# **UREAP Exit Report**

# Using a Virtual Reality to Establish an Investigator Decision Making Paradigm

### **Russell Walton**

Supervisors: Dr. Heather Price & Dr. Andrew Park

### Introduction

Suspect interviews can lead to false confessions or erroneous conclusions by interviewers. The main objective of this research project was to determine if altering certain physical traits of the suspect changes the decisions an interviewer makes in an interview in a criminal investigation. To gather data, a virtual interview scenario application was designed and developed. Upon development completion, a pilot study was run utilizing the application. This study was designed as a pilot, however, with the interview application now complete, additional trials can be conducted in the future.

A virtual interview scenario was created in which specific physical characteristics of the suspect, specifically, skin colour and facial tattoos, were altered while keeping all other elements consistent. A 3D suspect character was utilized and suspect response dialog was recorded to increase the interviewer's immersion. At certain points during the progression of the virtual interview, the participants, in the role of interviewer, are presented choices between two question types. At these decision points, the trial participant can choose to ask a hard, close-ended accusatory question, or a soft, open-ended questions. This created a decision tree for the interview scenario. The application tracks the path the interviewer takes through this decision tree, as well as their judgments of suspect guilt before and after the interview. The participants rate the perceived guilt of the virtual suspect, on a scale from 1 to 6. Accompanying the application, the participants evaluated the truthfulness, intelligence, honesty, and credibility of the virtual suspect. Demographic information about the trial participants was collected as well.

New technologies and the increased reality of virtual environments provide opportunities for new methods of study in the field of criminology and forensic psychology. The concept for this project was developed by reviewing studies that utilized virtual scenarios in the field of criminology, and through the analysis of studies that investigated the effects that physical traits and interview questions choice have on investigator decision making.

There are a number of studies that used virtual reality to explore questions in the field of law enforcement. In one study, undergraduates complete a computer simulation in which they had to determine whether a male suspect who appeared on screen was holding a gun or a neutral object. The participants were more likely to mistakenly shoot (i.e., shoot an unarmed suspect) when the suspect was black than when the suspect was white (Plant & Peruche, 2005). This study also demonstrated that after extensive practice with the program this racial bias was eliminated both immediately after training and 24 hours later.

A 2014 meta-analysis of the effects of suspect characteristics on arrest, found that suspect characteristics are an important contributor to the odds of arrest. Race, gender, and ethnicity matter when investigating the correlates of arrest (Daniel, 2014). This has been established in arrest, we were interested if it holds true for interviews.

There is also evidence that virtual environments are well-received by law enforcement personnel. In another study, a virtual environment system was constructed to support police officers becoming more observant and confident at spotting indicators of anti-social behaviour. Participants positively received the virtual environment, with the majority indicating a preference for virtual environment simulations over conventional paper-based activities (Smith & Carter, 2010). There was evidence that the virtual environment system was suitable for police-based practitioners as this group performed well and rated the virtual environment highly in the post-session questionnaire.

Multiple and different forms of evidence are presented in each case throughout an investigation or during an interview. In addition, the order of the presentation of the evidence can also influence an investigator decision. The first piece of evidence presented can decrease the influence of all following evidence (Charman, 2013; Price & Dahl, 2014). This is called the confirmation bias. As well, in other situations a recency effect can occur where the last piece of evidence presented strongly influences the investigator decision.

Our study utilized a decision tree within the virtual interview scenario application, wherein participants could choose to ask questions of hard or soft type. As part of this choice, participants were exposed to evidence about how the suspect's vehicle could potentially link them to the crime, and information about the suspect's presence in the area of the crime. Depending on the participant's choices, the order in which these two pieces of information were presented was determined.

Since this research created a virtual reality interview program, understanding the police interview processes was crucial. Obtaining complete, accurate and reliable information from suspected offenders is central to any criminal investigation, and the police interview is key to obtaining such information (Oxburgh et al., 2014). Investigators can use soft, open-ended questions or hard, accusatory close-ended questions during an interview. Adopting a free narrative approach may encourage the suspect to give a detailed and accurate account, however, closed questions are by far the most frequently used (Leahy-Hardland & Bull, 2016;

Westera et al., 2016). Although used less in practice, the open-ended, information-gathering approach has overwhelming empirical support, enables the gathering of more accurate information and has a lower probability of producing false confessions since suspects are exposed to significantly lower levels of psychological pressure (Areh, 2016). Training in specific interrogation methods is strongly associated with usage (Cleary & Warner, 2016). Virtual environment technology can provide a safe and controlled environment to assist police officer training (Power, 2011; Smith & Carter, 2010).

