Can biographical information bias decision-making when determining whether two people are siblings?

Brianna Do

This thesis is submitted in partial fulfillment

of the Honours degree of Bachelor of Psychological Science

School of Psychology

The University of Adelaide

October 2018

Word Count: 7,248

Table of Contents

List of Figures

List of Tables

Abstract

Previous literature has demonstrated that it is a difficult task for individuals to use faces to detect relatedness, regardless of whether it is a sibling or familial relationship. To make this task easier for the decision maker, it makes sense to provide them with some supplementary information. However, past studies also suggest that this additional information can bias the decision outcome. Therefore, the current study aimed to determine the impact of providing limited biographical information (such as name and date of birth) on human decision making. Using a within–subjects repeated measures design, participants (N = 49) were required to complete 144 sibling determination trials, where they had to look at two faces and determine whether they were siblings or not. There were three experimental conditions: (1) no biographical information, (2) correct biographical information, and (3) incorrect biographical information. The data were analysed using ANOVA and t-tests. Overall, sibling determinations were typically slowest, and made with less confidence and accuracy when incorrect biographical information was provided. However, when correct biographical information was provided, performance did not always significantly improve when compared with no biographical information given. Consistent with previous literature, the results also suggest that it is harder to declare that two people are siblings than it is to declare they are not. These findings can benefit a range of government agencies that have to make decisions based on facial imagery. Furthermore, with more diverse and publically available databases that are sibling-specific, future research in this field can be enhanced.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University, and to the best of my knowledge, contains no materials previously published except where due reference is made. I give consent for the digital version of this thesis to be made available on the web, via The University of Adelaide's digital thesis repository, the Library Search and through web search engines, unless permission has been granted by the School to restrict access for a period of time.

Brianna Do

October, 2018

Acknowledgements

Firstly, I would like to take this opportunity to say a massive thank you to my external supervisor Dr Rebecca Heyer, from the Defence Science and Technology (DST) Group. You have always seen my potential and believed that I can accomplish any goal I set my mind to. You are an incredible mentor and without you, this thesis would not have been possible and I am grateful for all your assistance and patience.

I would also like to thank my internal supervisor Dr Carolyn Semmler, from The University of Adelaide for being there for me. To my family, work colleagues, friends (both in and outside of psychology), I appreciate your endless unconditional love, positivity and support. I appreciate each and every one of you and I am grateful for all the memories that we have shared together throughout the year.

CHAPTER 1

Introduction

1.1 Rationale

Bias refers to when systematic errors compromise the authenticity of the decision making process (Hammersley & Gomm, 1997). Human judgement is subject to many different types of biases. We unconsciously pick up cues from our surrounding environment and factor them into our mental analyses on a daily basis (Croskerry, Sinhal & Mamede, 2013). Cognitive bias can be defined as the ways in which human perceptions and judgements can be shaped by extraneous information and external pressures (Kahneman & Egan, 2011). Cognitive bias can include: (1) contextual bias, (2) confirmation bias, and (3) avoidance of cognitive dissonance. Contextual bias occurs when people are influenced by potential irrelevant and misleading background or supplementary information (Dror, Charlton & Péron, 2006). If this supplementary information is correct, it can have a positive influence on the decision outcome, as people are able to be more confident in their decision and will be able to make correct decisions a lot quicker. However, if this information is incorrect, it can have a negative influence, as people will make more errors and take longer to make their decision. Confirmation bias refers to when people interpret information, or look for new evidence that helps them to conform to pre-existing beliefs or assumptions (Dror, 2018). Finally, avoidance of cognitive dissonance refers to when people are reluctant to accept new information that is inconsistent with their tentative conclusion (Frey, 1982).

Bias can have an impact on real life decisions in a range of different domains. The U.S. National Research Council of the National Academy of Sciences (2009) produced a report highlighting the weaknesses of a number of well-known forensic disciplines used in the criminal justice system. According to this report, one of the most serious, yet subtlest issues affecting the reliability of conclusions made in forensic science was cognitive bias (National Research Council, 2009). Following the release of this, there have been many other reports focusing on the different types of biases in the forensic sciences (Dror, 2018; Edmond, Tangen, Searston & Dror, 2014; Guarnera, Murrie & Boccaccini, 2017; Neal & Brodsky, 2016; The President's Council of Advisors on Science and Technology, 2016; Zapf & Dror, 2017). These studies all stress the need for cognitive psychologists to start working more closely with forensic scientists, in order to minimise as much bias as possible from impacting on one's decisions.

In real world domains, like passport processing, when visa and passport officers are making decisions about faces, whether they are matching faces or verifying biographical information, there is often supplementary information available that can affect their decision. Making decisions about faces, such as whether they match or not, is an error prone task (Megreya & Burton, 2008; White, Kemp, Jenkins, Matheson, & Burton 2014). Only a small amount of research has focused on contextual bias, and the impact of this supplementary information on face matching decision making within the field of forensic science (Heyer & Semmler, 2013; McCaffery & Burton, 2016). Heyer and Semmler (2013) reported on the impact of decision aids (such as match scores) within a facial recognition system on the decision maker's ability to decide whether faces matched or not. The decision aids were manipulated to sometimes suggest that a match was present when it was not, causing participants to make errors. The results showed that error rates were lower when no supplementary information (decision aids) was provided. A more recent study based on passport inspections also found that participants were poor at spotting errors in biographical information when asked to match pairs of faces presented with or without supplementary information (McCaffery & Burton, 2016). The study demonstrated that when faces were embedded in passport frames (providing the participants with supplementary information that could distract their decision making), the participant's ability to detect a face mismatch was reduced.

In addition to matching faces, government officials are sometimes required to make other judgements using facial imagery, such as determining relationships between people. Previous literature has demonstrated that it is a difficult task for individuals to use unfamiliar faces to detect relatedness, regardless of whether the relationship is a sibling or familial one (Kaminski, Dridi, Graff & Gentaz, 2009; Nesse, Silverman & Bortz, 1990; Park, Schaller & Van Vugt, 2008). A study conducted by Kaminski, Dridi, Graff and Gentaz (2009) investigated whether adults were capable of assessing relatedness of unrelated individuals using photographs, and whether visible facial cues varied according to the level of relatedness. The findings showed that adults were able to distinguish between individuals, however they were not outstandingly efficient at doing so, due to varying facial cues. Park, Schaller and Van (2008) explored the impact of heuristic cues and the inaccurate inferences that are drawn when making kin recognition. A signal detection analysis found that cue based recognition may have been biased due to false-positive errors causing implicit kinship inferences, regardless of the presence of non-kin relationships (Park, Schaller & Van, 2008). A group of researchers even looked at sex differences when judging one's ability to recognise family resemblance (Nesse, Silverman & Bortz, 1990). This study discovered that both women and men have equal ability to determine family resemblance and that this was not related to participant's age, martial status, number of siblings, number of children, or years of education. All of these studies confirmed that determining people's relationship to one another is a difficult task.

