fischer Exact test was negatively impacted due to a small sample size; however, descriptive statistics demonstrated a potential increase in accuracy by the trained group. Looking at fixation patterns, data shows a significant increase within the experimental group in regards to attention to the text boxes (salient features of the images) with respect to baseline entry times to the second viewing (post-training viewing) entry times.

TABLE OF CONTENTS

Title	i
Reader's Approval Page	ii
Acknowledgements	iii
Abstract	iv
Table of Contents	vi
List of Tables	viii
List of Figures	ix
List of Abbreviations	x
Introduction	1
Methods	5
Recruitment Procedure	6
Inclusions & Exclusions	6
Inclusions	6
Exclusions	7
Subjects	7
Detailed Procedures of Experiment	8
Experimental Group	8
Control Group	9
Data Collection Process	9
Subject Data Collection	11

Statistical Analysis	13
Results	15
Accuracy of Triaging	15
Fixation Patterns	17
Entry Time	17
First Fixation Duration	19
Fixation Total	21
Between Groups	24
Discussion	28
Entry Time	29
First Fixation Duration	31
Fixation Total	31
Conclusion	32
References	35
Curriculum Vitae	37

LIST OF TABLES

Table	Title	Page
1	Number of Correct Triaged Responses Within Groups	16
2	Number of Correct Triaged Responses Between Groups	17
3	All Images: Average Entry Time in ms	18
4	4 Images: Average Entry Time in ms	19
5	All Images: Average First Fixation Duration in ms	20
6	4 Images: Average First Fixation Duration in ms	21
7	All Images: Average Fixation Total in ms	22
8	4 Images: Average Fixation Total in ms	24
9	All Images & 4 Images Average Entry Time Between Groups	25
10	All Images & 4 Images Average First Fixation Duration Between Groups	26
11	All Images & 4 Images Average Fixation Total Between Groups	27
12	Number of Correct Responses All Images	29
13	Number of Correct Responses 4 Images	29

ABBREVIATIONS

ABC	Airway, Breathing, & Circulation
AOI	Areas of Interest
ATLS	Advanced Trauma Life Support
BP	Blood Pressure
BR	Breathing Rate
cm	Centimeters
EMT	Emergency Medical Technician
HR	Heart Rate
ms	Milliseconds
RPM	Respirations, Perfusion & Mental Status
RR	Respiratory Rate
START	Simple Triage And Rapid Treatment
WHO	World Health Organization

Introduction:

This experiment used BeGaze (a gaze tracking software) to track and record participants' fixation patterns upon triage victim images in which the participants were asked to accurately categorize trauma victims into one of the four triage classifications (green, yellow, red, black). The Simple Triage And Rapid Treatment (START) triage program was used as the basis of the training component for the experimental group. The START triage program has been used to train first responders since 1983 (START Triage, 2013). This program is designed and intended for those with basic first aid training. The first responder should be trained in basic first aid to be able to recognize the severity of injuries, as well as to understand how to reposition and re-open an airway if closed or blocked, a vital step during the triaging process. The START program is a useful tool during large-scale accidents or mass casualties where there are multiple injured persons. A mass casualty as defined by the World Health Organization (WHO) is "an event which generates more patients at one time than locally available resources can manage using routine procedures. It requires exceptional emergency arrangements and additional or extraordinary assistance" (WHO, 2007). The START triage system, however, is used in every situation to which first responders are called, no matter how minor or major the incident.

The START triage system is specifically designed to decrease confusion for the first responder and to assist in properly categorizing the victims in order to rapidly treat them based on the severity of their injuries and mental status. Once the initial group of first responders has arrived to the incident, they will quickly take on their predetermined

roles. The role of the triage unit leader is to strictly triage. Each victim should not take more than one minute to assess, triage and appropriately tag. Tagging is completed by the triage unit leader who will indicate the status of the victim by a small piece of paper with the colors order from black, red, yellow and green (Critical Illness and Trauma Foundation, 2001). Attempting to treat or assist any victims during the triaging period is prohibited as to not waste precious time needed to get help for severely injured victims. This will allow for prioritization of those with the smallest “golden hour” (Critical Illness and Trauma Foundation, 2001) and the highest need for medical attention. The golden hour can be described as the window of time (one hour) from the time of the injury that an individual victim has the best chance of recovery if they can get Advanced Trauma Life Support (ATLS).

It is imperative that this entire process runs smoothly and is coordinated ahead of time if possible. An important aspect of the process is obtaining information about the victim in order to properly triage them. This information is taught to the first responder by the START program and will be implemented at the scene of the incident. There are simple steps used to determine a victim’s triage level. The START program recommends three simple vital signs to be checked per victim when beginning the triage process: respirations, perfusion and mental status (RPM). These three vital signs were used as the basis for the three vital signs used on the images that the participants viewed. A similar and easy to remember mnemonic was used: ABC: airway, breathing, circulation (World Health Organization Department of Child and Adolescent Health and Development, 2005). Mental status was excluded from the images because it is difficult to assess via a

still image, as well as the fact that it can be subjective. Mental status is a secondary feature used to triage a person after the ABC's have been fully assessed. It adds another dimension of the status of the victim and cannot easily be determined without a question and answer assessment.

The first and by far easiest way to work through and organize a large crowd is to ask everybody that can hear and understand the triage unit leader to move to a safer location where they will be assisted by another first responder. This group of people is triaged as minor (green) as they do not demonstrate a loss in mental status nor any major injuries that impede them from doing what was asked of them. Typically these people will have an open airway, relatively stable respiratory rate (RR), blood pressure (BP) and heart rate (HR). The status of the minor (green) victims is unlikely to change over days, allowing victims with more severe injuries to be treated first. Usually, triaging green/minor victims will clear a large number of people who only suffered minor injuries from the chaos and confusion of the severely injured victims.

