University of Huddersfield Repository

Hollinshead, Jemma Louise

Administrator Effects on Respondent Choice

Original Citation

Hollinshead, Jemma Louise (2014) Administrator Effects on Respondent Choice. Doctoral thesis, University of Huddersfield.

This version is available at http://eprints.hud.ac.uk/id/eprint/24473/

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

http://eprints.hud.ac.uk/

Administrator Effects on Respondent Choice

Jemma Louise Hollinshead

A thesis submitted to the University of Huddersfield
in partial fulfilment of the requirements for
the degree of Doctor of Philosophy

July 2014

Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the "Copyright") and s/he has given The University of Huddersfield the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.

ii. Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the University Library. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.

iii. The ownership of any patents, designs, trademarks and any and all other intellectual property rights except for the Copyright (the "Intellectual Property Rights") and any reproductions of copyright works, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.

In memory of my mother, gone but not forgotten.

Acknowledgements

Firstly, I would like to acknowledge the support and guidance of my supervisors Professor David Canter and Dr Donna Youngs. I would like to thank them for allowing me to be a part of the International Research Centre for Investigative Psychology. The past four years have been extremely rewarding and challenging. The centre has provided me with fascinating tasks and invaluable opportunities, which have allowed me to grow both professionally and personally.

I would like to express my gratitude to my fellow members of the research centre. My fellow PhD students have not only provided me with an outlet for my frustrations; they have been a constant source of advice and reassurance. I would particularly like to acknowledge Dr Maria Ioannou and Dr Laura Hammond for their patience and encouragement.

I would also like to thank my family for their unwavering support and their conviction that this work would be completed. I would like to express my deepest gratitude to my partner for his continued emotional and financial support, for never allowing me to give up, and for boosting my confidence on occasions more numerous to mention. I apologise to you all for the times that I have been distracted, or stressed, or absent over the past four years.

Finally I would like to express my gratitude to all the people who gave up their valuable time to take part in my research, without them it would not have been possible.

Abstract

Research concerning administrator effects and the wider field of experimenter expectancy effects (Rosenthal, 1976) has established the theory that the hypotheses and knowledge held by an experimenter can unconsciously influence their results. Therefore, in a novel use of a photographic line-up from an actual police investigation, this research aims to explore the impact of administrator effects without the memory component of a recalled event. Previous research in this area has used a traditional memory paradigm to test administrator effects, however this has clouded the issue of whether the witness is being influenced by the administrator or is actually remembering the event. This research removes the memory component and therefore concentrates on the expectancy effect of the administrator.

In order to further the understanding of administrator effects, this research examines whether there is an aspect of interpersonal behaviour which predisposes some individuals to be more susceptible to inferences from others, or predisposes some to be more likely to influence individuals than others. In order to do this the Fundamental Interpersonal Relations Orientation: Behaviour (FIRO-B) instrument is utilised to examine the interpersonal relationship behaviour of the administrator and the participant. This research also identifies the cues emitted by the administrator by audio-recording the interaction between the administrator and the participant.

Using an experimental design, which manipulated the knowledge of the location of the target, five hundred and twenty six participants were asked to identify the person responsible for the Lockerbie bombing. Line-up administrators, who were either informed of the location of the suspect, informed of the location of an alternative suspect, or uninformed of the location of the suspect, presented the participants with the photographic line-up of twelve men, one of which is believed to be the person responsible for the Lockerbie bombing. Participants were asked to pick the person they thought was the suspect from the line-up, they then completed the FIRO-B questionnaire.

Analysis of the frequency of identifications suggests the presence of an experimenter expectancy effect. A chi-square goodness of fit test indicates that significantly more participants identified the target suspect when the administrator was informed of the location of the target, than when the administrator was uninformed. Analysis of the FIRO-B data found that target identifiers in the informed condition reported significantly higher received control scores than non-identifiers from an informed condition matched comparison group. Target identifiers also reported significantly higher social interactivity and received inclusion scores than non-identifiers.

Analysis of the FIRO-B data from the line-up administrators found subtle differences in the FIRO-B scores of the administrators achieving a high number of target identifications compared to administrators achieving a low number of target identifications. In particular, administrators achieving a high number of target identifications reported higher levels of

expressed control and lower levels of received control than administrators achieving a low number of target identifications. Analysis of the transcripts of the line-ups indicate that administrators in the informed condition interacted with their participants for longer, and exhibited more verbal cues. Administrators who obtained a target identification also spoke to their participants for longer. Those administrators who spoke to their participants for longer reported higher levels of expressed control and lower levels of received control.

The results of this study point to an experimenter expectancy effect. Beyond that though there appears to be an aspect of interpersonal behaviour that may be responsible for a predisposition to influence or to be influenced. This thesis, in line with previous research advocates the use of 'double-blind' line-up procedures in order to eradicate the possibility of an administrator effect. However, it also highlights the importance of considering the social interaction between the experimenter and the participant that is at the heart of social psychology research with human participants. In particular, the damning effect on the results of research conducted by an experimenter who assumes the dominant role in a social interaction, with a participant who assumes the submissive role.

Contents

Figures and Tables

Figures

Tables

Table of Appendices

Chapter 1: The Interpersonal Nature of the Social Psychology Experiment – The Role of the Experimenter

1.1. A brief history of the experimenter expectancy effect

Research concerning experimenter expectancy effects has a long history in Psychology, indeed the effect has been established in a number of social psychology studies. The term was coined by Robert Rosenthal following the research conducted by Rosenthal and colleagues during the 1950's and 1960's. It describes how the hypotheses of the researcher, through their unintentional behaviour towards their participants, may cause their hypotheses to be confirmed. However, the observation that the experimenter can have an effect on their data far precedes the establishment of the discipline of Psychology. Indeed, Sir Francis Bacon writing in the year 1620 highlighted the impartiality of the observer. He also predicted phenomena that are now recognised by cognitive psychologists as biased assimilation, selective attention, and confirmation bias (Risinger, Saks, Thompson, & Rosenthal, 2002). Whilst, nearly two hundred years ago astronomers observing stellar transit times, found that different times were being recorded by different individuals. As individuals used the same methods to measure these times, the difference it was realised must reside with the individuals making the observations (Saks, Risinger, Rosenthal, Thompson, 2003).

During the early years of the 20th Century the idea that the experimenter can affect their results was slowly seeping into researcher's consciousness. However, whilst researchers were reaching these conclusions separately, there had yet to be a concerted effort to combine the efforts of the separate researchers into one theory. One of these researchers Pfungst (1911) drew attention to observer effects by debunking the mystery of Clever Hans, the horse said to be able to solve arithmetic problems. Pfungst observed that rather than being a mathematical genius, Clever Hans was instead responding to the inadvertent cues exhibited by the questioners. Research in the 1920's (Roethlisberger & Dickson, 1939) attempting to study the productivity of workers under different conditions, was also an early example of the experimental procedure having a direct effect on the participants. In what became known as the Hawthorne effect it was found that receiving special treatment as part of the experimental procedure was the cause of the increased productivity, rather than the improvements made to the working conditions. Although the methodology of the

original experiment was subsequently criticised, the Hawthorne effect, as the knowledge that the experimental procedure can have unforeseen effects on the participants, has survived.

By the 1930's a young researcher (Rosenzweig, 1933) published what is now considered a seminal paper on the possible sources of error in the psychological experiment (Rosenthal & Rosnow, 1991). Firstly, he argued that the experimenter's attitude towards the participant and the participant's attitude towards the experiment can be a source of error. Secondly, twenty years before Orne would describe the 'good subject' effect, Rosenzweig cautioned that the motivations of the subject may lead to bias in the experiment. Finally, Rosenzweig argued that the characteristics of the experimenter and their verbal and non-verbal behaviours may affect the responses of the participants.

However, it was not until the 1960's that researchers began to take serious notice of experimenter effects. In the intervening years a vast body of research has accumulated in the fields of physical, biological, and behavioural science where experimenter effects have been noted. From the physical sciences there are numerous examples of scientists failing to see phenomena that do exist, for example Newton and the absorption lines in the prismatic solar spectrum. And further examples of scientists observing phenomena that do not exist, for example Rene Blondlot's N-rays. It is now widely accepted that the human being is not an objective observer, but rather that, their views, wishes and knowledge have an impact on the data they report. As Rosenthal and Rubin (1978) succinctly note, the experimenter expectancy effect can be described as:

> "the tendency for experimenters to obtain results they expect, not simply because they have correctly anticipated nature's response but rather because they have helped to shape that response through their expectations" (p. 377).

1.2. The nature of experimenter effects

Experimenter effects can be divided into two groups, firstly where the experimenter does not influence the subject's behaviour, and secondly, where the experimenter does influence the subject's behaviour. Rosnow and Rosenthal (1997) term these two groups as non-interactional bias and interactional bias respectively.

1.2.1. Non-interactional bias

In the first group, non-interactional bias, there are three types of experimenter effect to consider; observer effects, interpreter effects, and intentional effects. In the first group the experimenter is an observer and the effect in this group manifests itself as an error of observation, perhaps an overestimation or underestimation of some phenomenon. An example of observation error was found in research conducted in 1940. Berkson, Magath, and Hurn found that laboratory technicians counting blood cells only reported sample counts that were accurate 15-34% of the time.

The second type of non-interactional experimenter effect; the interpreter effect, involves an error when interpreting data. As the interpretation of data can be subjective it is harder to define an interpretative error than it is an observer error. It is also suggested that interpreter errors are less significant than observer effects (Rosenthal & Rosnow, 1991). Other researchers have the opportunity to re-interpret the data once it is in the public realm, however, it is impossible to return to the actual experiment and observe the subject again. The third experimenter effect; the intentional effect involves falsifying data in order to support hypotheses. Whilst research has suggested that undergraduate students may be susceptible to falsifying data in order to obtain the desired results (Weinstein, 1979), there are also examples of eminent scientists who have falsified their research (Rosenthal & Rosnow, 1991).