Therefore, to make this task easier for the decision maker, it makes sense to provide them with some supplementary information. However, this supplementary information has been found to bias decision making across a range of tasks including facial comparison (Heyer & Semmler, 2013), document examination (Stoel, Dror & Miller, 2014), firearms examination (Mattijssen, Kerkhoff, Berger, Dror & Stoel, 2016), latent fingerprint examination (Dror, Charlton & Péron, 2006), forensic anthropology (Nakhaeizadeh, Dror & Morgan, 2014), and familial determination (Bressan & Dal Martello, 2002). For example, a study looked at whether latent fingerprint experts would be affected by supplementary information when making a judgement (Dror, Charlton & Péron, 2006). Participants were to first make a judgement, without any contextual information, about whether the latent prints (from the crime scene) were a match with the print exemplar from the suspect, and then they were asked to repeat this process, but with context that suggested that the prints were not a match. This study provided evidence to suggest that even experts are prone to being misled by extraneous information, causing erroneous identifications to be made (Dror, Charlton & Péron, 2006). Another example of how supplementary information can influence decision making is from Nakhaeizadeh, Dror and Morgan (2014), who investigated whether providing contextual information would cause differences in interpretation and conclusions of skeletal remains amongst three groups: (1) male context, (2) female context, and (3) no contextual information at all. The researchers found that there was a strong confirmation bias when participants were able to access supplementary information about sex, ancestry and age at death (Nakhaeizadeh, Dror & Morgan, 2014). Another study conducted by Bressan and Dal Martello (2002) requested participants to estimate the facial resemblance of child-adult image pairs with either truthful or deceitful information about their relatedness. The findings indicated that participants were highly biased, as children were judged as more similar to their presumed parents than to presumed strangers, regardless of their true genetic relationship.

1.2 The Present Study

Passport eligibility and visa processing officers, as well as police officers and analysts working in government agencies have access to different types of supplementary information when they make determinations about identity and/or relationships using facial imagery (McCaffery & Burton, 2016). This supplementary information might not always be correct, due to data entry errors and/or identity fraud. Given that this may be the case, should people working within this field disregard such information and make their judgements purely based on facial imagery? What value does this supplementary information add? Can it bias their decision? Due to a lack of research in this area, the present study aims to determine if there is a significant impact on decision making performance when providing supplementary information (name and date of birth) to participants being asked to determine whether two people are siblings or not.

It is hypothesised that biographical information will have an impact on people's decision making when making a sibling determination. Here, the propensity of such supplementary information to bias the decision outcome will be further investigated between the different experimental groups. It is expected that when correct biographical data is provided, there will be improvements in performance (higher accuracy and confidence, and lower response latency) – inducing positive bias, whereas having incorrect biographical data will be detrimental to performance – inducing negative bias. The impact of providing no biographical information will also be explored, and these results will be compared and contrasted with instances when biographical information is provided. In line with previous work (e.g. Vieira, Bottino, Laurentini & De Simone, 2014), it is further hypothesised that it will be harder to declare that two people are siblings than to declare that they are not.

CHAPTER 2

Method

2.1 Ethics Statement

This study was approved both by the Defence Science and Technology (DST) Group Ethics Review Panel (approval number: NSID 04/2018) and the University of Adelaide Human Research Ethics Committee (approval number: 18/59).

2.2 Participants

The participants in this study consisted of 49 staff members from the DST Edinburgh site. The sample included 32 males and 17 females, with ages ranging from 21 – 80 years (M = 39.33, SD = 12.90). The majority identified as 'White/Caucasian' (92%), with a minority stating 'Asian' (6%), and 'Mixed European/Asian' (2%).

The inclusion criteria for participation in this study was: (a) proficiency in English, (b) aged over 18 years, and (c) normal or corrected-to-normal vision. Staff members were recruited via poster advertisements (Appendix A) placed around popular areas at the DST site and a news item appeared on the DST's internal media platform (SATURN) seeking volunteers. Incentives were not provided for participation; however, participants were compensated with snacks and water for their time.

2.3 Design

A within–subjects repeated measures design was employed for this computer-based study. The three experimental conditions in this study were: (1) sibling determination without any biographical information, (2) sibling determination with correct biographical information, and (3) sibling determination with incorrect biographical information. In condition (1), participants were presented with two images on the screen without any biographical information and asked to determine whether they were siblings or not. In conditions (2) and (3), participants were still presented with two images on the screen but they were also given limited (either correct or incorrect) biographical information (first/last name and date of birth) in order to see whether this information influenced their decision on whether the two people in the image pair were siblings or not.

Each experimental condition included 48 image pairs. Of these 48 image pairs, 24 were siblings and 24 were non-siblings. There was an equal number of mixed gender (24 pairs) and single gender image pairs (10 female and 14 male). Although it would have been optimal to achieve a balance in the same gender pairs, this was unobtainable due to the availability of imagery in the database used in this study, SiblingsDB (Vieira et al., 2014). Hence, performance as a function of gender (male versus female) was not examined for this reason. Using the difficulty ratings provided with the SiblingsDB, the 144 image pairs were divided into three difficulty levels (easy, moderate, hard) and then randomly allocated to each experimental condition to enable a balance of difficulty across all conditions.

2.4 Measures

Descriptive statistics were used in order to compare generic performance measures and signal detection measures between experimental conditions. The generic performance measures included: (a) accuracy, (b) confidence, and (c) response latency. The accuracy measure referred to the percentage of correct sibling determinations. The confidence measure referred to the level of certainty the participants had when making their judgements. This was analysed via a rating scale with 10% increments, ranging from 0–100% with 0% being no confidence and 100% being absolute confidence. The response latency measure referred to the overall time that the participants took to complete each individual trial (or sibling determination), which was recorded from the onset of each image pair displayed on the screen to the point where the participants submitted their decision.

The signal detection measures included: (a) hit, (b) miss, (c) false alarm, and (d) correct rejection. The hit measure referred to when the image pairs were siblings and the participants identified them as siblings, whereas the miss measure referred to when the image pairs were siblings but the participants identified them as not being siblings. The false alarm measure referred to when the image pairs were not siblings but the participants identified them as siblings, whereas the correct rejection measure referred to when the image pairs were not siblings and the participants correctly identified them as not siblings.