The triage unit leader would proceed to triage the remaining victims into the three remaining triage categories: delayed (yellow), immediate (red) and deceased (black). For each victim, the ABC's will be analyzed to ensure proper triage classification. If the victim's airway is compromised, the triage unit leader repositions it to open it—if possible—and breathing rate (BR) is assessed for normality. If the airway was compromised and subsequently cleared, a check is performed to determine if breathing is resumed. Victims whose airway was not compromised will fall under the delayed (yellow) category; some injury (ie: relatively minor such as a distal extremity injury) has

prevented them from either understanding or proceeding accordingly when asked to move in the first step. Those triaged as delayed (yellow) will require medical attention but immediate action is not contingent upon their survival. The first responder then checks for normal breathing rate (BR), which should fall within 10 and 29 breaths per minute (U.S. Department of Health & Human Services, 2013). If it is not within this range, the victim is tagged as immediate (red) and the triage unit leader can move to the next victim. If the victim's breathing rate does fall in between the 10-29 range, the triage unit leader will then evaluate circulation; primarily checking if a radial pulse is present. Victims without a strong radial pulse are triaged as immediate (red)—as they may have internal or external bleeding leading to low radial pulse, which decreases their chances of survival, and increases their urgency to receive medical attention. Victims who display compromised airway, decreased/increased BR and/or decreased circulation are triaged as immediate (red)—they require immediate medical attention within 60 minutes to increase likelihood of survival. Injuries will include proximal extremity injuries (crushed, broken, burns, etc.), compromised airway, crushed skull, etc. Lastly, are the victims that are triaged as black, are those that have not survived the accident due to the severity of their injuries or are unlikely to survive and receive only palliative care. Once all victims have been assessed and triaged, those with the highest priority level can be assisted until they are transported.

This study examines the effect of training individuals in the START triage system on the way they interact with images of casualty victims and their ability to appropriately triage these victims. We postulate that trained subjects would focus on the salient features

in the images. In particular, participants should fixate longer at the learned salient features and use them as their primary source before triaging the victim images. We hypothesize that subjects with training (experimental) would demonstrate a change in duration of fixation in regards to salient features versus non-salient features, compared to those subjects with no training (control). We also hypothesize that the trained group would demonstrate an increase in accuracy in properly triaging the victim triage images, and the control group's accuracy would either decrease or stay the same.

Previous findings have shown that viewing never before seen images takes the brain longer to process than viewing images previously seen (Yarbus, 1967). As the number of viewings increase, the location and duration of gaze upon salient features within the image changes. For instance, the first viewing of an image will contain many fixation points each at very short durations. As the number of viewings increases (more than 3-5 viewings), the number of fixations will decrease, and the duration of each fixation will increase (keeping the total viewing time constant) allowing the individual to focus on the salient features. In particular, entry time, first fixation duration and fixation total for each of the salient areas of interest (AOI's) were analyzed.

Methods

The following will describe the data collection process using the BeGaze program as well as the recruitment procedure, inclusion & exclusion, subjects, detailed procedures of experiment and statistical analysis.

Recruitment Procedure:

Participants were recruited by posting an advertisement on Boston University School of Medicine's Quickie Jobs website. Once they contacted the researcher, they were asked screening questions, which are located below in the inclusions and exclusion criteria. If the participants answered "yes" to any of the exclusion criteria questions, they were automatically excluded from the experiment, and no identifying information was recorded. If the subject meets the inclusion criteria and all of the answers were "no" to the exclusion criteria questions, their contact information was taken down and a time was scheduled for them to come participate. To ensure complete randomness, each subject was assigned to the experimental group or the control groups based on the order they arrived for testing.

Inclusions & Exclusions:

Inclusions:

- 18 - 64 years of age, healthy
- Willing to participate in the study (informed consent)
- Participants will only be English speaking. (Due to the requirements to carefully screen for vision loss, to be sure of informed consent and having a clear understanding of the instructions necessary to complete this study, only subjects who speak English, which is the PI and student investigator's language of competence can be safely enrolled.)

Exclusions:

- Have uncorrected loss of vision
- Neurological disease
- Brain injury of any kind
- Become uncomfortable seeing blood in images

Subjects:

- There were 24 subjects enrolled, 12 in each group.
- The two groups were:
 - 1) Training in the START triage protocol (Experimental Group)
 - 2) No Training (Control Group)

Participants were put into two groups: those who received training in the START triage program (the experimental group), and those who received unrelated training about patient transportation (the control group). Participants in the experimental group were trained with the START triage program, which included a PowerPoint presentation outlining the four categories of the START program (green; yellow; red; black), as well as necessary steps to properly categorize victims (ABC's: Airway; Breathing; Circulation). The control group received an equivalent PowerPoint presentation that trained them how to transport patients from the site of the accident to the appropriate hospital or clinic. This served as an equivalent training period that functioned to mimic each slide from the training of the experimental group. Both training sessions included a

step-by-step outline, followed by a schematic at the end of the training that outlined and highlighted each step for the respective training.

Detailed Procedures of Experiment:

Experimental Group:

Once calibrated, the participant was given a set of instructions on the computer screen informing them of the upcoming slides. Instructions informed the participants of their four choices to classify the images. They were asked to determine the level of severity for each image based on what is presented to them. Prior to the baseline phase, the only information given to them was a color bar with green, yellow, red, black colors labeled as “minor”, “moderate”, “severe” and “deceased”, respectively. Once they read through the instructions and were ready to proceed, they viewed (in random) a series of 8 baseline images. Once completed, another set of instructions was presented and the participant read through a set of PowerPoint slides outlining the START triage program with a thorough explanation of the steps to take in order to properly triage a victim. There were 11 slides in total in this training phase. The last two slides outlined in schematic form a summary of the steps explained in the earlier slides. A very similar schematic comes with the START triage program documents and is used in field triage (U.S. Department of Health & Human Services, 2013). Upon completion of the experiment, the participants viewed the same 8 images once more (random order again) using the knowledge they obtained from the training to triage the victims in the images.