Non-interactional observer effects have also been found where the expectation of the observer has been manipulated. Two groups of identical planaria (freshwater flatworms) were observed for how many head turns and body contractions they made. In the first condition the observers were told that they should expect a high number of head turns and contractions. In the second condition the same observers were told to expect a low number of head turns and contractions. The observers identified twice as many head turns and three times as many body contractions in the first condition, when they were expecting a high number (Cordaro & Ison, 1963).

Recently attention has turned to the possibility of observer and interpreter effects within the forensic sciences. Indeed it is suggested that a number of factors can pervade the consciousness of the forensic scientist and lead to bias. Firstly, it is suggested that if an

examiner assumes the role of working for the prosecution, their perceptions of the evidence before them will be affected by their need to aid the prosecution's case. Indeed, it is suggested that individuals who fail to retain their impartiality may be motivated to bias their results in favour of the prosecution's case.

In what has been termed an "adversarial allegiance" (Murrie, Boccaccini, Turner, Meeks, Woods, & Tussey, 2009, p. 19), Murrie, Boccaccini, Guarnera, and Rufino (2013) suggest that it is not only forensic scientists who are susceptible to the biasing effects of working for the prosecution. Indeed, they argue that forensic psychologists and psychiatrists, when employed as expert witnesses, are susceptible to the same bias. Field studies of actual civil-commitment trials have found that experts scored their clients on a measure differently according to whether they were employed by the prosecution or the defence. Furthermore, the direction of the scores favoured the case of the employer, either the prosecution or defence. This has been found even with the Psychopathy Checklist-Revised (PCL-R) (Hare, 2003), which consistently demonstrates very strong inter-rater agreement (Murrie, Boccaccini, Johnson, & Janke, 2008; Murrie et al., 2009).

In addition to assuming a role for the prosecution, forensic examiners can also be susceptible to contextual effects. As Risinger et al., (2002) point out, often, samples sent to the examiner are accompanied by extraneous information about the case. This may include the investigators hypothesis that they have caught the culprit and just need the examiner to confirm their hypothesis. Or the investigator may provide the examiner with details of other evidence they have collected (Giannelli, Weatherhead, & Weatherhead, 2010). In such circumstances the presence of this extraneous information may inadvertently effect the judgements made by the examiner.

Notwithstanding the comprehension of the general population that DNA analysis is free from bias and interpretation, it is necessary in some cases for analysts to provide some interpretation of the data. For example when the DNA sample is degraded, when there is a small quantity, or when there is a mixture of DNA from two or more individuals, it is necessary for the analyst to interpret the data and provide some judgement. In these cases it is suggested that the interpretation of the evidentiary DNA sample may be influenced by the analysts knowledge of the suspects DNA profile. Furthermore, it is argued that

information about the suspect, or details about the investigation can become known to the scientist, including whether eyewitness identifications of the suspect have been made, or whether the suspect has confessed to the crime, and the possible motives of the suspect for committing the crime (Krane et al., 2008). The authors stress that knowledge of the suspects DNA profile and knowledge of the case can result in confirmatory bias, thereby calling into question the reliability of the analysis.

Fingerprint comparison is also a subjective judgement on the part of the analyst. Although automated fingerprint identification systems search databases and find comparable samples, the actual matching of a latent fingerprint from a crime scene, to a suspect's fingerprints is conducted by an expert. Real-life examples of mistakes by multiple forensic experts suggest that subsequent experts may be susceptible to confirmation and context bias if they know that the previous analyst has matched the latent print to the suspect (Dror & Cole, 2010). Therefore, if an examiner assumes the role of working for the prosecution, and knows that the police have who they believe to be the suspect in custody they may be susceptible to confirmation bias. Whereby evidence confirming the hypothesis is accepted and disconfirming evidence is rejected.

1.2.2. Interactional bias

In this instance it is the second group, the interactional bias group, where the experimenter directly influences the response or behaviour of the participant, in which we are particularly interested. In moving to study this group we also move from the physical and biological sciences to the behavioural sciences to consider the expectations of the experimenter. The effect of a person's expectations on another person's behaviour, the "self-fulfilling prophecy" was aptly labelled by the sociologist Merton (1948). His theory proposes that a person's expectations and predictions about another person's behaviour will cause the person to act in the predicted or expected way. Although the 'self-fulfilling prophecy' term was coined by Merton in the 1940's the concept had already been utilised in a clinical setting by the 1890's (Rosenthal, 1976). Since this time research concerning self-fulfilling prophecy has been conducted in a wide variety of experiments, including, reaction time, psychophysics, animal learning, verbal conditioning, personality assessment, person perception, learning, and ability (Valentine, 1992). Whilst much of this research purported

to show experimenter expectancy effects, as Rosenthal (1976) points out, the methodology of some of these studies may have shown the effect of confounding variables not the experimenter's expectancy. In response to this, studies by Rosenthal and colleagues have adopted a methodology more suited to measure the experimenter expectancy effect, whereby the variables of the experiment are kept constant and only the experimenter's expectancy is manipulated.

1.3. Evidence for the experimenter expectancy effect

The self-fulfilling prophecy theory can be seen in action in the Pygmalion experiments (Rosenthal & Jacobson, 1992), where teachers expected certain students to bloom academically those students did indeed bloom academically. In these now famous experiments, children of an American elementary school were given an IQ test. The teachers at the school were led to believe that the test would predict academic growth. Twenty percent of the students were randomly selected into the academic growth condition, the teachers were then informed which of their students had scored highly on the test and therefore were expected to bloom academically. When the children were tested with the same IQ test eight months later; those who had been indicated as academic bloomers had scored higher on the IQ test than those in the control group. This effect however, was first documented with animal subjects. Studies with rats (Rosenthal & Fode, 1963a; Rosenthal & Lawson, 1964), have found that those experimenters who are led to expect bright rats on a maze or in Skinner boxes did indeed report bright rats, and those led to expect dull rats reported dull rats.

In order to study experimenter expectancy effects under methodologically sound conditions Rosenthal conducted a series of experiments (Rosenthal, 1966). One such experiment was conducted by Rosenthal and Fode (1961), and then replicated by them in 1963(b). For the first experiment ten undergraduate experimental psychology students were recruited as experimenters. Each experimenter recruited between 18 to 24 participants, resulting in a subject pool of 206 student participants. The experimenters were required to conduct a person perception task by showing the participants ten photos and asking them to rate each photo on a scale from extreme failure to extreme success. Each photo had previously been rated as a zero, neither a success nor a failure. However, five experimenters were told that

their participants would average a -5 rating (moderate failure), whilst the other five experimenters were told that their participants would average a +5 rating (moderate success). Additionally, experimenters were told they would be paid $1.00 per hour, or $2.00 per hour if their results were as expected.

In order to control the variables of the experiment, experimenters were asked not to discuss the experiment with anyone, and not to say anything to the participants other than the instructions supplied to them. The results of the first experiment showed that the experimenters expecting successful ratings (+5) obtained higher ratings than those experimenters expecting failure ratings (-5). Due to the exceptional nature of these results, and the scepticism which greeted these results from journal publishers (Rosnow & Rosenthal, 1997), the researchers conducted a replication. In the replication, twelve undergraduate industrial psychology students served as experimenters, who between them recruited 86 participants. The replication followed the same procedure as the first experiment with one important exception. In the first experiment the experimenter handed each photo to the participant when asking them to make their ratings. However, in the replication the photos were mounted onto cardboard and were not handled by the experimenter. Again the results were exceptional; as in the first experiment all of those experimenters expecting positive ratings did indeed obtain positive ratings, and those expecting negative ratings obtained negative ratings. Therefore, even in the replication experiment, where the researchers had strived to reduce the interaction between the experimenter and subject, the same effects of experimenter expectancy had been found.

Reviewing these studies, Rosenthal (1976) indeed argues for the presence of an experimenter expectancy effect: “Since the experimenters had all read from the identical instructions, some more subtle aspects of their behaviour toward their subjects must have served to communicate their expectations to their subjects” (p.149). Nevertheless, there are a number of questions that were raised by this research. Firstly, Rosenthal (1976) questioned whether the results obtained were due to the monetary incentive provided to the experimenters to produce the expected results. Secondly, it was noted that in one respect the experiments were ecologically invalid, as the experimenters were only tasked with one condition of expectancy, either success or failure. An experiment conducted by Laszlo (cited in Rosenthal, 1976) covered these two areas. Firstly, the experimenters were

not offered an extra incentive to achieve the expected results, and secondly experimenters conducted both success and failure rating conditions. The results of this experiment were in-line with the previous two, when experimenters were expecting higher ratings they obtained higher ratings. However, in contrast to the two previous experiments, there was some overlap in Laszlo's data, in that some of those expecting positive ratings actually received negative ratings, and vice versa. Nonetheless, these results do appear to suggest an experimenter expectancy effect regardless of cash incentive or a mixture of expectancies.

The implications of the motivations of the experimenter have received further attention in the research literature. In a study designed to measure the effect of both experimenter and subject motivation, Rosenthal, Fode and Vikan-Kline (cited in Rosenthal, 1976), conducted an experiment similar to those detailed above. In this instance twelve graduate students were recruited as experimenters. They were required to conduct the photo-rating task, but were led to expect mean ratings from their subjects of +7. The experimenters were randomly allocated to a moderately motivated group (paid $2 per hour) or a highly motivated group (paid $5 per hour). The subjects, of which there were 58, were also randomly allocated into a moderately motivated group (unpaid), and a highly motivated group (paid 50 cents for their participation). Contrary to the hypothesis of the study the moderately motivated experimenters received the highest mean ratings and therefore the greatest expectancy effects, with the moderately motivated subjects. Whilst the highly motivated experimenters received the lowest mean ratings with the highly motivated subjects.