# Method

### Participants/Design

Undergraduate psychology students from Thompson Rivers University were invited to participate in the project as mock investigators. Participating psychology students were given a percentage boost to their final mark in one of their psychology courses for participating in the study. The trial participants were between the ages of 18 and 30, and both males and females participated.

The project was designed with three main research questions in mind:

1. Will the investigator's question choice, soft open-ended (information gathering) versus hard closed (accusatorial), be affected by a suspect's skin colour?

- 2. Will skin colour of the suspect affect the investigator's decision of guilt or innocence?
- 3. Will tattoos on suspects affect the investigator decision of guilt or innocence?

Our hypothesis was that a suspect's skin colour, and whether or not they have tattoos, would affect interviewer question choice and the interviewer's judgements of guilt or innocence. We expected a suspect to be judged as more guilty, and to be asked questions that were harder and more accusatory if they had dark skin or tattoos.

Participants were randomly assigned to one of four treatment groups in a 2 by 2 design. Condition 1 suspects have a white skin colour and no tattoos, condition 2 suspects have white skin and tattoos, condition 3 have dark skin and no tattoos, and condition 4 have dark skin colour and tattoos. The participants asked the virtual suspect a series of questions, were presented with evidence, and asked to rate the guilt of the suspect both before conducting the interview and after the interview's completion.

Following Connolly, Price, Lavoie and Gordon (2008), the mock investigators were also asked to rate the suspects on credibility, accuracy, and honesty. This evaluation takes place as part of the survey after the participant completes the virtual interview using the developed application.

A virtual interview scenario application was developed as part of this project. The application required participants to select a start guilt rating, reflecting their judgment of the suspect's guilt after the participant was shown a description of the scenario. This scenario description included information on the crime that was committed, and the reason that the suspect had been brought in for an interview. When this start guilt rating was determined by a participant, they have not yet seen the suspect. As well, the description of the scenario were the same for all four conditions. Once this start guilt rating had been chosen, the virtual interview began, and the virtual suspect was displayed to the participant.

The virtual interview progressed through a series of scripted interview blocks and decision points. The structure of the virtual interview is displayed in Figure 1. Each scripted interview block consisted of between 4 to 10 sections of scripted lines for the participant to read outloud, and corresponding responses from the suspect. The participant was presented lines, displayed as text on the screen, which they read outloud to the virtual suspect. After the participant finished reading a section of lines, they clicked a "Continue" button, and the virtual suspect responded to the participant. The virtual suspect's responses were pre-recorded, and the audio responses played audibly to the participant.

There were two points during the virtual interview when a participant was presented with a decision. At these decision points, the trial participant chose to ask a hard or soft question. The choice the participant makes determined the next scripted block that was presented by the virtual interview application. This structure of scripted blocks and decision points created a decision tree for the interview scenario. The application tracked the path the participant took through this decision tree.

Regardless of the path the participant took through the interview decision tree, the interview scenario always concluded with the same ending scripted interview block. During the scripted ending, the virtual suspect asked for a lawyer, and the interview scenario concluded. Once the interview had concluded, the participant was again asked to rate the suspects guilt. This end guilt was also logged by the application, and the application terminated upon end guilt selection.



A script was developed for the virtual interview scenario. Before the script was written, ten transcripts of real suspect interviews were studied, and the recordings of additional interviews were evaluated. For the sake of participant immersion, we considered it of value that the script be based in reality as much as possible. The final script that was developed and used for the virtual interview application was based mainly on two real interviews, the interview of Russell Williams by Detective Sergeant Jim Smyth, and the interview of Philip Markov by Detective Dennis Harris and Detective Robert Kenny. Sections of these interviews were incorporated into the script, and then adapted to fit the coherent scenario presented by the interview application. It should be noted that though the script was inspired by these interviews, the finished script was a unique piece of writing created for this research project.

The virtual suspect was created using Morph3D's MCS Male. MCS Male is a pre-made 3D male model whose features can be modified relatively easily. The suspect model was created to be as "average" as possible. The virtual suspect was not especially attractive in appearance, nor particularly unattractive. The suspect was not overweight, nor skinny, and neither tall nor short. The virtual suspect's eyes were brown, a racially ambiguous eye colour. The suspect's nose was not short or long, and not wide or skinny. The suspect's eye shape was also intentionally racially ambiguous. The idea was to create a 3D character whose common physical features would be suitable with either dark or light skin colour, and would influence the participant's judgement as little as possible.

The virtual suspect character's responses were voiced by a male voice actor. A voice actor was chosen who did not have a strong regional accent. As well, the actor's voice did not indicate specific racial traits. In timbre, the actor's voice was not especially high nor especially low.