2.5 Materials

2.5.1 Image Source. Due to the lack of publicly available facial image databases in general, there has been limited research involving human ability to detect relatedness, especially in siblings. In saying this, previous research with a database of sibling imagery, the SiblingsDB assembled in Italy, found that it is much easier to determine that two people are not siblings than that they are (Vieira et al., 2014). As the database was intended for research purposes only, a license was granted to the DST group in order to conduct the present study. It is also important to note that those who appeared in the database had already supplied informed consent.

2.5.2 Image Selection. Images from the SiblingsDB were pre-screened for inclusion in the study and rating data provided with the database was used to generate three groups of image pairs based on the difficulty levels found in a previous study (Vieira et al., 2014). Using these difficulty levels, the 144 image pairs were then randomly allocated to each experimental condition to enable a balance of difficulty across all conditions.

2.5.3 Biographical Information Generation. Biographical information was generated for this study using a fake name and date of birth generator available online (https://www.fakenamegenerator.com/). Names were generated specific to gender, the Italian name set was chosen, and the country option was set to Australia, to reflect names of people of Italian origin who may have settled in Australia (given that the database was from Italy). The names and dates of birth were then randomly allocated to images by gender (male names allocated to males, females to females). The resulting name, date of birth and image set was then presented to three raters who checked that the name and date of birth was acceptable given the image. Adjustments to name and year of birth were sometimes required. Adjustments made to names were done via a secondary random generation whereas adjustments made to age were determined by seeking consensus from an additional five raters. The average of their estimates was used to calculate the year of birth. A total of 12 sets (image and biographical details) had to be manually changed in this way, and of these 12 sets, two additional raters checked the suitability.

For the purposes of this study, biographical information in the correct condition included the same surname and credible dates of birth for sibling pairs and different surnames and improbable dates of birth for non-sibling pairs (Figure 1).

Correct Biographical Information
(Leading the participant towards making the correct decision)

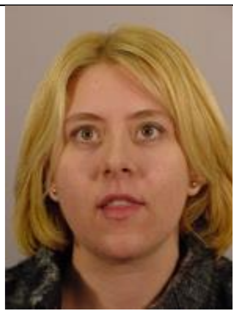
Siblings	 **Scott Summers** **12/11/1984**	 **Tim Summers** **04/09/1987**
Non-Siblings	 **Andrea Andrews** **21/07/1987**	 **Brianna Jones** **16/06/1987**

Figure 1: Siblings/non-siblings versus correct biographical information. Images for illustration purposes only (Phillips, Wechsler, Huang & Rauss, 1998).

The biographical information in the incorrect condition included same surnames and credible dates of birth for non-sibling pairs and different surnames and improbable dates of birth for sibling pairs (Figure 2).

Incorrect Biographical Information
(Leading the participant towards making an incorrect decision)

Siblings	 **Tamara Rose** **12/03/1991**	 **Sarah Winter** **14/04/1991**
Non-Siblings	 **Rebecca Smith** **23/11/1979**	 **Michael Smith** **24/12/1981**

Figure 2: Sibling/non-siblings versus incorrect biographical information. Images for illustration purposes only (Phillips, Wechsler, Huang & Rauss, 1998).

2.6 Experimental Application

This study was conducted on personal computers in the Biometrics Lab at the DST Group Edinburgh site using an application developed by the DST Group and images from the SiblingsDB (Vieira et al., 2014). The experimental interface included functions such as the ability to collect and store consent, answers to demographic questions, experimental data and registration for results.

2.7 Procedure

Participants were given a verbal briefing at the beginning of their allocated session. During this time, they were also given a copy of the Information Sheet (Appendix B), and the DST Group Guidelines for Volunteers (Appendix C). This allowed for participants to be primed about what they were required to do during the experiment and to have knowledge about the management of the collected data. Participants were encouraged to read the Information Sheet carefully and indicate their willingness to contribute to the study individually by clicking on the designated button on the first screen of the experimental application (Appendix D).

The initial screen of the experimental application required participants to answer some basic demographic questions such as age, gender, ethnicity, and questions regarding their vision. Practice tasks that mirrored the type of tasks that participants were asked to do during the experimental trials were performed. Before the practice and real tasks, participants were presented with a set of instructions that could be read on screen and this was also read aloud by the researcher before they commenced the activity. It was made clear to participants that there would be two images presented and it was their job to determine whether or not the people in the image pairs were siblings via selecting the yes or no option. They were also told to rate their confidence level on a scale with 10% increments, ranging from 0–100% when making those decisions and that there would sometimes be biographical information included underneath the images. Participants were also advised that the biographical information might not always be correct. Therefore, it was recommended that they try and make their judgements based on the image pairs only.

Participants were then told to take their time to begin the experiment. Presentation of the trials was randomised and counterbalanced by the experimental interface across all three experimental conditions. This meant that the participants were either shown an image pair without any biographical information, an image pair with correct biographical information, or an image pair with incorrect biographical information in no particular sequence. After the participants completed the 144 trials, they were given the opportunity to request their results via entering their email addresses into the experimental application.

CHAPTER 3

Results

3.1 Data Screening, Assumptions and Test Selection

Prior to performing statistical analyses, the data were screened for missing values and none were identified. The Shapiro-Wilk Test was then utilised to assess the dataset for normality. The results showed that skewness in the data ranged from -1.45 to 0.76, and the majority of the variables were normally distributed, as most of the variables produced a p-value < .001 (see Appendix E). Hence, the assumption of normality was met within this dataset so parametric tests that are appropriate for both a normal dataset and a within-subjects design were used for all analyses.

To compare the differences across the three experimental groups (no biographical information, correct biographical information and incorrect biographical information), an Analysis of Variance (ANOVA) test was used. To test for equality of variances, a Levene's test was conducted to check that the assumption had not been violated. Additionally, paired samples t-tests were conducted in order to assess the difference between two groups. The effect size of these tests was calculated using Cohen's *d*, whereby a small effect size = 0.2, medium effect size = 0.5, and large effect size = 0.8+ (Cohen, 1988). Furthermore, a Bonferroni correction was applied for all post-hoc Wilcoxon Signed-Rank Tests, reported at a .017 level of significance (two-tailed).