Control Group:

After calibration, the control participants were given the same set of two instruction screens and same 8 baseline images as the experimental group. After viewing the baseline images (in random order), the participants were given a different set of “training slides” which did not correspond to triage or the START program. Instead, these slides provided information on patient transport. A mock training was used for the control group to maintain consistency between the two groups, and prevent differential fatigue in the experimental group and/or prevent the control group from receiving an advantage of a shorter testing period. The control group training session taught transportation categories, as well as steps to properly categorize patients for transportation, followed by a schematic summarizing their patient transportation training. As much as possible, the number of slides as well the density of information per slide also remained relatively consistent for each group. Patient transport training for the control group consisted of 10 slides depicting how to transport victims from the scene of an accident. The last two slides also depicted a schematic summarizing patient transport.

Data Collection Process

To determine if accuracy in triaging occurred within the experimental group, the BeGaze software recorded participants triage selection. The camera and software also recorded participants’ gaze on the AOI features located within each image to determine fixation patterns. Each image (for both the baseline and the post-training viewing) contained three text boxes, one for each of the ABC’s individualized for the victim within

the image. These three text boxes designated as AOI's via the BeGaze software program allowed these salient features to be analyzed. For example, the airway AOI, would read either "Airway: Open" or "Airway: Closed"; breathing AOI would read "Breathing Rate: [number]/min"; and the circulation AOI would read "Circulation: Radial Pulse Present" or "Circulation: Radial Pulse Not Present". The fourth AOI (white space) encompasses the remainder of the screen that is not a part of the distinguished text box AOI's. Within the white space AOI is the victim and their injury. In order to test if fixation patterns changed post-training, the BeGaze software allowed for the analysis of fixation patterns upon the AOI's of interest within the images. Two of the key features that were recorded by BeGaze were saccades and fixations for each of the four AOI's per image, per participant; this study focused on the latter (fixations). A saccade is any eye movement from one location to another. For instance, when an individual looks at a picture and first fixates upon the face in the center, then, moves to the top left corner; this eye movement from the center to the top left corner is a saccade, or a linear eye movement. Once the eye's gaze stops upon a location for at least 80 milliseconds (ms), it is considered to be a fixation (Sensomotoric Instruments GmbH, 2010).

Images used in this study represented all 4 levels of the triage process (green, yellow, red and black). Images representing green/minor victims and black/deceased victims were not the central focus of this study due to their simplicity but were used in the analysis of the data. The two green and the two black victim images are easily triaged by the study participants without any prior knowledge of the triaging system. These images were included to completely simulate a realistic field triage situation in which

there are green and black victims that must be triaged. A real field triage situation would have many green and black victims that will stand in the way of triaging and providing care to yellow and red victims—those who need the most help. For this reason, the green and black victim images were not disregarded from this study, but were used to completely embody a realistic field triage experience to the participant.

Entry time (ms), first fixation duration total (ms) and fixation total (ms) were used to compare each AOI from baseline to second viewing in the following ways. Entry time provides information regarding the time at which the participant's gaze first entered each particular AOI. Comparing this variable from the baseline viewing to the second viewing demonstrates if the participants in fact looked at specific AOI's sooner, later or had no change, when presented with the victim images in the second viewing. First fixation duration total time was used to reveal how long (milliseconds) the participant's first fixation was upon each AOI. This variable can be used to determine if participant's first fixation upon a particular AOI changed in duration during their second viewing as compared to their first viewing. Lastly, fixation total time was the sum of the amount of time (in milliseconds) that participants fixated upon an AOI. These three variables (entry time; first fixation duration; fixation total) were used to compare within groups to analyze if there was a statistically significant change within the two groups from pre- to post-training.

Subject Data Collection:

For gaze tracking, the subject sat in front of the computer connected to the BeGaze eye-tracking device. Subjects sat roughly 30 cm away from the computer and the BeGaze eye-tracking device. Next, the BeGaze program ran a calibration test to align the participant's gaze to the 45-degree angle of the gaze-tracking device. During the calibration sequence, the participant was asked to follow a black dot upon a white screen to different locations within the screen. The computer displayed the calibration results, which were presented in x and y coordinates representing how much the participant's gaze deviated from the stimulus presented on the screen. The x and y coordinates are measured during this process and displayed upon completion of calibration. This process was repeated until both the x and the y coordinates were below 1.00 (gaze deviated less than 1 degree on both the x and y coordinates). Once appropriately calibrated, the experiment began and the BeGaze device directly sent the gathered information to the computer upon completion of the experiment per participant.

The visual stimulus information was presented as a series of images, each corresponding to the levels of injury severity of the START triage program (green, yellow, red and black) on the computer screen. There were two images per triage category, comprising a total of eight images. Each of the 8 resulting images were presented in random order. Each image included three small text boxes each containing vital signs (ABC's) to properly triage the victim. These text boxes were designated as salient areas of interest in the subsequent analysis. Participants were asked to make their selection of triage category using the number pad on the keyboard of the laptop in front of them. The numbers 1 through 4 were indicative of the triage colors: 1 = green; 2 =

yellow; 3 = red; 4 = black. The computer recorded every keypad selection per participant. The images appeared one at a time on the screen for 15,000 milliseconds each for the subjects to categorize into triage categories. Between each image was a white screen with a black crosshair, which remained on the screen for 1,000 ms to refocus the participants gaze to the center of the screen. The 8 images were displayed before training to serve as the baseline data. The experimental group continued the experiment by reviewing the simulated START triage program of PowerPoint slides provided for them, while the control group reviewed a comparable mock training (patient transport training), which simulated a nearly identical experience to that of the experimental group. Once the training sessions were complete, the same 8 images used for the baseline, were randomly viewed once again for a direct comparison both within and between groups for experimental and control groups as a result of the START triage training.