Early research in this area has also considered the characteristics of both the experimenter and the subject in an attempt to determine whether some experimenters are more likely to influence their subjects, and whether some subjects are more susceptible to influence. The biosocial attributes of the experimenter and subject, for example their gender was considered. Recent research with rodents suggests that the gender of the experimenter does indeed have an impact. Mogil et al., (2014, cited in "Lab mouse test results depend on scientist's gender,") found that lab rodents behaved differently according to the gender of the experimenter. In this study they did not exhibit the usual signs of pain when injected with an inflammatory agent, but only when male experimenters were present.

Nevertheless, with human subjects, contradictory research findings abound, with some studies reporting a difference between female and male experimenters (Binder, McConnell & Sjoholm, 1957; Sarason & Harmatz, 1965), and other studies finding no difference (Ferguson & Buss, 1960). Rosenthal (1967) suggested that the gender of the experimenters and the subjects are important variables to be considered in an experiment because of the differing behaviours male and female experimenters are found to exhibit. He found that male experimenters exhibit more friendly behaviour than female experimenters, whilst female subjects are smiled at more often than male subjects.

When considering whether male or female experimenters are more likely to influence their subjects a number of studies have been conducted with varying results. Rosenthal (1976) reports two experiments using the photo rating task where male experimenters demonstrated expectancy effects with both male and female subjects. Female experimenters demonstrated expectancy effects with female subjects, however, with male subjects female experimenters obtained ratings opposite to those expected. When the experimenter gender was held constant (only male experimenters) female subjects were more susceptible to the expectancy effect than the male subjects (Rosenthal, Persinger, Mulry, Vikan-Kline, & Grothe, cited in Rosenthal, 1976). Research then seems to suggest that male experimenters are more likely to influence their subjects, whilst female subjects are more susceptible to influence. However as Rosenthal (1976) highlights, whilst some studies have found male and female subjects to have no difference in their susceptibility to influence, they have never found male subjects to be significantly more susceptible to influence than female subjects.

Research conducted to test the Pygmalion hypothesis (increasing a teacher's expectation of a pupils performance will lead to an increase in that performance), outside of the classroom has consistently found the Pygmalion effect with male subjects in military and industrial settings. However, when male and female subjects are compared in a military setting, the Pygmalion effect was found for the male subjects but for the female subjects was dependent on the gender of the trainer. The research conducted by Dvir, Eden, and Banjo (1995) found that when the trainer was male there were statistically significant differences in the performance of male and female trainees in the Pygmalion condition compared to male and female trainees in the control condition. However, when the trainer was female

there were no statistically significant differences between female trainees in the Pygmalion condition and female trainees in the control condition.[1] This research in-line with the research reviewed above suggests that females are susceptible to the Pygmalion effect, but are not apt at delivering it.

Whilst research has considered readily assessable and stable characteristics of both the experimenter and the subject, studies have also examined those less readily assessable characteristics, namely psychosocial attributes, including anxiety, need for approval, hostility, authoritarianism, intelligence, dominance, status, and warmth (Rosenthal, 1976). It has been suggested that experimenters who exude differences in these attributes elicit different responses from their subjects. However, in a review of seven studies of experimenter and subject anxiety Rosenthal (1976) found that levels of anxiety in the experimenter and subject are significantly related to expectancy effects but not in a discernible pattern. Low, medium and high levels of anxiety in both experimenters and subjects are found to be significantly associated with expectancy effects. The effect of the status of the experimenter was directly tested by Vikan-Kline (cited in Rosenthal, 1976). The experimenters, either graduate students (lower status) or faculty members (higher status), were required to try to influence their subjects to give ratings of success or failure on the photo-rating task. Overall the higher status experimenters were more influential than the lower status experimenters. However, when the results were analysed in terms of chronology, they found that lower status experimenters were more influential than higher status experimenters with the first half of subjects. This relationship was reversed for the second half of subjects.

In addition to the characteristics of the experimenter, the situational factors of the experiment are also considered, for example the experimenter's prior acquaintanceship with the subject. Again the research in this area is not definitive. A study by Sacks (1952) found that not only did prior contact between the experimenter and the subject increase the subject's performance on an IQ test, but the warmth of that contact was also important. It is widely thought that the prior contact serves to reduce the anxiety felt by the subject, thereby improving their performance on the task. However, further research has found

[1] There was no condition of female trainers training male trainees. The authors highlight the need for research in this area.

contradictory results, Stevenson, Keen and Knights (1963) state that strangers were more influential experimenters than those known to the subject. Rosenthal (1976) argues that it may be the type of experiment that mediates the effect of the relationship between the experimenter and the subject. Where the task is more onerous for the subject, such as an intelligence task, as in the Sacks (1952) study, there may be an increase in anxiety levels, resulting in a greater effect of prior contact. Whereas with a more simple motor task, as in the Stevenson et al., (1963) study, there is less anxiety and therefore less of an effect of the relationship between the experimenter and subject. When considering expectancy effects, research has suggested that prior acquaintanceship between the experimenter and subject leads to expectancy effects (Rosenthal, Persinger, Vikan-Kline, & Mulry, cited in Rosenthal, 1976). However, this relationship is not so straightforward when the gender of the experimenter is taken into account. Male experimenters were more influential with acquainted subjects, whereas female experimenters were more influential with unacquainted subjects (Persinger cited in Rosenthal, 1976).

Further situational factors have been studied in this respect, including the experimenter's experience of conducting such experiments, and how the behaviour of the subject during the experiment affects the behaviour of the experimenter. Research has firstly suggested that participants respond differently in experiments if their experimenter is experienced in conducting the experiment. Furthermore, it has been observed that the experimenter's behaviour is influenced by the responses obtained from their participants. If the first group of participants respond in a manner consistent with the hypothesis then the experimenter will change their behaviour to successive participants, thereby influencing those participants to also respond in the same manner (Rosenthal, 1976).

Due to the pervasive nature of experimenter expectancy effects and the vast amount of research that this topic has spurned from areas as diverse as education and the workplace, Rosenthal and Rubin (1978) conducted a meta-analysis of the research conducted. The body of the meta-analysis consisted of 345 studies of interpersonal expectancy effects conducted in eight research areas; reaction time, inkblot tests, animal learning, laboratory interviews, psychophysical judgments, learning and ability, person perception, and everyday situations. The aim of the meta-analysis was to determine whether the first 345 studies were actually measuring an experimenter expectancy effect at a greater than chance level, or whether the

345 studies constitute a biased sample. The results of the meta-analysis found that in all of the eight research areas a much greater proportion of studies achieved statistical significance than would be expected by chance. Furthermore, it was calculated that non-significant results from 65,000 studies would be required to reduce the probability values to levels of significance that would not be acceptable. Overall, analysis of probability values and effect sizes found unwavering evidence for the presence of experimenter expectancy effects. Indeed, a contemporary of Rosenthal, Rosenberg (1980) suggests that the meta-analysis indicates that the expectancy effect is "replicable ... it operates over a broad range of substantive research areas [and is] capable of contaminating research procedures of the non-experimental as well as the experimental variety" (p. 473).

1.4. The communication of an experimenter's expectancy

Indeed from the beginnings of research concerning experimenter effects conducted in the laboratory and the classroom, the concept has been extended to fields as diverse as the judicial system and the healthcare system (Rosenthal, 2003). Research has evolved from studying the experimenter-subject and teacher-pupil relationship to also considering the judge-juror, doctor-patient, and manager-employee relationships. Furthermore, as the research field has expanded so has the understanding of the "covert communication" (p. 151), the means of transmitting the expectations of the experimenter to the subject. Early research, attempting to explain how experimenters communicate their expectancies to their subjects, involved experimenters intentionally trying to influence their subjects. It was hoped that the cues used by experimenters would be the same overstated cues used in an unintentional condition. In a replication of the photo-rating task described above, where the experimenter tried to influence their subjects and observers tried to discern the experimenters' expectancy, it was found that observers are very accurate at discerning experimenters' expectancies (Rosenthal, 1976). The observers of this experiment suggested two modes of expectancy communication, a visual-kinesic mode and an auditory-paralinguistic mode.

Research was then conducted to try to determine the importance of these two modes of communication (Rosenthal & Fode 1963b). A variation of the photo-rating task described above, where the photos were mounted onto card and not handled by the experimenter

was used. Eighteen experimenters were led to expect ratings of -5 for the photos and were randomly assigned to three experimental conditions; visual only condition, auditory only condition, and visual and auditory condition. In the visual only condition the experimenters handed the subjects written instructions and did not speak to them during the photo-rating task, although they remained in view of the subject. In the auditory only condition the experimenters read the instructions to the subjects but sat behind a screen during the photo-rating task. In the visual and auditory cues condition the experimenters read the instructions to the subjects and remained in their view for the photo-rating task. Rosenthal and Fode (1963b) found differences in the expectancy effects for the three conditions. There was no expectancy effect in the visual only condition; but there was a significant expectancy effect in the auditory only condition. However, the visual and auditory condition showed the most significant expectancy effect. The authors therefore conclude that auditory cues are more important for communicating expectancy than visual cues, although the most effective method appears to be a combination of visual-kinesic and auditory-paralinguistic modes.