3.2 Overall Performance of Sibling Determinations

In order to determine the impact of biographical information on decision making, descriptive statistics for accuracy, confidence and response latency across all experimental groups were calculated. These are displayed in Table 1. The responses to sibling and non-sibling pairs were collated into 'overall' scores on each variable depending on which corresponding group they belonged to (see Table 1). The following analyses relate to differences across and between the three experimental groups with respect to these overall scores on accuracy, confidence and response latency. Signal detection measures are presented in Table 2 and analysed to further explore differences in accuracy between sibling and non-sibling pairs, across the three experimental groups.

Table 1

Descriptive Statistics for Accuracy, Confidence and Response Latency by Group

Statistics	Group 1			Group 2			Group 3		
	A (%)	C (%)	RL (s)	A (%)	C (%)	RL (s)	A (%)	C (%)	RL (s)
N	49	49	49	49	49	49	49	49	49
M	71.00	61.42	6.64	75.00	62.72	7.5	55.00	59.42	7.95
Median	71.00	59.58	6.59	75.00	60.42	7.67	58.00	58.75	8.07
SD	6.00	10.83	2.05	12.00	10.86	2.18	15.00	11.29	2.37
Variance	0.42	117.21	4.19	1.33	117.91	4.77	2.12	127.50	5.62
Minimum	56.00	43.96	3.36	50.00	46.25	3.63	10.00	39.79	3.51
Maximum	85.00	87.29	12.85	100.00	89.17	13.72	79.00	83.75	14.46

Note. Group 1 = sibling determination without biographical information. Group 2 = sibling determination with correct biographical information. Group 3 = sibling determination with incorrect biographical information. A = accuracy. C = confidence. RL = response latency.

The hypothesis predicted that there would be an impact on performance when biographical information was provided. It was expected that when correct biographical data was provided (Group 2), there would be improvements in performance, whereas providing incorrect biographical data (Group 3) would be detrimental to performance.

3.2.1 Accuracy. The results showed that accuracy was greatest when participants made sibling determinations using correct biographical information (Group 2), followed respectively by no biographical information (Group 1) and incorrect biographical information (Group 3) (see Table 1). An ANOVA test indicated that accuracy was significantly different across the three groups ($F(2,144) = 46.00$, $p < .001$, $\eta^2 = 0.39$). Paired samples t-tests revealed a large significant difference in accuracy between Groups 1 and 3 ($t = 7.42$, $p <.001$, $d = 1.06$), and Groups 2 and 3 ($t = 6.08$, $p <.001$, $d = .87$). A moderate, significant difference also existed between Groups 1 and 2 ($t = -2.50$, $p = .016$, $d = -.36$).

3.2.2 Confidence. Confidence was greatest when participants made decisions using correct biographical information (Group 2), followed respectively by no biographical information (Group 1) and incorrect biographical information (Group 3) (see Table 1). An

ANOVA test indicated that confidence was not significantly different across the three groups ($F(2,144) = 1.12$, $p = .33$, $\eta^2 = 0.02$). Paired samples t-tests revealed a moderate and significant difference in confidence between Groups 1 and 3 ($t = 3.64$, $p < .001$, $d = .52$), and Groups 2 and 3 ($t = 5.46$, $p < .001$, $d = .78$). There was no significant difference in confidence between Groups 1 and 2 ($t = -1.96$, $p = .06$, $d = -.28$). Thus, overall, confidence was lowest when participants made a sibling determination using incorrect biographical information, but correct biographical information did not significantly improve confidence compared to providing no biographical information.

3.2.3 Response Latency. Response latency was greatest when decisions were made using incorrect biographical information (Group 3), followed respectively by correct biographical information (Group 2) and no biographical information (Group 1) (see Table 1). An ANOVA test indicated that response latency was significantly different across the three groups ($F(2,144) = 4.49$, $p = .013$, $\eta^2 = 0.06$). Paired samples t-tests revealed a significant and relatively large difference in response latency between Groups 1 and 2 ($t = -6.09$, $p < .001$, $d = -.87$), and Groups 1 and 3 ($t = -8.87$, $p < .001$, $d = -1.27$). A moderate and significant difference also existed between Groups 2 and 3 ($t = -3.42$, $p < .001$, $d = -.49$). Thus, decisions were slowest, overall, when participants made a sibling determination with incorrect biographical information, faster with correct biographical information, and fastest when no biographical information was provided.

3.2.4 Signal Detection Measures. Table 2 presents the descriptive statistics for the signal detection measures across all experimental groups. The following exploratory analyses will only focus on the differences in correct rejection and hit rates between and across groups because the results for false alarm and miss rates can be logically inferred from these measures. Accuracy for non-sibling pairs is represented by correction rejection rates while accuracy for sibling pairs is represented by hit rates.

Table 2

Descriptive Statistics for Signal Detection Measures by Group

	Hit			Miss			False Alarm			Correct Rejection		
Statistic	G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
N	49	49	49	49	49	49	49	49	49	49	49	49
M	.61	.68	.41	.39	.32	.59	.19	.18	.32	.81	.82	.68
Median	.62	.71	.42	.38	.29	.58	.17	.12	.29	.83	.88	.71
SD	.17	.20	.21	.17	.20	.21	.13	.14	.20	.13	.14	.20
Variance	.03	.04	.04	.03	.04	.04	.02	.02	.04	.02	.02	.04
Minimum	.29	.29	.04	.08	.00	.12	.00	.00	.04	.42	.29	.08
Maximum	.92	1.00	.88	.71	.71	.96	.58	.71	.92	1.00	1.00	.96

Note. G1 = sibling determination without biographical information. G2 = sibling determination with correct biographical information. G3 = sibling determination with incorrect biographical information. Theoretically, the hit, miss, false alarm and correct rejection rates can range from 0–1. Higher scores indicate a greater rate.

3.2.4.1 Correct Rejection. The correct rejection rate was greatest when provided with correct biographical information (Group 2), followed respectively by providing none (Group 1) and providing incorrect biographical information (Group 3) (see Table 2). An ANOVA test indicated that the correct rejection rate was significantly different across the three groups ($F(2,144) = 11.52$, $p < .001$, $\eta^2 = 0.14$). Paired samples t-tests revealed a moderate and significant difference in the correct rejection rate between Groups 1 and 3 ($t = 5.16$, $p <.001$, $d = .74$), and Groups 2 and 3 ($t = 4.31$, $p <.001$, $d = .62$). However, there was no difference between Groups 1 and 2 ($t = -1.11$, $p = .273$, $d = -.16$). Thus, participants were better able to identify that non-sibling pairs were different people when there was either correct biographical information or none provided, in comparison to when they were given incorrect biographical information.