Statistical Analysis

Three statistical tests were used to assess the two hypotheses. The Fischer Exact test was used to examine whether there was a change in triage accuracy within the experimental group or control group. One-tailed paired t-tests were used to analyze if gaze pattern variables changed from baseline to second viewing, within both the experimental group and control group. One-tailed t-tests with unequal variance were used to compare between groups (the control group vs. the experimental group) in the testing phase.

Fischer Exact tests were performed to determine if there were differences in accuracy of triaging between the experimental group and control group both within subjects and between subjects. Since all of the participants (both the experimental group and control group) saw both baseline and second viewing of the images, all of the participant results across all images were summed together to create the data used within the Fischer Exact test. The following Fischer Exact tests were performed for all images and for the 4 complex images:

1. *Control group*: baseline to second viewing [to examine the effect of the non-pertinent training session]
2. *Experimental group*: baseline to second viewing [to examine the effect of the pertinent training session]
3. *Baseline*: control group to experimental group [to determine that the groups were similar before training]
4. *Second viewing*: control group to experimental group [to compare the behavior of the two groups with and without training in START]

One-tailed paired t-tests were completed within groups using the gaze data collected from all 8 images and the 4 complex images. One-tailed paired t-tests were conducted for the experimental group and control group on the following parameters, to determine the effect of the training sessions on each of these variables:

1. *Entry Time*: baseline to second viewing
2. *First Fixation Duration*: baseline to second viewing
3. *Fixation Total*: baseline to second viewing

One-tailed unpaired t-tests were completed between groups using the data collected from all 8 images and the 4 complex images to compare the control group to the experimental group. The following 1-tailed unpaired t-tests were conducted:

1. *Entry Time*: control group to experimental group
2. *First Fixation Duration*: control group to experimental group
3. *Fixation Total*: control group to experimental group

Results

The following section displays the results from the Fischer Exact tests (accuracy of triaging), the 1-tailed paired and unpaired t-tests (fixation patterns) that were performed. Throughout this section, the four complex images (two yellow and two red triage images) are highlighted, as these are the images that are likely to present the biggest challenge to the subjects.

Accuracy of Triaging:

Fischer Exact tests were conducted within groups from baseline to second viewing for both the control group and the experimental group from data depicted in Table 1. In all cases, no significant p-values were achieved. As Table 1 illustrates, there was a slight increase in accuracy of triaging for the control group for both all images, and the 4 complex images. Similarly, the experimental group demonstrated improved accuracy in triage decision-making. However the experimental group improved their triage decision-making at a greater rate. Due to the relative closeness of response patterns

between the control group and experimental group, Fischer Exact tests did not achieve significant p-values.

<u>Control</u>	Baseline	2nd View	Difference
All Images	60	72	12
4 Images	24	30	6
<u>Experimental</u>			
All Images	66	82	16
4 Images	28	37	9

Table 1: Number of Correct Triage Responses Within Groups. “Four images” refers to the four complex images (two yellow and two red triage images).

Fisher Exact tests were completed between groups from control to experimental for both baseline and second viewing and did not provide significant p-values (Table 2). As shown in Table 2, the experimental group performed slightly better in triage decision-making both at baseline and in the second viewing. Due to small sample size and relative closeness in numerical sums of correctly triaged images between control and experimental group, significance was not achieved within the Fischer Exact tests. These numerical outcome differences show the need for a larger sample size. As demonstrated in Table 2 there was a proportional increase from baseline to second viewing in the number of correctly triaged images (all images and 4 complex images) for the experimental group as compared to the control group. The experimental group showed a greater improvement from baseline to second viewing.

Control vs Experimental: <u>Baseline</u>	Control	Experimental	Difference
All Images	60	66	6
4 Images	24	28	4
Control vs Experimental: <u>2nd Viewing</u>			
All Images	72	82	10
4 Images	30	37	7

Table 2: Number of Correct Triage Responses Between Groups. “Four Images” refers to the four complex images (two yellow and two red triage images).

Fixation Patterns:

There were two methods of analyzing the fixation pattern data: using all 8 images and using the 4 complex images (the two yellow and the two red). The data will be presented in the following order: 1) Entry Time; 2) First Fixation Duration; 3) Fixation Total; 4) Between Groups; within each subgroup all 8 triage images and the 4 complex (2 yellow and 2 red) triage images will be analyzed.

Entry Time

One-tailed paired t-tests were completed within groups from baseline to second viewing for both the control group and the experimental group; the resulting p-values are shown in Table 3. As expected, entry time for the second viewing of the control group remained relatively consistent compared to their baseline, yielding insignificant changes; while the experimental group demonstrated statistically significant changes baseline to second viewing. The experimental group shows a decrease of entry time into all of the AOI's, rendering significant p-values for the airway AOI ($p = 0.009$), the breathing AOI ($p = 0.008$) and the circulation AOI ($p = 0.0001$). The white space AOI did not result in a

significant difference of entry time from baseline to the second viewing for the experimental group.

ALL IMAGES: Average Entry Time in ms			
	Baseline	2nd Viewing	p-value
Airway:			
Control	4369	4163	0.333
Experimental	4657	2877	0.009
Breathing:			
Control	4022	4016	0.264
Experimental	3866	2302	0.008
Circulation:			
Control	3795	4169	0.133
Experimental	4418	2537	0.0001
White Space:			
Control	193	224	0.186
Experimental	795	413	0.071

Table 3: All Images: Average Entry Time in ms. Significant p-values are in bold face font

This same pattern persisted when the 2 green and 2 black images were removed from the data analysis, and the 2 yellow and 2 red images were analyzed alone. Table 4 outlines the p-values for each of the AOI's from baseline to second viewing. The experimental group shows a significant decrease of entry time for two of the AOI text boxes (breathing AOI: $p = 0.002$; circulation AOI: $p = 0.006$), and insignificant differences in the airway AOI and the white space AOI. Similarly to all images, the control group remains insignificant for all AOI's.