Rosenthal (2003) has furthermore demonstrated the process of experimenter expectancy effects by three linked variables. Firstly there is the expectancy for the behaviour of the other person (E). Secondly, there is the communication or mediating variables (M). Thirdly, there is the response of the person for whom there are expectations (O). The link between E and M explains the effect of the expectancy on the mediating communication variables, whilst the link between M and O describes how the mediating communication variables affect the outcome variable. The E-M link and the M-O link described above has been used to explore judges expectations of guilt, how this expectation affects their instructions to jurors and how jurors decisions about the defendants' guilt are affected by the instructions (Halverson, Hallahan, Hart, & Rosenthal, 1997). It is argued that in their capacity as arbiters of criminal justice proceedings, judges may have made a decision about a defendants' guilt or innocence, which they then convey, through verbal or non-verbal means to the jurors. Research concerned with the relationship between judges belief in a defendant's guilt and the jurors decisions of guilt or innocence (Hart, 1995) found a causal relationship between these two factors. After hearing audio-taped testimony participants were then shown a video-tape of a judge giving the jury instructions, the judge's expectation of the defendant's

guilt was manipulated. The participants who viewed the judges who had guilty expectations were more likely to return a guilty verdict.

Halverson et al., (1997) propose that jurors may be particularly susceptible to the non-verbal leakage of the judge's beliefs about the defendant's guilt for two reasons. Firstly, being part of a jury is a novel experience for most people. In times of uncertainty as how to act people will look to the authority figure in a situation, in this case a judge, and follow their lead as how to behave. Secondly, if a juror is confused by the instructions given by the judge, they will attend more to the judge's non-verbal behaviour and take their cues from that. Halverson et al., (1997) therefore devised an experiment which reduced the complexity of the judge's instructions, and made them more understandable. They hypothesised that with clearer instructions jurors would attend less to the non-verbal behaviour of the judge, and therefore be less influenced by it. The experiment was conducted with two populations, students and adults. With the student population no discernible biasing effect of the judge's expectations was found, either with the clearer instructions or the standard instructions. However, with the adult population decisions of guilt were strongly linked to judge's expectations of guilt with the standard instructions but this relationship was weaker with the revised instructions.

Following the Pygmalion experiments and those that succeeded it Rosenthal and colleagues suggested a theory of four factors to explain teacher expectancy effects (Rosenthal, 1994). Firstly Rosenthal argues that teachers create a warmer socio-emotional *climate* for those students who they have high expectations for. Secondly, he argues that teachers will give more *input* and teach more material to those high expectation students. The third factor, *output*, suggests that teachers will ask the high expectation students more often for answers to questions in the classroom. Fourthly, Rosenthal suggests that teachers will give more differentiated and personal *feedback* to high expectation students. The validity of the four-factor theory proposed by Rosenthal was vindicated by a meta-analysis (Harris & Rosenthal, 1985 and 1986) which found that the four factors which mediate the expectancy effects are significantly related to the teacher's expectations and the student's performance. Although the meta-analysis concentrates on the expectancy effects through the teacher-pupil relationship, because the majority of the studies were conducted with this interaction, some of the mediating behaviours may also be pertinent to other types of

interaction. Of the twelve behaviours found to associate strongly with teacher expectations, creating a less negative climate, maintaining closer physical distances, creating a warmer socio-emotional climate, having longer interactions, engaging in more eye contact, and smiling more, may also relate to other expectancy interactions.

Indeed the warmer socio-emotional climate may explain why rats that were expected to perform better in a maze or Skinner box did indeed perform better. When asked to rate their attitudes and behaviour towards their rats, those experimenters who thought they had 'bright' rats rated them as more pleasant and likeable than those with 'dull' rats. The 'bright' rat experimenters also handled their rats more, and more gently, and were more pleasant and friendly towards their rats than the 'dull' rat experimenters. Furthermore, the 'bright' rat experimenters watched their rats more closely than 'dull' rat experimenters. The researchers suggest that the reinforcement provided by the extra handling and closer observation may explain the superior performance of the rats in the maze and the Skinner box (Rosenthal & Fode, 1963a; Rosenthal & Lawson, 1964).

1.5. Moderating experimenter expectancy effects

Despite the prevalence of experimenter expectancy effects, researchers have also identified possible aspects which moderate the expectancy effect. Firstly it has been suggested that initial impressions of the character of a person, that they are lazy for example, cannot then be overcome by a subsequent expectancy (Jones, 1991). Indeed, it has been argued that where there is greater contact between the teacher and the student, where the teacher may have already formed an impression of that student, the Pygmalion effect is reduced (Raudenbush, 1984). Secondly, it is suggested that experimenter expectancy effects may be moderated because some individuals may not have the ability to understand or interpret the cues of the experimenter, therefore rendering the expectancies obsolete (Rosnow, Skleder, Jaeger, & Rind, 1994). It is furthermore imperative to note here that the vast majority of people are unaware of how susceptible they are to the subtle cues advertently or inadvertently communicated to them by other people (Wright, Carlucci, Evans, & Compo, 2010). Indeed as Wright et al., (2010) point out, people are more likely to believe that a horse can do mental arithmetic, rather than believe that it was due to cues from the audience or experimenter. Even in situations where an experimenter is actively trying to

influence the response of the subject, 70% of the subjects, when asked if they thought the experimenter was trying to influence them, gave the lowest possible rating of influence (Clark, Marshall, & Rosenthal, 2009).

1.6. Chapter summary

There has been a gradual assimilation of the knowledge that the experimenter can inadvertently affect the results of their research. Research in this area did not begin in earnest until Rosenthal and colleagues began studying experimenter expectancy effects in the 1950's and 1960's. Since then there has been a proliferation of research in the physical, biological and behavioural sciences as researchers have taken note of the effect they can have on their research. A vast quantity of research has been conducted by Rosenthal and colleagues finding support for experimenter expectancy effects with human and animal subjects. A finding which was further validated by the results of a meta-analysis of 345 studies of experimenter expectancy effects. The effects of experimenter's expectancies have also moved beyond the laboratory and now permeate educational, judicial, healthcare, and management arenas. Theories of how expectancies are communicated have also been highlighted, and research has attempted to determine whether some people are more susceptible to influence or whether some people are more able to influence others.

Chapter 2: The Interpersonal Nature of the Social Psychology Experiment – The Role of the Subject

The previous chapter dealt with the effect that an experimenter's expectancy can have on the subject; in this chapter we turn our attention to how the subject can in turn provide influence through their expectations. This chapter therefore aims to answer a question proposed by Rosenthal (1976), if an experimenter communicates their expectations to the subject "why do subjects act so as to confirm these expectations?" (p. 180).

As Orne (1962) highlights, the experimental procedure described in the previous section can reduce the participant to a passive responder who is to be exposed to stimuli, and observed for differences in their reactions or behaviour. Orne (1962) contends that human subjects cannot be manipulated in the same way as stimuli in the physical sciences. Instead humans are conscious and thinking individuals, with motivations, perceptions, and expectations of their own to be accounted for. As Carlsmith, Ellsworth, and Aronson contend, man as the experimenter is "imperfectly standardised...he is imperfectly calibrated" (1976, p. 280). The same must also be said of man as the experimental subject.

2.1. The unique nature of the experimental situation

The demand characteristics of the experimental procedure were identified by Orne (1962) at approximately the same time as Rosenthal and colleagues were investigating experimenter expectancy effects. In essence Orne suggested that participants modified their behaviours in-line with the cues from the experimenter as to the 'true' meaning of the experiment, and what the experimenter really wanted to find from the participant. In medical research it has long been established that patients report an alleviation of symptoms or an aggravation of side effects when they believe they are taking a particular medication, even if they are in the placebo condition. Subjects, whether in medical or psychological research, therefore respond to the experimental procedure and behave as they believe they should behave.

Orne (1962) also draws attention to the acquiescence of free will of subjects in experimental situations, and the control that is given to experimenters. This submission to the experimental procedure is demonstrated in an example by Orne (1962), who details an attempt to devise a task so boring and meaningless that subjects would give up the task

after a short while. However, he reports that it could not be done; it was not possible to devise a task so boring or meaningless that a subject would refuse to do it, or would be discontinued quickly. In a monotonous addition task, where subjects were required to rip up each page of work after they had completed it, subjects would still continue for several hours without any hostility towards the experimenter. Furthermore, subjects have been shown to place themselves in physical danger (picking up a poisonous snake or putting their hands in nitric acid) in order to fulfil the wishes of an experimenter (Orne & Evans, 1965).

Orne (1962) argues that there is something fundamentally unique about the experimental situation and "taking part in an experiment" (p. 777) which means people acquiesce to the will of the experimenter without question. Orne (1962) suggests that the experimenter-subject relationship is unequal in power, and likens it to a parent-child or doctor-patient relationship. That fundamental unique attribute is a shared belief by the experimenter and the subject that the experiment is important, that it will benefit science and humanity, and that the experimental procedure is justified. Orne (1962) demonstrated the acquiescence to the experimental method with a very simple procedure. When acquaintances were asked to do push-ups as a 'favour' they generally questioned the validity of the request. However, when another group were asked to do push-ups as part of an 'experiment' they did not question the request and were more than happy to participate. It is suggested that this is because of the unequal nature of the relationship between the experimenter and the subject. The experimenter holds a position of power as they hold the knowledge of what constitutes a correct response, whilst the participant can only guess as to how they are supposed to act (Rosenberg, 1965).