3.2.4.2 Hit Rate. The hit rate was greatest when correct biographical information was provided (Group 2), followed respectively by providing none (Group 1) and providing incorrect biographical information (Group 3) (see Table 2). An ANOVA test indicated that the hit rate was significantly different across the three groups ($F(2,144) = 24.44$, $p < .001$, $\eta^2 = 0.25$). Paired samples t-tests revealed a relatively large and significant difference in the hit rate between Groups 1 and 3 (t = 7.23, p <.001, d = 1.03), and Groups 2 and 3 (t = 6.52, p <.001, d = .93). A moderate and significant difference also existed between Groups 1 and 2 (t = -2.99, p = .004, d = -.43). Thus participants were better able to accurately identify sibling pairs when there was correct biographical information, followed by no biographical information and incorrect biographical information.

3.2.5 Summary. In terms of overall performance, the results supported the hypothesis. As expected, providing incorrect biographical information (Group 3) yielded the worst performance in terms of accuracy, confidence and response latency. Performance was better when participants were given correct biographical information (Group 2) than when asked to make a sibling determination with no biographical information (Group 1). However, providing participants with correct biographical information to help them make their decision only helped improve accuracy when determining whether two people were siblings, but did not make a difference when determining that two non-siblings were not siblings.

3.3 Performance Differences Between Sibling and Non-Sibling Pairs

Previous studies have suggested that it is a lot easier to determine that two people are not siblings than that they are (Vieira et al., 2014). In order to determine whether participants found it easier to determine if two people were siblings, or two people were not siblings in this experiment, descriptive statistics for accuracy, confidence and response latency for sibling versus non-sibling pairs were calculated. Table 3 displays these results.

Table 3

Descriptive Statistics for Accuracy, Confidence and Response Latency on Sibling and Non-Sibling Pairs

Statistics	Sibling Pairs			Non-Sibling Pairs		
	A (%)	C (%)	RL (s)	A (%)	C (%)	RL (s)
N	49	49	49	49	49	49
M	56.83	60.05	7.63	77.04	62.33	7.09
Median	56.94	59.31	7.64	79.17	60.42	7.2
SD	14.77	11.64	2.17	12.53	10.34	2.17
Variance	2.18	135.4	4.72	1.57	106.88	4.72
Minimum	26.39	35.97	3.56	37.5	45.69	3.22
Maximum	87.5	85	12.76	85	88.47	13.17

Note. A = accuracy. C = confidence. RL = response latency.

3.3.1 Accuracy. Accuracy was significantly better for determining that two people were not siblings, than determining that they were (t = -5.508, p < .001, d = -.79).

3.3.2 Confidence. Participant's reported levels of confidence were significantly higher when determining that two people were not siblings, than determining that they were (t = -3.2, p = .002, d = -.47).

3.3.3 Response Latency. Response latency was significantly greater when participants were determining that two people were siblings, than when determining that they were not (t = 4.38, p < .001, d = .63).

3.3.4 Summary. The results examining the difference in performance between sibling and non-sibling pairs in this experiment supported the hypothesis. It was found that individuals were better able to declare that two people were non-siblings than that they were siblings. In doing so, participants were significantly more confident and were able to make a non-sibling judgement a lot quicker than a sibling judgement.

CHAPTER 4

Discussion

The broad aim of this study was to investigate whether providing biographical information would impact human decision making when determining whether two people were siblings or not. It was also an opportunity to determine if it was harder for the decision maker to declare that two people were siblings than not. The results showed that the hypotheses in this study were supported.

4.1 Sibling Determinations

When visa and passport officers process applications, they often have access to supplementary information, other than facial imagery, to assist them in the task (McCaffery & Burton, 2016; Quigley-McBride & Wells, 2018). In this study it was hypothesised that biographical information would have an impact on people's decision making when making a sibling determination (better performance with correct information and worse performance with incorrect information). The results supported the hypotheses. Overall, when individuals were provided with incorrect biographical information (Group 3) they made more mistakes, were less confident and took longer to make their decisions than when presented with correct information (Group 2). An interesting finding was that performance was not always significantly better when participants were provided with correct biographical information (Group 2) than when asked to make a sibling determination with no biographical information (Group 1). That being said, the overall results concurred with previous research by Bressan and Dal Martello (2002), where the impact of supplementary information (labels) on judging familial relationships was investigated. The conclusions made from this study demonstrated that when resemblance is evaluated using faces, it is biased by the presumed relationship between the compared individuals (Bressan & Dal Martello, 2002).

There have been many studies that suggest the removal of this supplementary information would be beneficial in order to prevent contextual bias and the likelihood of miscarriages of justice, when making judgements (e.g. Dror & Charlton, 2006; Dror, Charlton & Péron, 2006; Dror & Cole, 2010; Dror, Péron, Hind & Charlton, 2005; Kassin, Dror & Kukucka, 2013). These previous studies have looked at supplementary information across a range of different decision making tasks. In particular, a similar study based on fingerprints, examined the effects of asking participants to make a match or non-match decision when pairs of fingerprints were associated with either crime-related photographs, neutral photographs, or no photographs. That study showed that when high emotion crime-related contextual information was provided, match decisions (and errors) increased, but there was no significant difference in performance between the neutral and no photograph groups (Osborne & Zajac, 2015). The suggestions from much of this research has been to remove such information from the decision making process. Therefore, is it really worth including biographical data when it does not necessarily improve or add value to individuals when a sibling determination is to be made? The results of this study suggest that it is not.

4.2 Sibling And Non-sibling Pairs

The overall performance between sibling and non-sibling pairs across accuracy, confidence and response latency was analysed. The results found that it was indeed easier for participants to declare that two people were non-siblings than it was to declare that they were siblings. Although there is little research on the detection of siblings in image pairs, the findings of this study support the main findings of Vieira, Bottino, Laurentini and De Simone (2014). These researchers found that computer algorithms for detecting sibling pairs could outperform human raters, as it is harder to differentiate facial features of siblings. Since it is easier to determine non-sibling pairs, individuals are able to be more confident and make their decision a lot faster in comparison to making a sibling determination.

4.3 Influential Factors

An influential factor that could have contributed to the performance of sibling determination was own-race bias. Own-race bias refers to when people are able to accurately recognise people of the same race in comparison to people of different races (Meissner & Brigham, 2001). This bias can increase confidence, however studies have only found a small significant effect that this is true (Wright, Boyd, & Tredoux, 2003). Previous researchers have suggested that when analysing faces of people from the same race, faster processing is involved (Chance & Goldstein, 1987; Valentine, 1991), enabling people to make quicker decisions.