### Materials

A custom application was built for use in the project trials. The virtual interview scenario application was developed using the Unity game engine. Custom functionality was coded using the C# programming language. A number of 3rd party Unity add-on were evaluated for use in the application. These included pre-built 3D models, such as Morph3D's MCS Male, the 3D character model that was used for the completed application. As well, a number of animation add-ons were evaluated for 3D character lip sync. These add-ons include LipSync Pro and SALSA with RandomEyes. SALSA with Random Eyes was used for the completed application due to it's ease of integration with the MCS Male character model, and accuracy of animation triggering. Dialogue engine add-ons were also evaluated for project use, these included Dialogue System, Extra Dialogue, and THE Dialogue Engine. These dialogue engines were evaluated for use in structuring the interview's progression through the script and implementing the decision tree. However, the completed application utilized none of these 3rd party dialogue systems due to the lack of ease adaptability of these products. Instead, a custom event system was developed to progress through the script, decision tree and to trigger the virtual suspect's audio responses.

All of the virtual suspect's scripted response lines were read by a voice actor and recorded using Avid's ProTools. These responses were stored as individual audio files, and triggered by the application after a participant, in the role of investigator, had completed reading a section of their lines and clicked a button to signal that they had completed asking the virtual suspect a question. Having these recording responses, voiced by an actor, was intended to increase the participants' immersion in the virtual interview.

### Procedure

Participants arrived at the lab and were greeted by a research assistant. Participants were managed one-on-one with a research assistant, who gave them preliminary information about trial completion, and had them sign a consent form. Participants were then led to a private room that contained a laptop computer with the virtual interview scenario already loaded, and the participant's assigned scenario (1-4) pre-selected. Although the room was private, the door to the participant's room was left ajar so that the research assistant could monitor the session from the adjacent room without being intrusive.

The total trial time for each participant was approximately 30 minutes. This trail time included the introduction by the research assistant, the participants progression through the virtual interview using the computer application, the post-interview questionnaire, and the debriefing.

Data was collected by the application during execution, and in the written survey. The computer application collected the participants start guilt and end guilt ratings, and the choices made the the decision points. The survey collected participant demographic information, including age, sex, background, judgments of the virtual suspect's honesty, credibility, and accuracy. As well the questionnaire asked participants for report what they felt worked during the interview, what didn't work, and provided a section for any additional info the participant wanted to report.

# Results

### Condition and StartGuilt Predicting Decision 1

A binary logistic fixed regression analysis was used to determine if the condition (white skin with tattoo, white skin without tattoo, dark skin with tattoo, dark skin without tattoo) affected the decision to ask soft or hard questions to the suspect at the first decision point. Start guilt was also entered into the analysis to control for pre-interview differences based on the participant reading of the scenario. As the factors were entered concurrently, any significant effects would represent independent contributions as the fixed regressions controls for the effects of all of the other factors.

As can be seen in the Table below, condition or start guilt did not predict the decision to ask hard or soft questions (all p's > .05).

#### Table 1

Variables in the Equation							
		В	S.E.	Wald	df	Sig.	Exp(B)
Step 1ª	Start_Guilt	.624	.477	1.712	1	.191	1.867
	Condition			2.777	3	.427	
	Condition(1)	1.728	1.063	2.642	1	.104	5.628
	Condition(2)	.702	1.075	.427	1	.513	2.018
	Condition(3)	.930	1.103	.712	1	.399	2.535
	Constant	-4.225	2.332	3.283	1	.070	.015

Variables in the Equation

a. Variable(s) entered on step 1: Start\_Guilt, Condition.

Figure 2 displays the count of soft or hard questions chosen at the first decision point for the four conditions. The first decision point occurred after the introduction scripted block of the interview.



### Figure 2

Effects of Condition on End Guilt

A four condition (white skin with tattoo, white skin without tattoo, dark skin with tattoo, dark skin without tattoo) repeated measures ANOVA with two time intervals (start guilt, end guilt) was used to examine possible effects of condition on end guilt. As can be seen in Table 2, the condition by time interaction was not significant (p = .900) and so were no differences in condition over time.