2.2. The demand characteristics of the experimental procedure and why they are attended to

Participants it is argued also want to play the role of the "good subject" (Orne, 1962: p. 778), with participants anxious to perform well and in accordance with the aims of the experiment. The role of a good subject will mean different things to different subjects, for some it may involve appearing intelligent and giving the 'correct' answer, to others it may mean appearing normal or healthy (Orne, 1969). The 'good subject' effect was investigated initially by Orne (1959) with experiments using hypnosis. He contended that a person's

behaviour whilst hypnotised may be determined by how they believe a hypnotised person should act. A group of introductory Psychology students were informed that catalepsy of the dominant hand is a normal reaction during hypnosis. The other group of students were not informed about catalepsy. Students from the two groups were then asked to volunteer to be hypnotised by experimenters who were blind to the groupings of the students. Of the nine students in the first group to be hypnotised, five demonstrated catalepsy of the dominant hand. Out of the second group of students however, none demonstrated the reaction. This according to Orne suggests that in an attempt to be 'good subjects' the students adapted their trance behaviour to their knowledge of what is expected during the hypnotic state.

Therefore, in order to be good subjects, participants will attempt to uncover the true rationale for the experiment and condition their responses or behaviour to fit the experimenter's hypothesis. Orne (1962) suggested that there are a number of cues which he labelled 'demand characteristics' that the subject will attend to which will give them insight into the purpose of the experiment. They include rumours about the research the subject may have heard, the information given to the subject when recruiting them, the setting of the laboratory, and the communication during the experiment whether implicit or explicit. The experimental procedure itself can provide cues to the subject. A repeated measures design, where a subject is tested twice on the same measure, can lead a subject to expect some change. However, demand characteristics may not be perceived in the same way by all subjects. The perception will depend upon the experience and expectations of the subject. It will also depend on the amount of time and effort invested by the subjects in the experimental procedure, the greater the investment in the procedure the greater the investment in the outcome of the experiment.

Orne (1962) suggests that the inherent nature of demand characteristics in human subjects means that it is not possible to devise an experiment where these characteristics are not a factor. As Orne highlighted in the monotonous addition task, subjects will always ascribe some meaning to the task they have been set. In the addition task for example, participants surmised they were part of an endurance task. Orne contends that where participants ascribe meaning they will also ascribe a hypothesis for the experiment; therefore instead of trying to remove demand characteristics, he instead suggests acknowledging their presence

and trying to determine the effect of the characteristics. However, determining the demand characteristics that the subject was aware of may not be a simple task to accomplish. Interviewing the subjects after the experiment, about what they perceived to be the aim of the experiment is one possible way advocated by Orne (1962) to achieve this. However, he also states that this action in itself could have demand characteristics, and a correlation between the participant's behaviour and their perception of the hypotheses may not explain the causes of the behaviour.

In order to counter these problems Orne (1962) suggested a pre-experimental inquiry, whereby participants from the same population as the experimental participants are shown the experimental materials, and the procedure is explained to them, but they do not take part in the experiment. They are then asked what they think the aims of the research and the hypotheses are. The demand characteristics of the experimental condition can then be estimated. Orne (1962) also advocates the use of simulating subjects, where a group of participants do not take part in the experimental condition, but are asked to act as if they have. Using his work on hypnosis to illustrate this point, Orne (1959) contends that a blind administrator, unaware that some participants are hypnotised whilst some are only pretending, treats the two groups the same. Moreover, experimenters have been found to be unable to distinguish the hypnotised participants from the simulators, even when they are highly trained hypnotists. Therefore, it can be argued that the simulators are using the cues from the experimental situation and the experimenter's behaviour to act how they think a hypnotised person should act. Orne (1962) then maintains that if the simulating group are using these cues, then the same cues may also be responsible for the behaviour in the experimental group, not the hypnotised state.

Following Orne's (1962) theory of an altruistic 'good subject', other researchers have attempted to explain the power of the experiment and the experimenter. In particular, Rosenberg (1969) argued that participants when faced with the experimental procedure are apprehensive about how they will be perceived and evaluated, and therefore developed the term 'evaluation apprehension'. Especially in social psychology experiments, where participants are aware that their inner psyche (intelligence, mental health, and competence) is being evaluated, they are anxious to appear 'normal' to the experimenter. The subjects therefore attempt to discern the rationale of the experiment, and then behave in a manner

which will gain them a positive evaluation from the experimenter (Rosenberg, 1980). It is also suggested that feedback to the participants as to their performance in the experiment, results in participants further changing their behaviour in order to secure more complimentary feedback.

Utilising Rosenthal's person perception task, Rosenberg (1969) empirically measured the impact of evaluation apprehension on expectancy effects by manipulating the level of evaluation apprehension. Firstly, Rosenberg had participants read a communication before participating in the person perception task. In the high evaluation apprehension condition participants read a communication which indicated a link between poor performance on the perception task and psychological maladjustment, thereby increasing the apprehension of the participant that if they fail the perception task they will appear maladjusted. In the low apprehension condition the communication indicated that the task was to collect base-line data, participants were actively discouraged from linking their performance to their psychological attributes. The results of this experiment found that a strong expectancy effect was found in the high apprehension condition, but no such effect was found in the low apprehension condition.

Whilst this research seemed to suggest that the expectancies of the experimenter are more appreciable to participants experiencing higher levels of evaluation apprehension, Rosenberg was concerned that the experiment did not include a control condition where the apprehension of the participants was not manipulated. Duncan, Rosenberg, and Finkelstein (1969) therefore conducted a replication with high and low evaluation apprehension conditions, and a control condition with no apprehension manipulation which Rosenberg likened to the level of evaluation apprehension in Rosenthal's experiments. The results of this replication found a strong linear relationship between apprehension and expectancy. In the low apprehension condition as before, no expectancy effect was identified. The expectancy effect was identified in the control condition, therefore supporting Rosenthal's research, and also to a greater extent in the high apprehension condition. Thereby providing evidence of a link between evaluation apprehension, and the experimenter expectancy effect.

In a series of experiments which have become infamous in the chronicles of psychology, Milgram (1974) proposed the 'obedient participant'. Orne (1962) had already identified the unequal balance of power between the experimenter and the subject, however, Milgram disagreed with Orne's argument that participants attempt to work out the rationale of the experiment and act accordingly. Instead he suggested that the participant acts in accordance with the demands of the experiment because they are obedient to the authority of the experimenter. Indeed even Milgram was surprised by the level of obedience that the participants would demonstrate on the encouragement of an experimenter (Van Avermaet, 2001).

In the initial experiment subjects were offered a financial reward to participate in a study concerned with memory and learning. Participants were led to believe that they had been randomly selected to be the 'teacher' and the other participant (actually a confederate of the experimenter), would be the 'learner'. The teachers were instructed to administer electric shocks, of increasing intensity, to the learner if they gave an incorrect answer, or failed to provide an answer to a paired association task. The electric shocks ranged from 15 volts to 450 volts. If the teacher hesitated or refused to administer an electric shock they were prompted by the experimenter. Milgram (1974) reported that almost two thirds (62.5%) of the participants administered the highest level of electric shock to the learner (450 volts), and on average the maximum shock was 368 volts. Participants were therefore willing to administer painful punishments to the learner on the encouragement of the experimenter, even though they could hear the learner scream in pain.

Milgram (1974) conducted a number of variations of the original experiment, all of which cemented the power of the authority of the experimenter. When the experimenter was not in the same room as the participant and gave instructions to administer the electric shocks over the phone, only 21% of participants gave the learner the maximum voltage shock. Furthermore, when the experimenter left the room and the authority for the experiment was passed from the experimenter to a confederate participant, some participants physically prevented the confederate participant from administering shocks, and only 20% of the participants administered the maximum electric shock. It would therefore appear that as Orne contended, the scientific experimenter occupies a uniquely privileged position which invokes unparalleled levels of obedience.

As with many aspects of human behaviour, it is difficult to find a definitive and absolute cause of such behaviour. As with many of the explanations of behaviour it is often true that it may be a combination of theories which provide the best explanations. It may be so with the cause of experimenter expectancy effects. It may be that the theory of altruism purported by Orne, or evaluation apprehension argued by Rosenberg, or obedience to authority suggested by Milgram all have a place in explaining the behaviour of participants during the experimental procedure (Rosnow & Rosenthal, 1997). Indeed, the above proposition was validated in a study conducted by Aiken and Rosnow (as cited in Rosenthal & Rosnow, 1975). College students were asked to rate how similar differently motivated behaviours are to being a subject in a psychology experiment. There were two altruistically motivated behaviours 'give anonymously to charity' and 'work free as a lab assistant.' Two evaluation apprehension motivated behaviours 'taking a final exam' and 'being interviewed for a job.' And two obedience to authority motivated behaviours 'obeying a no smoking sign' and 'not arguing with the professor.'

Using Multidimensional scaling the researchers created a visual map of the psychological differences between the motivated behaviours. The map revealed that the college students rated the altruistically motivated behaviours as closest to being a subject in a psychology experiment, thereby providing evidence for Orne's 'good subject' effect. However, this does not mean that the other two theories should be discounted. Indeed, the evaluation apprehension and obedience to authority motivated behaviours were also associated with being a subject in an experiment. It would therefore appear then that subjects attend to the demand characteristics of an experiment, because they want to be a good subject, they are anxious at how they will be evaluated, and are obedient to the authority of the experimenter.

Regardless of the reasons for attending to demand characteristics, Rosenthal and Rosnow (1991) suggest that there are three mediatory steps which determine whether subjects comply with them. The first step they propose is receptivity, whether the subject is receptive to the expectations of the experimenter. They suggest that there are two possibilities; subjects are either receptive or non-receptive. Non-receptive subjects at this stage demonstrate non-compliant behaviour to the demand characteristics, they did not understand the demand characteristics therefore they could not comply with them.