As mentioned, the dataset used in this experiment consisted of people of Italian decent. In the study, the majority of the participants were White/Caucasian, thus the contact hypothesis could have affected the results. The contact hypothesis refers to when people form and maintain close and sustained contact with other people of different racial and ethnic groups, allowing them to differentiate facial features from their own race and that of others (Walker & Hewstone, 2006). It has been shown that by being surrounded by people of other races promotes positive, tolerant attitudes towards those groups (Powers & Ellison, 1995). Thus, because Australia is a very multicultural country, participants in this study may have already been exposed to people of Italian decent and may have not used the biographical information to help them make a sibling determination, as the Italian faces may have been familiar enough for them to just rely on the facial information.

4.4 Strengths

One of the strengths of this study is that it was one of the first to explore the impact of biographical information on relatedness decision making, using sibling and non-sibling pairs. Thus, conclusions drawn from this study can add value to the limited psychological research within the field of forensic comparison science to reduce the prevalence of errors in decision making. Few past studies have looked at how biographical information can influence people's abilities to make a face matching decision (e.g. McCaffery & Burton, 2016), but none had looked at how this contextual information could impact people's ability to make a judgement about a face in general. Therefore, another strength of this study is that it further reinforces the argument that supplementary information should be removed when people are making facial imagery judgements, whether it is a matching judgement or a determination on relatedness. This is because the conclusions drawn from this study are consistent with previous findings that suggest that there is no point including this supplementary information even if it is completely reliable and valid, as it can change the way in which people make decisions, and cause errors to be made (Bressan & Dal Martello, 2002; Dror, Charlton & Péron, 2006; Dror & Cole, 2010; Dror et al., 2005; Heyer & Semmler, 2013).

4.5 Limitations

There are some limitations that may have affected this study. One of the limitations involved how the experimental data were collected. Participants completed the experimental application under supervision, in one sitting, amongst other participants. Although this was a very convenient way to collect data in a timely manner, it would have been more effective for participants to be placed in a quiet, distraction-free environment, where they were unable to view the screens of other participants, or engage in conversations, and focus on the task at hand. Sibling determinations and participants' level of perceived confidence levels may have been influenced by those in close proximity. By placing individuals in separate rooms, participants who were anxiety ridden would have been able to alleviate the pressure of not being the last participant to complete the experiment.

Another potential limitation involved the sample size and representativeness of the recruited participants. There was only a small sample of 49 participants who contributed in this study. As mentioned, the participants were recruited from the DST Edinburgh site, and the majority of these participants had a tertiary education and background in scientific research. Thus, this limited the generalisability of the study's findings. To improve on the reliability and validity of future studies, a larger sample of participants from a wider population should be employed, including a sample of visa and passport processing officers.

The way this experiment was designed could have affected the results from this study and thus, could be improved. As each experimental condition had 48 image pairs, participants had to determine whether two people were siblings out of a total of 144 image pairs. This took participants, on average, an hour to complete. A limitation of this study was that there was no progress tracker. By including a progress tracker to see where participants are up to in this study would enable the prevention of fatigue and increase motivation to complete the study as accurately as possible.

4.6 Implications and Suggestions for Future Research

This study has significant implications for research about how contextual bias can affect human decision making. It is recommended that people working within government agencies should consider removing supplementary (contextual) information before they are to make a judgement. This will help to prevent any bias and will allow people working within this space to change or develop training policies to mitigate the influence of biographical information on human decision making. The findings also highlighted the importance of bias awareness training, as people might still insist on relying on this information to make their decision. The results provide further evidence to suggest that biographical information can be negatively influential without people realising. Even though the participants were told that this information would sometimes be incorrect, the results showed that, overall, participants did not always disregard the supplementary information.

As this was one of the first studies to identify that biographical information can influence human decision making when asked to make a sibling determination, it provides baseline understanding on this topic. Thus, future research is required in order for researchers to have a greater understanding and knowledge of how the different types of supplementary information can affect decision making, and hence are able to identify the factors that effect bias and improve accuracy outcomes.

Future studies should include a question regarding participants' use of supplementary information, in order to gauge participants' perspectives about how often they referred to the biographical information. This study should be followed up with an eye tracking study in order to objectively assess and compare these results with participants' perceived use of biographical information to confirm whether the information was used knowingly or not, and perhaps how much of the decision making time was taken up by looking at the biographical information versus looking at the faces.

Future research may also benefit from the development of a new publically available sibling database that consists of a range of ethnicities, genders and age ranges. By having a more diverse range of facial imagery, researchers would be able to investigate more about the own-race bias and contact hypothesis issues, as well as the impact of age, on sibling determination decisions. A study found that there were no discrepancies between men and woman when making a family resemblance judgement (Nesse, Silverman & Bortz, 1990). Hence, gender differences in the ability to recognise whether two people are siblings or not could also be explored. It would also be quite interesting to sample visa and passport processing officers who work in this field on a day-to-day basis and compare these results with the general population.

4.7 Conclusion

This study aimed to understand the impact of providing biographical information on human decision making when participants were asked to make a sibling determination. The results found that overall, sibling determinations were generally slowest, and made with less confidence and accuracy when incorrect biographical information (Group 3) was provided. It also showed that when correct biographical information (Group 2) was provided, sibling determinations were not always significantly better in comparison to not having any biographical information at all (Group 1). In line with previous research, this study further demonstrated that it is indeed harder for the decision maker to declare that two people are siblings than that they are not. The findings from this study can benefit a range of government agencies that have to make decisions based on facial imagery, suggesting that any supplementary information that may be available to assist decision making, may actually be changing the way the decision is made and, in some cases, increasing error. Developing more diverse and publically available databases that are sibling-specific, will allow more comprehensive research to be completed in this field.

References

Bressan, P., & Dal Martello, M. (2002). Talis Pater, Talis Filius: Perceived Resemblance and the Belief in Genetic Relatedness. *Psychological Science*, *13*(3), 213-218. doi: 10.1111/1467-9280.00440

Chance, J., & Goldstein, A. (1987). Retention interval and face recognition: Response latency measures. *Bulletin Of The Psychonomic Society*, *25*(6), 415-418. doi: 10.3758/bf03334728

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences.* New Jersey: Lawrence Erlbaum Associates.

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 1: origins of bias and theory of debiasing. *BMJ Quality & Safety*, *22*(Suppl 2), ii58-ii64. doi: 10.1136/bmjqs-2012-001712

Dror, I. (2018). Biases in forensic experts. *Science*, 360(6386), 243-243. doi: 10.1126/science.aat8443

Dror, I. E., & Charlton, D. (2006). Why experts make errors. *Journal of Forensic Identification,* 56, 600-616.