4 IMAGES: Average Entry Time in ms			
	Baseline	2nd Viewing	p-value
Airway:			
Control	5279	4870	0.206
Experimental	4447	3059	0.400
Breathing:			
Control	4323	3557	0.124
Experimental	3888	1833	0.002
Circulation:			
Control	3924	3863	0.330
Experimental	3998	2325	0.006
White Space:			
Control	278	365	0.451
Experimental	401	233	0.452

Table 4: 4 Images: Average Entry Time in ms. Significant p-values are in bold face font.

First Fixation Duration

One-tailed paired t-tests were used to compare subjects' first fixation duration from baseline and second viewings for all AOI's (Table 5). The duration of the subjects' first fixation on the circulation AOI was significantly longer in the second viewing for the control group ($p = 0.003$) and significantly shorter for the experimental group ($p = 0.01$), compared to baseline. During the first viewing of this AOI, all participants read this text box. In the second viewing, the experimental group appeared to use their knowledge from their training and move on to quickly after reading the AOI to triage the image during their second viewing. Conversely, the control group participants did not receive the training to inform them as to what this text box was signifying, and on average spent more time on this AOI during their second viewing. The control group's first fixation on the white space was also significantly longer from baseline to the second viewing. The

control group focused significantly more on the injury within the image during the second viewing, rather than the AOI text boxes, causing their fixation on the white space AOI to be significantly longer ($p = 0.001$). The experimental group spent significantly less time fixating upon the white space AOI as compared to baseline first fixation duration ($p = 0.024$).

ALL IMAGES: Average First Fixation Duration in ms			
	Baseline	2nd Viewing	p-value
Airway:			
Control	216	191	0.059
Experimental	220	206	0.204
Breathing:			
Control	190	180	0.448
Experimental	183	183	0.363
Circulation:			
Control	172	210	0.003
Experimental	199	162	0.010
White Space:			
Control	260	331	0.001
Experimental	261	225	0.024

Table 5: All Images: Average First Fixation Duration in ms. Significant p-values are in bold face font.

The following analyzes the subject's initial fixation on the AOI's of the more complex images, the 2 yellow and 2 red images (Table 6). For these images, the control group demonstrated a significant increase ($p = 0.004$) from baseline to the second viewing of their first fixation on the white space AOI. The experimental group demonstrated a decrease in first fixation duration upon the circulation AOI (as seen within the all images data), which is supportive of demonstrating the training effect. The

control group took significantly more time to look at the circulation AOI ($p = 0.018$). The control group demonstrated a significant increase ($p = 0.004$) in first fixation on the white space AOI from baseline to second viewing, demonstrating their reliance on the white space AOI rather than the AOI text boxes (circulation, breathing, airway) during triage decision-making. The experimental group showed a statistically significant decrease ($p = 0.029$) in first fixation duration into the white space, as was expected from their training.

4 IMAGES: Average First Fixation Duration in ms			
	Baseline	2nd Viewing	p-value
Airway:			
Control	216	191	0.135
Experimental	182	196	0.226
Breathing:			
Control	198	175	0.253
Experimental	175	178	0.353
Circulation:			
Control	167	203	0.018
Experimental	231	177	0.162
White Space:			
Control	253	365	0.004
Experimental	250	213	0.029

Table 6: 4 Images: Average First Fixation Duration in ms. Significant p-values are in bold face font.

Fixation Total

For fixation total (sum of all fixation durations on the AOI), 1-tailed paired t-test p-values comparing baseline to second viewing for the control and experimental group were completed and subsequent p-values are located within Table 7. Control group data depicts a decrease of fixation total time from baseline to second viewing for all AOI's except the white space AOI, but the only significant reduction was in the fixation time on

the circulation AOI ($p = 0.005$). Since the airway AOI is the least complex AOI (“Airway: Open” or “Airway: Closed”), it did not cause significant fluctuation (increase or decrease) in fixation for either groups. In addition, the control group white space AOI resulted in a significant increase in the fixation total time yielding significance ($p = 0.026$) from baseline to second viewing.

The trends for the experimental group are very different. As a result of the triage training the experimental group fixation totals, on average, increased from baseline to second viewing for each of the trained AOI text boxes (airway, breathing, circulation) though they increased for the airway AOI. However, these changes were not significant. Conversely, the experimental group spent significantly less time fixating on the white space AOI ($p = 0.00001$). The white space AOI data demonstrates the change in focus areas between the control group and the experimental group from the white space to the AOI text boxes (airway, breathing, circulation) as a result of the triage training.

ALL IMAGES: Average Fixation Total in ms			
	Baseline	2nd Viewing	p-value
Airway:			
Control	675	600	0.332
Experimental	1009	875	0.262
Breathing:			
Control	664	506	0.059
Experimental	765	825	0.491
Circulation:			
Control	963	729	0.005
Experimental	782	964	0.147
White Space:			
Control	7138	7964	0.026
Experimental	7263	5333	0.00001

Table 7: All Images: Average Fixation Total in ms. Significant p-values are in bold face font.

Using the 4 complex images to analyze the fixation total times also depicts very consistent results shown in Table 8. While not all resulting in significance, except in the case of the breathing AOI ($p = 0.032$) the control group second viewing, on average, decreased their fixation total times for the AOI text boxes (airway, breathing, circulation). The experimental group, on average, increased their fixation total times on these same AOI's though these changes were only significant for the airway AOI ($p = 0.034$) and the circulation AOI ($p = 0.0008$). Further evidence for the change in behavior between the control and experimental group gaze pattern is seen in the white space AOI which shows a decrease in total fixation during the second viewing, as compared to the baseline, for the experimental group (resulting in significance: $p = 0.000003$). Finally, while not reaching statistical significance, the control group demonstrated slight, but not significant increase in total fixation on the white space AOI during the second viewing as compared to baseline.