Receptive subjects then move to step two, motivation. It is argued that subjects are either acquiescent to the demand characteristics, non-acquiescent to them, or counter-acquiescent to them. At this stage non-acquiescent subjects are non-compliant to the demand characteristics, they are not motivated by the demand characteristics therefore they could not comply with them. Acquiescent and counter-acquiescent subjects then move to the final step, capability. Subjects are either capable of acting on their motivation or they are not. Both acquiescent and counter-acquiescent subjects who are incapable of acting on their motivation are non-compliant to the demand characteristics. They could not act acquiescent or counter-acquiescent to the demand characteristics therefore they could not comply with them. Counter-acquiescent subjects who are capable of acting on their motivation demonstrate counter-complaint behaviour; that is they behave in an opposite manner to the demand characteristics. Acquiescent subjects who are also capable of acting on their motivation demonstrate compliant behaviour; they cooperate with the demand characteristics. It is therefore suggested that demand characteristics can only affect the subject's behaviour when the subjects are receptive to them, are motivated towards or against them, and are capable of acting on their motivation.

2.3. Chapter summary

Research has attempted to explain why subjects in experiments behave so as to validate the experimenters' hypothesis. It is suggested that subjects attune to the demand characteristics of the experiment to determine the aims and hypotheses of the experiment. A number of reasons to explain this have been postulated. Orne (1962) suggests that subjects attend to the demand characteristics because they wish to be a good subject, Rosenberg (1969) suggests that it is because they are apprehensive of being negatively evaluated, and Milgram (1974) suggests that it is because subjects are obedient to the authority of the experimenter. Regardless of why subjects attend to the demand characteristics of the experiment it is argued that they must be receptive to them, motivated by them, and capable of acting on them, for demand characteristics to be a factor in an experiment.

From the review of the literature in the first two chapters it is clear that the social psychology experiment is indeed a social interaction between the experimenter and the

subject. As such, the interpersonal behaviour, personality, beliefs, knowledge and a seemingly endless list of factors of both the experimenter and the subject all become variables in the experiment. The social psychology experiment is therefore particularly susceptible to bias from both the experimenter and the subject (Carlsmith et al., 1976).

Chapter 3: Experimenter Expectancy Effects in Eyewitness Identification Procedures

In Chapter 3 we move from examining the role of the experimenter and the subject in the social psychology experiment, to considering a specific area of research; the effect of experimenter expectancy and demand characteristics in eyewitness identification procedures.

3.1. The empirical evidence of bias in eyewitness identification procedures

Before the police line-up became a common feature of the criminal justice system, eyewitness identification consisted of show-ups and courtroom identifications. The aim of instituting the line-up procedure was to protect innocent suspects; in order to do this the construction of the line-up must be fair and unbiased towards the suspect. For the line-up to be fair the original rules for conducting a line-up, derived from case-law, stated that the suspect should not stand out to the witness. Therefore a number of fillers who physically resemble the suspect and each other were required. The specific number of fillers needed or the specificities of the resemblance were not mandated, and different law enforcement agencies developed their own standards (Malpass, Tredoux, & McQuiston-Surrett, 2006).

In Chapter 1 the possibility of role bias by members of the criminal justice system was highlighted. It is clear that police officers may also be susceptible to the same bias. Police officers are the front-line of the adversarial criminal justice system and therefore identify their role as working for the prosecution. This bias may result in police officers conducting biased line-ups in order to obtain an identification (Roberts, 2009). The presence of biased eyewitness identification procedures has been accepted by members of the criminal justice system for many years, as the quote from Supreme Court Justice William Brennan attests:

> "The vagaries of eyewitness identification are well-known... A major factor contributing to the high incidence of miscarriage of justice from mistaken identification has been the degree of suggestion inherent in the manner in which the prosecution presents the suspect to witnesses for pre-trial identification... Suggestion can be created intentionally or unintentionally in many subtle ways."(*United States v. Wade*, 1967).

Whilst eyewitness identification has always been a popular topic for experimentation by Psychologists, the introduction of DNA technology to the criminal justice arena, highlighting

the inadequacies of such identification evidence, through the number of false identifications, has resulted in a resurgence of interest (Wells, Small, Penrod, Malpass, Fulero, & Brimacombe, 1998). In particular there has been a resurgence in research focussing on the accuracy and reliability of eyewitness identifications. Utilising the research conducted by Rosenthal and colleagues, studies have attempted to uncover the effect that an administrator can have on an identification procedure.

In order to achieve this Wells and Luus (1990) proposed that an analogy be drawn between a social psychology experiment and a line-up procedure. They argue that the important aspects of an experiment have comparable elements in a line-up. For example the experimenter is the police officer conducting the line-up, the subjects are the eyewitnesses. The hypothesis of the experiment is that the suspect in the line-up is the perpetrator, and that the eyewitness will identify the suspect. Therefore, as with any experiment the types of bias discussed above, confirmation, response, and expectancy, can infiltrate the line-up procedure. From the stand-point of a line-up as an experiment, Wells and Luus (1990) suggest that the best practice for conducting line-ups can be empirically tested. Indeed, there are a number of examples of studies which seek to improve the line-up procedure by utilising the experimental framework.

Firstly, Wells and Luus (1990) state that mock witness line-ups can be used to test whether the construction of the line-up is biased towards the person the police believe is the suspect. In such situations the witness did not view the crime (hence the term mock witness) but has instead read a description of the suspect (Doob & Kirshenbaum, 1973). It is argued that if the mock witness identifies the suspect above chance level then the line-up construction is biased towards the suspect, or the experimenter's hypothesis of who the suspect is has become apparent to the mock witness. The second test of witness susceptibility to inference from the line-up administrator is called the blank line-up control (Wells, 1984). In this type of line-up an actual eyewitness is shown a suspect absent line-up, if the eyewitness picks a person from the line-up then they can be excluded from the line-up containing the actual suspect, as they have been shown to be either susceptible to response bias, or have a poor memory of the event.

The decision process of the witness has been extensively studied, and Wells (1984) has argued that the witness undertakes one of two decision processes, an absolute judgement or a relative judgement. In an absolute judgement the witness compares each member of the line-up with their memory for the suspect. If no member of the line-up matches the memory of the suspect no identification is made. In a relative judgement the witness picks out the person that best fits their memory of the suspect relative to the other members of the line-up. The evidence for the relative judgement model is most apparent in target-absent line-up procedures. As the witness is comparing each line-up member to each other rather than their memory of the suspect, they will identify the person who looks most like the culprit even if the culprit is not in the line-up (Wells, 1984). It is argued that a number of factors can affect the fairness of a line-up procedure by shifting the witnesses' willingness to make a relative judgement rather than an absolute judgement.

Using the line-up as experiment analogy and the research on experimenter expectancy effects and demand characteristics reviewed in Chapters 1 and 2, it is clear that a line-up procedure can be susceptible to both of these procedures. Firstly the line-up administrator, fulfilling the role of the experimenter, knows the location of the suspect in the line-up and may be motivated to obtain an identification if there is a belief among the law enforcement officials that they have found the culprit. Therefore, the administrator may unconsciously influence the witness into identifying the suspect. Secondly, the three explanations for demand characteristics may also apply to line-up procedures. Eyewitnesses may be susceptible to the good subject effect by identifying someone from a line-up because they want to help the police and catch the 'bad guy'. Furthermore, a witness may be apprehensive that the line-up is an evaluation of their memory, and therefore want to prove that their memory of the event is good by identifying someone from the line-up. Finally, as line-up procedures are normally conducted by a law enforcement professional it is clear that witnesses may be susceptible to the cues exhibited by the line-up administrator because they are in a position of authority (Phillips, McAuliff, Kovera, & Cutler, 1999). The line-up procedure when viewed as an interaction between the administrator and the witness is therefore open to bias which can affect the reliability of an identification. Below the sources of bias in a line-up procedure are discussed.

3.1.1. Instruction bias

The line-up procedure has been the subject of numerous experiments, where aspects of the procedure have been varied. For example, in one early study the type of instruction given to the witness and the presence or absence of the suspect in the line-up were manipulated (Malpass & Devine, 1981). In this experiment undergraduate students viewed a staged act of vandalism and were then asked to take part in a line-up to identify the culprit. The line-up procedure was varied; some of the witnesses were given biased instructions; they were led to believe the culprit was in the line-up, and they did not have the option not to make an identification. The other witnesses were given unbiased instructions; they were told the culprit may be in the line-up, but they were also given the option of not identifying anyone from the line-up. Furthermore, in half of the line-ups the suspect was present; in the other half the suspect was absent. The results of this study indicate that the instructions given to witnesses are extremely important. Those witnesses given biased instructions were more likely to choose a suspect from the line-up, even when it was a suspect absent line-up. Therefore, the biased witnesses were more likely to make false identifications.

The study conducted by Malpass and Devine (1981) has been criticised for lacking ecological validity. Cutler and Penrod (1995) suggest that the biased instructions do not represent the instructions that would be given by police officers in the real world. Therefore, Cutler, Penrod, and Martens (1987a) devised an experiment using more subtle suggestions. Student eyewitnesses again viewed a staged crime, this time a videotape of a liquor store robbery, and again half of the eyewitnesses received bias instructions before viewing the line-up. However, in this experiment the biased instructions did not indicate that the suspect was in the line-up, they were told to pick the person they believed was the suspect. The other half of the witnesses were given unbiased instructions, and therefore given the option to not make a choice from the line-up. Even with more subtle biasing instructions this experiment found comparable results to Malpass and Devine (1981), those eyewitnesses who received biasing instructions were more likely to make a false identification from a target absent line-up.