Dror, I., Charlton, D., & Péron, A. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International*, *156*(1), 74-78. doi: 10.1016/j.forsciint.2005.10.017

Dror, I., & Cole, S. (2010). The vision in "blind" justice: Expert perception, judgment, and visual cognition in forensic pattern recognition. *Psychonomic Bulletin & Review*, *17*(2), 161-167. doi: 10.3758/pbr.17.2.161

Dror, I., Péron, A., Hind, S., & Charlton, D. (2005). When emotions get the better of us: the effect of contextual top-down processing on matching fingerprints. *Applied Cognitive Psychology*, *19*(6), 799-809. doi: 10.1002/acp.1130

Edmond, G., Tangen, J., Searston, R., & Dror, I. (2014). Contextual bias and cross-contamination in the forensic sciences: the corrosive implications for investigations, plea bargains, trials and appeals. *Law, Probability And Risk*, *14*(1), 1-25. doi: 10.1093/lpr/mgu018

Frey, D. (1982). Different levels of cognitive dissonance, information seeking, and information avoidance. *Journal Of Personality And Social Psychology*, *43*(6), 1175-1183. doi: 10.1037//0022-3514.43.6.1175

Guarnera, L., Murrie, D., & Boccaccini, M. (2017). Why do forensic experts disagree? Sources of unreliability and bias in forensic psychology evaluations. *Translational Issues In Psychological Science*, *3*(2), 143-152. doi: 10.1037/tps0000114

Hammersley, M., & Gomm, R. (1997). Bias in Social Research. *Sociological Research Online*, *2*(1), 1-13. doi: 10.5153/sro.55

Heyer, R., & Semmler, C. (2013). Forensic confirmation bias: The case of facial image comparison. *Journal Of Applied Research In Memory And Cognition*, *2*(1), 68-70. doi: 10.1016/j.jarmac.2013.01.008

Kahneman, D., & Egan, P. (2011). *Thinking, fast and slow* (p. 10). New York: Random House Audio.

Kaminski, G., Dridi, S., Graff, C., & Gentaz, E. (2009). Human ability to detect kinship in strangers' faces: effects of the degree of relatedness. *Proceedings Of The Royal Society B: Biological Sciences*, *276*(1670), 3193-3200. doi: 10.1098/rspb.2009.0677

Kassin, S., Dror, I., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal Of Applied Research In Memory And Cognition*, *2*(1), 42-52. doi: 10.1016/j.jarmac.2013.01.001

Mattijssen, E., Kerkhoff, W., Berger, C., Dror, I., & Stoel, R. (2016). Implementing context information management in forensic casework: Minimizing contextual bias in firearms examination. *Science & Justice*, *56*(2), 113-122. doi: 10.1016/j.scijus.2015.11.004

McCaffery, J., & Burton, A. (2016). Passport Checks: Interactions Between Matching Faces and Biographical Details. *Applied Cognitive Psychology*, *30*(6), 925-933. doi: 10.1002/acp.3281

Megreya, A., & Burton, A. (2008). Matching faces to photographs: Poor performance in eyewitness memory (without the memory). *Journal Of Experimental Psychology: Applied*, *14*(4), 364-372. doi: 10.1037/a0013464

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy and Law*, 7(1), 3-35. doi:10.1037//1076-8971.7.1.3

Nakhaeizadeh, S., Dror, I., & Morgan, R. (2014). Cognitive bias in forensic anthropology: Visual assessment of skeletal remains is susceptible to confirmation bias. *Science & Justice*, *54*(3), 208-214. doi: 10.1016/j.scijus.2013.11.003

National Research Council (2009). *Strengthening Forensic Science in the United States: A Path Forward*. National Academies Press, pp.8, 19 - 20.

Neal, T., & Brodsky, S. (2016). Forensic psychologists' perceptions of bias and potential correction strategies in forensic mental health evaluations. *Psychology, Public Policy, And Law*, *22*(1), 58-76. doi: 10.1037/law0000077

Nesse, R., Silverman, A., & Bortz, A. (1990). Sex differences in ability to recognize family resemblance. *Ethology And Sociobiology*, *11*(1), 11-21. doi: 10.1016/0162-3095(90)90003-o

Osborne, N., & Zajac, R. (2015). An Imperfect Match? Crime-related Context Influences Fingerprint Decisions. *Applied Cognitive Psychology*, *30*(1), 126-134. doi: 10.1002/acp.3180

Park, J., Schaller, M., & Van Vugt, M. (2008). Psychology of human kin recognition: Heuristic cues, erroneous inferences, and their implications. *Review Of General Psychology*, *12*(3), 215-235. doi: 10.1037/1089-2680.12.3.215

Phillips, P., Wechsler, H., Huang, J., & Rauss, P. (1998). The FERET database and evaluation procedure for face-recognition algorithms. *Image And Vision Computing*, *16*(5), 295-306. doi: 10.1016/s0262-8856(97)00070-x

Powers, D., & Ellison, C. (1995). Interracial Contact and Black Racial Attitudes: The Contact Hypothesis and Selectivity Bias. *Social Forces*, *74*(1), 205. doi: 10.2307/2580629

Quigley-McBride, A., & Wells, G. (2018). Fillers can help control for contextual bias in forensic comparison tasks. *Law And Human Behavior*, *42*(4), 295-305. doi: 10.1037/lhb0000295

Stoel, R., Dror, I., & Miller, L. (2013). Bias among forensic document examiners: Still a need for procedural changes. *Australian Journal of Forensic Sciences*, *46*(1), 91-97. doi: 10.1080/00450618.2013.797026

The President's Council of Advisors on Science and Technology. (2016). *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (pp. 31 - 34). Washington, D.C.: United States Congress Press.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A, 43*(2), 161-204. doi: 10.1080/14640749108400966

Vieira, T., Bottino, A., Laurentini, A., & De Simone, M. (2014). Detecting siblings in image pairs. *The Visual Computer*, *30*(12), 1333-1345. doi: 10.1007/s00371-013-0884-3

Walker, P., & Hewstone, M. (2006). A perceptual discrimination investigation of the own-race effect and intergroup experience. *Applied Cognitive Psychology*, *20*(4), 461-475. doi: 10.1002/acp.1191

White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport officers' errors in face matching. *Plos ONE* 9(8), e103510. https://doi.org/10.1371/journal.pone.0103510

Wright, D.B., Boyd, C.E., & Tredoux, C.G. (2001). A field study of own-race bias in South Africa and England. *Psychology, Public Policy, and Law*. *7*(1), 119-133. doi: 10.1037//1076-8971.7.1.119

Zapf, P., & Dror, I. (2017). Understanding and Mitigating Bias in Forensic Evaluation: Lessons from Forensic Science. *International Journal Of Forensic Mental Health*, *16*(3), 227-238. doi: 10.1080/14999013.2017.1317302

Appendix A

Poster Advertisement

Do you think these two people are siblings?