4 IMAGES: Average Fixation Total in ms			
	Baseline	2nd Viewing	p-value
Airway:			
Control	604	589	0.389
Experimental	648	823	0.034
Breathing:			
Control	690	396	0.032
Experimental	766	982	0.139
Circulation:			
Control	946	834	0.129
Experimental	682	1224	0.0008
White Space:			
Control	7355	7752	0.258
Experimental	7571	5600	0.000003

Table 8: 4 Images: Average Fixation Total in ms. Significant p-values are in bold face font.

Between Groups

Tables 9 - 11 depict p-values from 1-tailed unpaired t-tests comparing the gaze patterns of the control group to the experimental group at comparable stages of the experiment. Significant differences can be seen primarily within the second viewing data, and are represented by boldface font. Baseline data is included within Tables 9-11, however, this data will not be described in detail but was put in to be complete. Baseline data is used for comparisons within groups in the preceding section.

Table 9 depicts all entry time data from between groups. Though there are no differences at baseline between the groups, early entry time for the experimental group, as compared to the control group (all images) results in statistically significant differences for all AOI text boxes. White space AOI (all images) on the other hand, was not significantly different between groups at second viewing. However, data reveals an

earlier entry time for the control group, and a later entry time into the white space AOI for the experimental group.

As for the 4 complex images data analysis, the control group and experimental group had significantly different entry times for all AOI text boxes (experimental group earlier than control group): airway AOI: $p = 0.049$; breathing AOI: $p = 0.007$; circulation AOI: $p = 0.006$. The white space AOI resulted in significance within the all images analysis for the baseline data entry times ($p = 0.04$) when comparing the control group to the experimental group. While both groups viewed the white space AOI during the baseline viewing, the experimental group looked at it later compared to the control group.

ALL IMAGES: Average Entry Time in ms				4 IMAGES: Average Entry Time in ms			
	Control	Experimental	p-value		Control	Experimental	p-value
Airway:				Airway:			
Baseline	4369	4657	0.323	Baseline	5279	4447	0.175
2nd Viewing	4163	2877	0.018	2nd Viewing	4870	3059	0.049
Breathing:				Breathing:			
Baseline	4022	3866	0.397	Baseline	4323	3888	0.306
2nd Viewing	4016	2302	0.004	2nd Viewing	3557	1833	0.007
Circulation:				Circulation:			
Baseline	3795	4418	0.149	Baseline	3924	3998	0.455
2nd Viewing	4169	2537	0.0002	2nd Viewing	3863	2325	0.006
White Space:				White Space:			
Baseline	193	795	0.040	Baseline	278	401	0.278
2nd Viewing	224	413	0.145	2nd Viewing	365	233	0.429

Table 9: All Images & 4 Images Average Entry Time Between Groups.

Data from first fixation duration between groups is located within Table 10. First fixation duration between-group analysis (all images) revealed statistical significance for

the white space AOI second viewing ($p = 0.0003$) as a decrease from the control group compared to the experimental group. The experimental group participants resulted in a significant decrease from first viewing to second viewing in first fixation duration for the circulation AOI ($p = 0.003$). This can be interpreted as a decrease in first fixation duration for the circulation AOI when comparing the control group and the experimental group. As expected, the experimental group resulted in a statistically significant decrease in white space AOI from baseline to second viewing for first fixation duration for the 4 complex images ($p = 0.001$). The circulation AOI for the baseline viewing (all images) reveals a statistically significant difference of first fixation duration for both groups which resulted in significance ($p = 0.043$), however, the data displays similar first fixation durations.

ALL IMAGES: Average First Fixation Duration in ms				4 IMAGES: Average First Fixation Duration in ms			
	Control	Experimental	p-value		Control	Experimental	p-value
Airway:				Airway:			
Baseline	216	191	0.435	Baseline	216	182	0.063
2nd Viewing	220	206	0.165	2nd Viewing	191	196	0.270
Breathing:				Breathing:			
Baseline	190	183	0.365	Baseline	198	175	0.068
2nd Viewing	180	183	0.420	2nd Viewing	175	178	0.447
Circulation:				Circulation:			
Baseline	172	199	0.043	Baseline	167	231	0.089
2nd Viewing	210	162	0.003	2nd Viewing	203	177	0.160
White Space:				White Space:			
Baseline	260	261	0.491	Baseline	253	250	0.448
2nd Viewing	331	225	0.0003	2nd Viewing	365	213	0.001

Table 10: All Images & 4 Images Average First Fixation Duration Between Groups.

T-tests for fixation total produced significant p-values for between-group analysis for all second viewing AOI's for all images and the 4 complex images. Second viewing of the AOI text boxes from the control group to the experimental group consisted of a significant increase (all images and 4 complex images) and a significant decrease in fixation total in the white space AOI (all images: $p = 0.0000001$; 4 images: $p = 0.0002$). It was expected that the experimental group decrease their total fixation time within the white space AOI demonstrating their triage training. The experimental group on average, spent more time on the airway AOI at baseline as compared to the control group resulting in a statistically significant increase on the airway AOI for all images ($p = 0.049$). The 4 complex images data analysis reveals significance from the circulation AOI baseline viewing ($p = 0.025$), caused by the experimental group, on average, spending less time on the circulation AOI in the baseline, as compared to the control group.