This research and more in the same vein (Cutler, Penrod, & Martens, 1987b; Cutler, Penrod, O'Rourke, & Martens, 1986; O'Rourke, Penrod, Cutler, & Stuve, 1989), has highlighted the

effect of biased instructions. However, other researchers have argued that laboratory studies of instruction bias cannot be generalised to real world police line-ups as the laboratory studies cannot adequately imitate the seriousness and stress for the witness viewing an actual crime. Indeed, the ecological validity of laboratory studies in this area has been questioned, and has led some researchers to move away from the mock crime scenario. One such study (Kohnken & Maass, 1988), whilst still using a staged crime, only informed half of the eyewitnesses that the event had been staged, the other half believed the event to be a real crime. This study again varied the bias of the instructions and the presence and absence of the suspect. Kohnken and Maass (1988) reported that those eyewitnesses who believed they were in a real situation were more likely to not identify anyone from the line-up. The reality of the situation therefore appeared to render the eyewitness more cautious about making an identification.

A meta-analysis of biased instructions from eighteen studies was conducted by Steblay (1997). The results of this analysis indicate that the performance of an eyewitness is significantly affected by biased line-up instructions. In particular, instructions that imply the suspect is in the line-up and not providing an option of not choosing from the line-up are most influential. Moreover, the meta-analysis addresses the issue of the validity of laboratory studies. Steblay (1997) stresses that eleven of the studies included in the analysis involved “high-realism” (p. 295) events, and whilst the effect sizes are lower in these events than lower realism events, the effect of biased versus unbiased instructions is still significant.

The majority of studies in this area have one feature in common; they employ a comparison between target present line-ups and target absent line-ups, simulating police investigations conducted in the real world. Target present line-ups represent line-ups where the police suspect is in the line-up, and target absent line-ups represent line-ups where the police suspect is innocent (Clark, 2005). The meta-analysis conducted by Steblay (1997), found that biased instructions had differing effects according to the presence or the absence of the suspect in the line-up. Notably, in target absent line-ups there is a consistent increase in false identifications, whereas in target present line-ups there is an inconsistent effect on correct identification rates.

Clark (2005) argues that whilst the increase in false identifications through biased instructions in target absent line-ups is understandable, as the biased instructions lower the witnesses' decision criteria, there should also be a concomitant increase in the number of correct identifications in target present line-ups. In a review of Steblay's (1997) meta-analysis, Clark (2005) suggests that there are more correct identifications in target present line-ups with biased instructions, and the biased witnesses are accurate above chance. Therefore, whilst Steblay's (1997) analysis indicated that biased instructions were inherently bad (producing more false identifications, and no increase in correct identifications), Clark (2005) has found that biased instructions can increase correct identifications. He does not suggest that biased instructions should be allowed though, as they can lead to the false identification of an innocent suspect.

3.1.2. Suggestibility

Research has shown that suggestive remarks made by line-up administrators, can have a substantial effect on how a witness perceives what they saw, and can induce an eyewitness to make an identification (Wells & Quinlivan, 2009). In an experiment conducted by Wells and Bradfield (1998), witnesses were shown a simulated crime, and then were asked to identify the culprit from a suspect absent line-up. Some of the witnesses were then given a suggestive remark confirming that they had identified the culprit, whilst the other witnesses were given no such remark. In the group given the suggestive remark, 27% reported that their view of the culprit was good or excellent, and 20% reported easily making out the details of the culprits face. This is compared to the group who were not given a suggestive remark, where none of the witnesses reported a good or excellent view, and none could easily make out the details of the culprits face.

Research in this area has devised the term the 'post identification feedback effect', and this effect is not confined to direct feedback about whether they had identified the correct person, but also concerns feedback as to whether they had chosen the same person as other witnesses. Luus and Wells (1994) found that witnesses informed that they had made the same choice as other witnesses were more confident in their identification than those who had received no such information. The post identification feedback effect has been found to be particularly robust and reliable through meta-analytic study (Douglass &

Steblay, 2006). This meta-analysis of 14 studies which included 20 experimental tests found that witnesses reported significantly better witnessing conditions; and stronger and sharper memories if they received post-identification feedback confirming the accuracy of their identification. This finding, the authors contend, has important implications for real world situations. As post-identification feedback increases confidence, it also distorts the original memory and compromises the reliability of the memory. Therefore witnesses will appear more confident and compelling at trial, but this will be based on the post-identification feedback and not the accuracy of their memory, which may be quite poor. Research has also suggested that the post-identification feedback effect can be moderated, whereby confidence ratings are not inflated if the witness is asked to consider their confidence before feedback is given (Wells & Bradfield, 1999).

Research conducted by Clark et al., (2009) investigated how different statements during the line-up procedure can influence eyewitness identification decisions. They employed three statement conditions, in the control condition the line-up administrator did not speak to the witness during the procedure. In the subtle-influence condition the line-up administrator did not speak for the first 12 seconds of the procedure, after 12 seconds they made statements including “take your time” “look at each photograph carefully” and “there’s no rush” (p. 66). In the third and final condition, the similarity-influence condition, again the administrator did not speak for 12 seconds, then they asked the witness “Is there anyone in the line-up who looks more similar to the person you saw than anyone else in the line-up” (p. 66). In the two influence conditions, if the witness had failed to make an identification after the statements above, the line-up administrators prodded the witnesses with one more statement “If you’re unable to make an identification, that’s ok, just let me know” (p. 66).

The results of this study found that despite administrators in all conditions giving unbiased instructions; that the suspect may or may not be in the line-up, witnesses’ identification decisions were influenced by the administrator in the two influence conditions. Furthermore, the influence differed between the subtle-influence condition and the similarity influence condition. In the similarity-influence condition there was an increase in overall identification rates, correct identifications, and false identifications, in both target-present and target-absent line-ups. These results are in-line with the researcher’s

hypothesis, and they argue, can be explained by the criterion shift and relative judgment models of eyewitness decision making.

The subtle-influence condition however, did not conform to the hypothesis that subtle influences would change the decision process, resulting in a greater willingness to make an identification, and therefore cannot be explained by the criterion shift and relative judgement models. In this condition there was only an increase in identifications in the target-absent line-ups, false identifications increased, but positive identifications did not increase. In order to explain this contradiction, the authors propose a model similar to the Feature Contrast Model (Tversky, 1977), which explains how the way memory is matched to a line-up member, can be altered. It is suggested that the comments made in the subtle-influence condition allow the subject to reassess and reconsider the importance and weighting of the matching and mismatching features of the suspect. Clark et al., (2009) argue that the weight-shifting model can explain not only why false identifications but not positive identifications increased in the subtle-influence condition, but can also explain the same results found in other research (Clark, 2005; Steblay, 1997).

3.1.3. Foil bias

The aim of a fair and unbiased line-up is to protect against mistaken identification, whereby the foils match the description of the suspect, and the suspect does not stand out as different from the other members of the line-up. Anecdotal examples of biased line-up construction include a line-up constituted of one black man and five white men, where the witness had described the suspect as a black man (Ellison & Buckhout, 1981). In another case the witness described three Indian suspects and then took part in a line-up containing three Indian men and three white men (Pelwani v. S. 1963, cited in Malpass et al., 2006).

Although the above real-world examples may seem extreme and overtly biased, studies with more subtle differences in the line-up construction, line-ups with very similar looking foils compared to very dissimilar looking foils, have shown notable results. Research manipulating the quality of the foils used in a line-up has shown that in target present line-ups there is not a considerable difference in identification accuracy between high similarity line-ups and low similarity line-ups. However, in target absent line-ups, 70% of eyewitnesses made false identifications from a low similarity line-up, compared to 31% in the high

similarity line-up (Lindsay & Wells, 1980). However, it is also argued that it is important that the fillers in a line-up are not too similar to each other, as the witness will not be able to differentiate between a line-up of clones (Malpass et al., 2006).

The homogeneity of the appearance of the line-up extends to the clothing worn by the members of the line-up. A study conducted by Lindsay, Wallbridge and Drennan (1987), attempted to examine whether the type of clothing worn can affect the number of false identifications. In a series of experiments Lindsay et al., (1987) conducted line-ups with biased clothing. In condition 1 (biased clothing) the suspect wore the same clothing worn during the crime and the foils wore different clothing. In condition 2 (usual clothing) all line-up members wore different clothing to that worn during the crime. In condition 3 (dressed alike) all line-up members wore similar clothing. Whilst for correct identifications the clothing condition was not a significant factor, the number of false identifications were significantly different according to the clothing condition. For the dressed alike condition they constituted 10% of the identifications, for the usual clothing condition they constituted 21% and for the biased clothing condition they constituted 38% of the identifications.

Clark and Godfrey (2009) reviewed the literature on foil selection, and compared seven experiments which manipulated foil similarity over both target present and target absent line-ups. They reported that in six of the seven experiments both correct and false identification rates increased when the similarity of the foils was low. The issue of foil selection has long been a contentious one in the literature, with arguments raging about the best practice for selecting foils. Two competing procedures for foil selection have been advocated. In a suspect-matched selection, foils are selected to coincide with the appearance of the suspect. This is by far the most popular method for constructing a line-up used by law enforcement agencies (Wogalter, Malpass, & McQuiston, 2004). However, this type of construction has been criticised for both resulting in line-ups where the members are too similar (Luus & Wells, 1991), and in line-ups where the suspect stands out (Wogalter, Marwitz, & Leonard, 1992). This approach has also been criticised as it relies on the suspect being the actual culprit and not an innocent suspect, this can result in increased mistaken identifications (Clark & Tunnicliff, 2001).