We are looking for people to come and test their skill at detecting pairs of siblings (exactly what you did above). This is part of a study to help us understand how different methods of presenting information can impact on performance.

Details

Dates: XX June 2018

Location: Bay 16, 75L (Lab B7)

Session times: 0930, 1030, 1130, 1300, 1400, 1500 (for up to one hour)

What's in it for you?

- Up to an hour away from your desk, with refreshments
- A copy of your results, if you want them
- That buzz you get from contributing to research that will inform a range of government agencies that use facial recognition technology and humans just like you to process their output

Take a post-it note and contact me if you'd like further information or would like to participate:

[redacted]

This study has been approved by the DST Group Human Research Ethics Committee and University of Adelaide s School of Psychology Ethics Panel

Appendix B

Participant Information Sheet

UNCLASSIFIED

INFORMATION SHEET

SIBLING DETECTION USING FACIAL IMAGERY: HOW CAN WE MAKE THE TASK EASIER?

Brief description of the Study. Previous research examining human ability to detect relatedness (siblings, or other familial relationships) has determined that this is a difficult task. This study is being conducted by an Honours student in Psychology from the University of Adelaide (Ms Brianna Do), under the supervison of researchers from the Defence Science and Technology (DST) Group (Dr Rebecca Heyer and Dr Veneta MacLeod). It is designed to examine the impact of the different ways in which we can present information to people to assist them in this difficult task. Findings of this research are likely to benefit a range of government agencies that have to make identification or other decisions based on inspection of facial imagery, and can be used to inform training and development, as well as user interface and task design.

Your part in the Study. You will be asked to conduct a series of computer-based tasks where you will determine whether the two people on the screen are siblings or not. Sometimes there will be biographical data (a name and date of birth) included, other times there will not. Biographical data is not always reliable so please keep that in mind when making your decisions. These tasks will take you up to one hour to complete. Participation in the study is entirely voluntary; there is no obligation to take part in the study, and if you choose not to participate there will be no detriment to your career or future health care. You also have the right to withdraw at any time with no detriment to your career or future health care.

Risks of participating. There are no risks to your health or wellbeing as a result of participating in this study. Any occupational health and safety issues will be identified on site and appropriate measures will be taken to control risks to participants. Should you experience an adverse event during the conduct of the study, please let one of the researchers know, so that appropriate support can be provided (e.g., referral to first aid, counselling services).

Statement of Privacy. All data collected during the experiment will be treated in the strictest confidence and stored on password protected computers. Besides some basic demographic information and performance data, no other identifying information will be collected. However, you will also have the opportunity to receive a summary of the research findings by providing us with your email address at the conclusion of the experiment. Your email address will be stored separately to the other data you provide and linked by the researchers for this purpose only. All data you provide will be used for this project only, and stored on password protected computers accessible by the DST researchers and Honours student only. Once the data are no longer required they will be destroyed (within 2 years of data collection being completed). Results will be aggregated for reporting purposes to preserve anonymity.

Other relevant human research ethics considerations. In addition to receiving a copy of your own results, this research will be reported in the open literature in due course (Honours thesis and/or conference or journal publication). You have the right to withdraw your data at any time and can do this by contacting the lead investigator below and quoting the Unique ID number you will be assigned during the experiment. However, please note that while we can remove your data from the database at any time, once data analysis has been completed (approximately 6 months after data collection) we will not be able to re-analyse the data to exclude yours.

Consent. If you are willing to participate, please sign the consent form, and for data collection purposes also indicate this by clicking on the first screen of the experimental application, as instructed by the researcher.

[redacted]

Appendix C

Defence Science and Technology (DST) Guidelines

Appendix D

Experimental Application Consent Form

The primary aim of this study is to understand how difficult it is to determine whether two people are siblings or not by looking at their faces.

Participation in the study is entirely voluntary and there are no risks to your health or wellbeing as a result of participating in the study. All data collected during the experiment will be treated in the strictest confidence and stored on password protected computers.

To indicate your consent to participate, please click on the consent button below.

I CONSENT TO PARTICIPATING IN THIS EXPERIMENT

Appendix E

Testing Normality with Shapiro-Wilk

Variable	*W*	*P*	Skewness	Kurtosis
Accuracy				
Group 1 Overall	.98	.527	0.05	-0.03
Group 2 Overall	.97	.361	0.05	-0.74
Group 3 Overall	.90	< .001	-1.17	1.23
Hit				
Group 1	.96	.063	-0.18	-1.04
Group 2	.96	.093	-0.28	-0.88
Group 3	.97	.360	0.09	-0.88
Correct Rejection				
Group 1	.92	< .001	-0.90	0.31
Group 2	.87	< .001	-1.45	2.52
Group 3	.90	< .001	-1.17	1.3
Confidence				
Group 1 Overall	.94	.012	0.70	-0.24
Group 2 Overall	.94	.017	0.64	-0.52
Group 3 Overall	.96	.073	0.44	-0.55
Group 1 Siblings	.97	.169	0.36	-0.10
Group 2 Siblings	.98	.517	0.02	-0.40
Group 3 Siblings	.97	.224	0.40	-0.23
Group 1 Non-Siblings	.95	.034	0.73	0.11
Group 2 Non-Siblings	.94	.015	0.76	-0.09
Group 3 Non-Siblings	.97	.167	0.24	-0.16
Response Latency				
Group 1 Overall	.97	.248	0.48	0.17
Group 2 Overall	.98	.599	0.30	0.03
Group 3 Overall	.98	.718	0.28	-0.14
Group 1 Siblings	.97	.202	0.58	0.61
Group 2 Siblings	.98	.545	0.34	-0.08
Group 3 Siblings	.99	.905	0.24	0.07
Group 1 Non-Siblings	.97	.262	0.55	0.07
Group 2 Non-Siblings	.97	.330	0.54	0.86
Group 3 Non-Siblings	.98	.517	0.44	0.20

Note. *df* = 49 for all analyses. *SE* = 0.34 for all skewness output. *SE* = 0.67 for all kurtosis output.