ALL IMAGES: Average Fixation Total in ms				4 IMAGES: Average Fixation Total in ms			
	Control	Experimental	p-value		Control	Experimental	p-value
Airway:				Airway:			
Baseline	675	1009	0.049	Baseline	604	648	0.365
2nd Viewing	600	875	0.005	2nd Viewing	589	823	0.010
Breathing:				Breathing:			
Baseline	664	765	0.197	Baseline	690	766	0.318
2nd Viewing	506	825	0.002	2nd Viewing	396	982	0.0002
Circulation:				Circulation:			
Baseline	963	782	0.077	Baseline	946	682	0.025
2nd Viewing	729	964	0.007	2nd Viewing	834	1224	0.006
White Space:				White Space:			
Baseline	7138	7263	0.426	Baseline	7355	7571	0.387
2nd Viewing	7964	5333	0.0000004	2nd Viewing	7752	5600	0.0002

Table 11: All Images & 4 Images Average Fixation Total Between Groups.

Discussion:

The experimental design of this study was created to test the hypothesis that the experimental group would demonstrate an increase in accuracy in properly triaging the victim triage images as compared to the control group; also, subjects with training would demonstrate a change in duration of fixation in regards to salient features verses non-salient features compared to those subjects without training. While the Fischer Exact tests did not result in statistical significance in the answer accuracy, numerical differences indicate that this may have resulted with the use of a larger sample size. Finally, the second hypothesis of change in fixation patterns was determined valid.

To test the first hypothesis, Fischer Exact tests were performed to check if there was a difference in performance between the control group and the experimental group in accuracy when triaging victim images. The Fischer Exact tests did not result in significance but resulting trends indicate a numerical difference in accuracy between the control group and the experimental group that may result from a larger sample size. This sample size was insufficient to produce significance, but if the sample size was simply doubled, and the same trends in observed data remained, significance would prevail. The raw data values for number of correctly triaged responses (all images) can be seen in Table 12 shown below. The experimental group performed 33% better than the control group (experimental group-16; control group-12). As demonstrated in Table 13, the number of correct responses from the 4 complex images during the second viewing for the experimental group increased to 9 more correctly triaged victim images and for the control group, only 6 correctly triaged victim images. It is obvious that both of the two

groups improved in accuracy in triage categories during the second viewing. However, the experimental group resulted in more correctly triaged victim images, just not enough to produce statistical significance.

<u>ALL Images</u>	Baseline	2nd View	Difference
Experimental	66	82	16
Control	60	72	12

Table 12: Number of Correct Responses All Images. Sums of all participants’ number of correctly triaged images.

<u>4 Images</u>	Baseline	2nd View	Difference
Experimental	28	37	9
Control	24	30	6

Table 13: Number of Correct Responses 4 Images. Sums of all participants’ number of correctly triaged images.

In regards to the second hypothesis—fixation patterns will change for salient features in the experimental group from the baseline to their post training viewing of the victim images—each of the observed measurables will be discussed on it’s own for simplicity purposes, and will proceed in the following order: 1) Entry Time; 2) First Fixation Duration; 3) Fixation Total.

Entry Time

There is enough data to conclude an obvious change in gaze patterns and fixation times within the experimental group not found within the control group. The earlier entry times for the experimental group from baseline to second viewing reveals the change in behavior and each groups level of importance associated with the AOI’s. The resulting

data collected for the three salient AOI's (airway, breathing, circulation) demonstrates that participants with training (experimental group) focused on the ABC information that was presented to them within their triage training earlier than participants who did not receive this training (control group). This gaze pattern supports the hypothesis that participants would change their fixation patterns on the salient features (AOI's) when exposed to the images post training, as compared to the control group. The data presented for all images as well as that for the 4 complex images, provide almost identical data, further supporting the analysis. It can be determined from this, that regardless of the complexity of the image, the trained participants (experimental group) changed the order of salience of visual information to focus on the trained salient AOI's earlier than the non-trained participants (control group).

When comparing the two groups, the experimental group participants fixated upon salient features significantly earlier when presented with images post-training compared to the controls. The white space AOI data depicts a sharp delay in entry time for the experimental group when compared to the control group for the baseline ($p = 0.04$) and in the second viewing (not resulting in significance but noteworthy). The experimental group demonstrates significantly earlier entry times into the AOI text boxes as compared to the control group entry times (airway AOI: $p = 0.018$; breathing AOI: $p = 0.004$; circulation AOI: $p = 0.0002$). This same pattern persists within the 4 complex images data analysis: airway AOI: $p = 0.049$; breathing AOI: $p = 0.007$; circulation AOI: $p = 0.006$). The experimental group presented sooner entry times into the salient AOI text boxes and later entry times into the white space, as compared to the control group.

First Fixation Duration

As expected, the control group demonstrated a statistically significant increase in first fixation duration on the white space AOI (all images: $p = 0.001$; 4 images: $p = 0.004$). The experimental group demonstrated in a significant decrease in first fixation time on the white space AOI for the all images data analysis ($p = 0.024$) and the 4 complex images data analysis ($p = 0.029$). This demonstrated that there was a change in behavior in both the control group and the experimental group. For those AOI's that were significant—circulation AOI and white space AOI—the experimental group showed a decrease in first fixation duration from baseline to second viewing.

Between group analysis (all images and 4 complex images) demonstrates a shorter first fixation durations within the experimental group as compared to the control group—demonstrating the experimental group was affected by the triage training and thus changed their fixation patterns upon less salient features: white space AOI (all images: $p = 0.0003$; 4 images: $p = 0.001$).

Fixation Total

Evaluation of fixation total on the white space AOI was useful in demonstrating two things. First whether or not the experimental group changes their fixation patterns, and are fixating on the AOI text boxes rather than the white space. Secondly, to determine if there was a change in behavior in the control group with respect to the white space AOI from baseline to second viewing.