Tests of Within-Subjects Effects							
Measure: MEASUR	RE_1						
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
time	Sphericity Assumed	1.512	1	1.512	3.692	.063	.093
	Greenhouse-Geisser	1.512	1.000	1.512	3.692	.063	.093
	Huynh-Feldt	1.512	1.000	1.512	3.692	.063	.093
	Lower-bound	1.512	1.000	1.512	3.692	.063	.093
time * Condition	Sphericity Assumed	.237	3	.079	.193	.900	.016
	Greenhouse-Geisser	.237	3.000	.079	.193	.900	.016
	Huynh-Feldt	.237	3.000	.079	.193	.900	.016
	Lower-bound	.237	3.000	.079	.193	.900	.016
Error(time)	Sphericity Assumed	14.750	36	.410			92
	Greenhouse-Geisser	14.750	36.000	.410			
	Huynh-Feldt	14.750	36.000	.410			
	Lower-bound	1 <mark>4.7</mark> 50	36.000	.410			

#### Table 2

Table 3 shows the mean and standard deviation of the start guilt and end guilt ratings for the four conditions.

**~** · · · ·

Descriptive Statistics							
	Condition	Mean	Std. Deviation	Ν			
	1	4.00	.667	10			
	2	4.20	.919	10			
Start_Guilt	3	3.80	1.033	10			
	4	4.50	.527	10			
	Total	4.13	.822	40			
	1	4.30	1.059	10			
	2	4.60	.966	10			
End_Guilt	3	3.90	1.287	10			
	4	4.80	.789	10			
	Total	4.40	1.057	40			

#### Table 3

This project was conducted as a pilot study. Samples from 40 participants were collected. Given this limited sample size, the evaluation of the results should be considered preliminary.

### Discussion

Analysis conducted on the data collected for this pilot project indicated that suspect skin colour, and whether or not they had tattoos, did not have a significant effect on the type of question participant's chose to ask the suspect, or on the judgement of the suspect's guilt. This runs contrary to the hypothesis. However, due to the limited sample size of this pilot study, it is difficult to make any definitive determinations. Moving forward, additional trails will be conducted

that will allow for further, more accurate analysis. These findings can be explained by the demographics of the trial participants. Trial participants were are relatively young university students that study psychology. The participants' age, education level, and knowledge of psychology could affect trial results.

There are some limitations to using a virtual interview application. A virtual interview is not a real interview, and the questions that the participant's could ask the virtual suspect were controlled. In a real interview, an interviewer can ask a wide variety of questions, and can chose the exact wording they use. As well, a suspect's responses in real life are unpredictable and adaptable. However, even though the trials don't feature a real-life interview, people's biases are still present, and these biases still affect judgements.

The virtual reality interview built during this research will be used to conduct additional trials to increase the accuracy of the data and the strength of evidence. The data collected about interviewer characteristics, such as sex, age, or background, could be analyzed in the future to determine their effects on the results of the interview process. Once additional data has been collected, further analysis will be conducted in future projects.

### References

- Areh, I. (2016). Police interrogations through the prism of science Horizons of Psychology, 25, 18–28.
- Charman, S. D. (2013). The forensic confirmation bias: A problem of evidence investigation, not just evidence evaluation. *Journal of Applied Research in Memory and Cognition, 2,* 56-58.
- Cleary, H. & Warner, T. (2016). Police training in interviewing and interrogation methods: A comparison of techniques used with adult and juvenile suspects. *Law and Human Behavior, 40*, 270-284.
- Daniel, L. J. (2015). The effects of suspect characteristics on arrest: A meta-analysis. *Journal of Criminal Justice*, *42*, 589–597.
- Leahy-Harlandm, S. & Bull, R. (2016). Police strategies and suspect responses in real-life serious crime interviews. *Journal of Police and Criminal Psychology*, doi:10.1007/s11896-016-9207-8
- Oxburgh, G., Ost, J., Morris, P. and Cherryman, J. (2014). The Impact of Question Type and Empathy on Police Interviews with Suspects of Homicide, Filicide and Child Sexual Abuse. *Psychiatry, Psychology & Law, 21*, 903-917.
- Plant, A. and Peruche, M. (2005). The consequences of race for police officers' responses to criminal suspects. *Psychological Science, 16*, 180-183.
- Power, P. (2011). Enhancing the student learning experience through interactive virtual reality simulation. *Journal of Paramedic Practice, 3*, 447-449.
- Price, H. L. & Dahl, L. C. (2014). Order and strength matter for evaluation of alibi and eyewitness evidence. *Applied Cognitive Psychology*, *28*, 143-150.
- Smith, S. P. & Carter, T. (2010). A virtual environment to test police and public awareness of anti-social behaviour indicators. *International Journal of Police Science & Management*, *12*, 548-56.
- Westera, N., Kebbell, M., & Milne, B. (2016). Want a Better Criminal Justice Response to Rape? Improve Police Interviews With Complainants and Suspects. *Violence Against Women*, 22, 1748–1769.