What is important for this experiment is focusing on the gross changes from baseline to the second viewings. In the all images analysis, the control group consistently shows a decrease in fixation total from baseline to second viewing within all AOI text boxes, and an increase in white space AOI. On average, the control group decreased their fixation total times from baseline to second viewing for the AOI text boxes, and showed a statistically significant increase in viewing time ($p = 0.026$) from baseline to second viewing on the white space AOI (all images). Fixation total times for the experimental group, on average, increased from baseline to second viewing for the salient AOI's (circulation, breathing, airway). As expected, the white space AOI shows a significant decrease (all images: $p = 0.00001$; 4 complex images: $p = 0.000003$) within the experimental group from baseline to second viewing.

Analysis of all images and 4 complex images between groups demonstrates statistically significant differences for all second viewing AOI's including white space AOI when comparing the control group to the experimental group fixation total. This increase in the experimental group fixation total times is astoundingly even across all AOI text boxes, and a similarly significant difference can be seen in the white space AOI as well; the white space AOI decreases fixation total for the experimental group when compared to the control group. This clear representation of change in fixation patterns fully supports the stated hypothesis in which this experiment set out to test.

Conclusion

Overall, this study was able to provide support that training in the START triage program results in a change in naïve subjects assessment of triage images. Analysis of the data for fixation patterns supports the hypothesis that training would change the experimental group participants fixation patterns. The Fischer Exact tests performed did not result in statistically significant findings to support the hypothesis that the experimental group would demonstrate an increase in triage accuracy compared to the control group. However, raw data shows variation between the experimental and control groups that a change in accuracy patterns between groups might be achieved with larger sample sizes. This study would also benefit from an interactive rather than passive training lesson for both the control and experimental group. This would force participants to remain focused on the training, rather than possibly skimming through the PowerPoint slides presented to them. Future research would benefit from the use of different pictures; the images used focused on one victim at a time, while a more realistic situation would be to have many victims in the surrounding, which adds to the complexity of triage.

To further investigate triage decision-making, adding a stress response indicator to this study would be a good way to analyze if stress plays a role in decision-making. This stress response indicator could also help research whether or not triage situations cause stress and what this stress imposes on triage decision-making.

Lastly, the gaze tracker was not able to capture every fixation or saccade per participant, due to a multitude of complications. One major issue is calibration—if the participant moves too far away from the desired location for data collection, the BeGaze eye tracker was apt to stop tracking until the participant moved back into a more

desirable location. However, this was common among almost all participants—tracking ratios did not exceed 96%. For future studies, using a chin mount, forcing the participants to remain still and in constant focus, could easily solve this.

REFERENCES

- Critical Illness and Trauma Foundation. (2001). *Background*. Retrieved December 26, 2012, from START: <http://citmt.org/Start/default.htm>
- Critical Illness and Trauma Foundation. (2001). *START*. Retrieved December 26, 2012, from The Triage Tag: <http://citmt.org/Start/tag.htm>
- SensoMotoric Instruments GmbH. (2010). *BeGaze 2.4 Manual*. Boston, MA.
- START Triage. (2013). *START-Simple Triage And Rapid Treatment "The Race Against Time"*. Retrieved November 12, 2012, from Simple Triage And Rapid Treatment: <http://www.start-triage.com/>
- U.S. Department of Health & Human Services. (2013, March 6). *START Adult Triage Algorithm*. Retrieved August 16, 2012, from Radiation Emergency Medical Management: <http://www.remm.nlm.gov/startadult.htm>
- WHO. (2007). *World Health Organization*. Retrieved November 26, 2012, from Mass Casualty Management Systems: http://www.who.int/hac/techguidance/MCM_guidelines_inside_final.pdf
- World Health Organization Department of Child and Adolescent Health and Development. (2005). Manual for Participants. *Emergency Triage Assessment and Treatment (ETAT)*. Switzerland: WHO Press.
- Bichot, N. P., & Schall, J. D. (1999). Effects of similarity and history on neural mechanisms of visual selection. *Nature neuroscience*, 2(6), 549-54. doi:10.1038/9205
- Kocak, E., Ober, J., Berme, N., & Melvin, W. S. (2005). Eye motion parameters correlate with level of experience in video-assisted surgery: objective testing of three tasks. *Journal of laparoendoscopic & advanced surgical techniques. Part A*, 15(6), 575-80. doi:10.1089/lap.2005.15.575
- Leong, J. J. H., Nicolaou, M., Emery, R. J., Darzi, a W., & Yang, G.-Z. (2007). Visual search behaviour in skeletal radiographs: a cross-specialty study. *Clinical radiology*, 62(11), 1069-77. doi:10.1016/j.crad.2007.05.008
- Matsumoto, H., Terao, Y., Yugeta, A., Fukuda, H., Emoto, M., Furubayashi, T., Okano, T., et al. (2011). Where do neurologists look when viewing brain CT images? An eye-tracking study involving stroke cases. *PloS one*, 6(12), e28928. doi:10.1371/journal.pone.0028928

- Myles-Worsley, M., Johnston, W. a., & Simons, M. a. (1988). The influence of expertise on X-ray image processing. *Journal of experimental psychology. Learning, memory, and cognition*, 14(3), 553-7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2969946>
- Savelsbergh, G. J. P., Williams, a M., Van der Kamp, J., & Ward, P. (2002). Visual search, anticipation and expertise in soccer goalkeepers. *Journal of sports sciences*, 20(3), 279-87. doi:10.1080/026404102317284826
- Sowden, P. T., Davies, I. R. L., & Roling, P. (2000). Perceptual Learning of the Detection of Features in X-Ray Images : A Functional Role for Improvements in Adults ' Visual Sensitivity ?, 26(1), 379-390. doi:10.1037//0096-1523.26.1
- Yarbzl, A. L., & Haigh, B. (1967). Eye Movements & Vision. *Journal of Experimental Psychology: Human Perception & Performance*.
- Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & psychophysics*, 57(6), 787-95. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7651803>

Curriculum Vitae

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]