The second procedure is description-matched selection, whereby foils are selected to coincide with the verbal description of the culprit by the eyewitness. It is argued that this procedure produces a less biased line-up, where the suspect does not stand out, but there is still some diversity within the line-up (Wells, Seelau, Rydell, & Luus, 1994). Research has suggested that more correct identifications are obtained with this type of line-up (Wells, Rydell, & Seelau, 1993). However, unless the eyewitness can provide a very detailed description of the suspect's face it is difficult to construct a line-up that adequately matches the suspect (Malpass et al., 2006). The contention between suspect-matched and description-matched foil selection shows no sign of abating. Indeed, a meta-analysis conducted by Clark, Howell, and Davey (2008) reviewed studies which compared the two types of line-up construction, and found mixed results. The description-matched method produced more correct identifications, but also more false identifications. Nevertheless, researchers have called for the adoption of description-matched procedures for foil selection (Luus & Wells, 1991; Clark & Tunnicliff, 2001), although, Clark and Godfrey (2009) stress the need for further research.

3.1.4. Presentation bias

Line-up procedures have historically followed a simultaneous format, where the witness views the suspect and the foils at the same time. In the first study of its type, Lindsay and Wells (1985) varied the line-up procedure shown to undergraduate eyewitnesses. Half of the eyewitnesses were given the standard simultaneous line-up containing six photos, whilst the other half were shown the same six photos sequentially (although witnesses in this condition were unaware of how many photos they would view). Both line-ups were also varied according to the presence of the suspect. Whilst the type of line-up did not affect the number of correct identifications in the target present line-up, the results in the target absent condition were remarkable. In the simultaneous presentation 43% of eyewitnesses made a false identification, whereas in the sequential presentation false identifications dropped to 17%.

In a replication of this study, Cutler and Penrod, (1988) informed witnesses in the sequential condition how many photos they would view, however the results did not change, the authors therefore concluded that the sequential line-up still produced less false

identifications even with knowledge of the number of photos to be viewed. This result was also reported by Lindsay, Lea, and Fulford (1991), who concede that whilst the effect of the sequential line-up was not as pronounced in their study, it was still apparent when the size of the sequential line-up was known to the witness.

The full force of this effect was highlighted in a series of experiments conducted by Lindsay, Lea, Nosworthy, Fulford, Hector, Le Van, and Seabrook (1991). In these experiments line-ups were biased according to clothing, foil selection and instruction given. The results suggested that the sequential line-up reduced the effect of the bias on false identifications. Furthermore, when the biases were combined, Lindsay et al., (1991) found that the sequential line-up still had a remarkable effect on reducing false identifications (84% in the simultaneous condition compared to 25% in the sequential condition). A meta-analysis of studies comparing simultaneous and sequential line-ups was conducted by Steblay, Dysart, Fulero, & Lindsay (2001). Twenty three papers (providing 30 tests of the hypothesis) were included in the analysis. The results of the analysis found for the complete data-set only a small effect of the superiority of the sequential line-up. When target-present line-ups alone were compared, simultaneous line-ups provided more correct identifications. However, in target-absent line-ups sequential line-ups provided more correct rejections of the line-up. The results also indicated that witnesses are more likely to make a choice from a simultaneous line-up.

The results of the meta-analysis also provided some insight into factors which reduce the superiority of the simultaneous line-up in target-present line-ups. Firstly, it is argued that there is a smaller difference in correct identifications between simultaneous and sequential line-ups when the stimulus used in line-up experiments is more realistic. Secondly, simultaneous line-ups display even higher rates of correct identification when biased instructions are employed. However, as unbiased instructions are widely employed it would appear that the different rates for correct identifications between the two line-ups, is not so great in real-world situations. Moreover, a verbal description of the suspect by the witness before the line-up has been reported to have the strongest effect on simultaneous line-ups, as it completely removes its superiority for correct identifications over the sequential line-up. As it is not practical for the police not to obtain a verbal description from a witness again, it appears that the difference between the two line-ups will not be as great in real-

world situations. Finally, the sequential line-up proved superior in crime situations as compared to non-crime situations. In sum Steblay et al., (2001) report that the correct rejection rate for sequential line-ups increases whilst the correct identification rate for simultaneous line-ups decreases as real-world conditions are maximised.

It is argued that the success of the sequential line-up stems from the fact that it forces witnesses to make an absolute judgement; they must compare each photo individually with their memory of the suspect and make a decision before they see the next photo. Whereas in the simultaneous line-up witnesses can make a relative judgement, the person most similar to the suspect, relative to the other members of the line-up, is chosen. Ebbesen and Flowe (2002) however, have argued against the absolute relative judgement process, and instead propose the criterion shift model. In this model it is argued that the sequential line-up produces a higher criterion for making a positive identification, but does not increase the accuracy of the memory. Whilst the studies presented above seem to imply the superiority of the sequential line-up, some researchers have issued words of caution (Phillips, et al., 1999), arguing that sequential line-ups are particularly susceptible to investigator bias, which will be reviewed in due course.

Acceptance of the superiority of the sequential line-up has not been universal. In a divergence from traditional laboratory studies the Report to the Illinois Legislature: Illinois Pilot Program on Sequential Double-Blind Lineup Procedures (Illinois Pilot Report) conducted a year-long field study to examine the sequential line-up. The results obtained were contradictory to those expected; sequential double-blind line-ups actually resulted in more known false identifications than the simultaneous single-blind line-ups. The authors of the Illinois Pilot Report stress the importance of the field study methodology in this area for providing “real world factors” (Mecklenburg, Bailey, & Larson, 2008, p. 22) that laboratory studies are lacking. In particular the field study incorporates the psychological experience of viewing a real crime, authentic viewing conditions, the impact of the line-up procedure on the witness, and an understanding of the consequences by the witness of identifying or not identifying someone from the line-up.

Although the design of the Illinois Pilot Report has received criticism (Schacter, Dawes, Jacoby, Kahneman, Lempert, Roediger, & Rosenthal, 2008), namely for the confound of

comparing blind sequential line-ups with non-blind simultaneous line-ups, Mecklenburg et al. (2008) maintain that it still provides a "significant and valuable contribution to the study of eyewitness identification" (p. 23), for three reasons. Firstly, it is argued that the studies which demonstrate the "sequential superiority effect" themselves contain confounds, for example, different line-up presentational formats (Cutler & Penrod, 1988), line-up instructions, foil similarity (Lindsay, et al. 1991), randomization of photographs (Sporer, 1993), and size of photos (Morgan, Hazlett, Doran, Garrett, Hoyt, Thomas, Baranoski, & Southwick, 2004), between simultaneous and sequential line-ups.

Secondly, Mecklenburg et al. (2008) argue that the data from the Illinois Pilot Report can still be evaluated regardless of the confounding factor. If the comparison between the double-blind sequential line-up and the single-blind simultaneous line-up is removed, the known error rate of 9.2% for the double-blind sequential line-up can be compared to field data that has controlled for investigator bias. An example of this, as noted by Mecklenburg et al. (2008), is the data collected by the Office of the District Attorney in Queens, New York, which found a known error rate of filler identifications between .58% and 5.62%. The third reason for the efficacy of the Illinois data is that it does not necessarily follow that the low filler identifications in the non-blind simultaneous line-ups are due solely to the non-blind procedure. Again Mecklenburg et al. (2008) point to field data, this time from the New York City Police Department, which compared blind simultaneous line-ups with non-blind simultaneous line-ups, and found no difference in identification rates.

Recently, borrowing from the medical literature, Mickes, Flowe, and Wixted (2012) have used Receiver Operating Characteristic (ROC) analysis to compare simultaneous and sequential line-up practices. As the sequential line-up has been found to decrease the rate of false identifications, whilst at the same time decreasing the rate of correct identifications, Mickes et al., (2012) have argued that the traditional method of comparing diagnosticity ratios (proportion of guilty suspects correctly identified / proportion of innocent suspects incorrectly identified) does not always indicate the superior procedure. Instead they suggest that ROC analysis can determine which type of line-up is diagnostically more accurate. After conducting three experiments comparing simultaneous and sequential line-ups, ROC analysis found no sequential superiority effect, even when biased foils were used. The authors concede that a decrease in the number of false identifications (as shown with the

sequential line-up) may be worth the corresponding decrease in the number of correct identifications. However, they argue that using a simultaneous procedure with a more conservative decision rule (for example only counting identifications made with over 70% confidence in that identification) can result in lower false identification rates, but higher correct identification rates.

3.1.5. Investigator bias

The line-up as experiment analogy has also been used to support the 'investigator bias' hypothesis, which posits that through verbal and non-verbal means the line-up administrator, with their knowledge of who the suspect is, can intentionally or unintentionally influence who the witness identifies from the line-up. This is a relatively new application of the work conducted by Rosenthal that has only recently gained attention in the eyewitness psychological literature. Indeed by the early 1990's researchers (Wells, 1993; Wells & Luus, 1990) were beginning to hypothesise about the possibility of investigator bias, although research had yet to substantiate the hypothesis (Cutler & Penrod, 1995).

A real-life example of investigator bias is provided by Loftus and Ketcham (1991) and details the case of Howard Haupt, who was wrongly identified by a witness from a photo-spread for the kidnap and murder of a young boy. In this instance the line-up procedure was tape-recorded therefore the researchers were able to analyse the transcripts. This analysis revealed that whilst the witness had reservations about being able to identify the suspect; the officer conducting the line-up pressed the witness and repeatedly referred to Haupt's photo, only ending the line-up when the witness had made some concession that Haupt's photo was closest to the suspect.

Whilst the above case provides one anecdotal example of investigator bias, researchers have attempted to empirically test the phenomenon. In one such experiment, Phillips et al., (1999), manipulated the administrator's knowledge of the identity of the suspect in the line-up and the presentation of the line-up. After viewing a live staged crime involving two perpetrators, witnesses were asked to identify the suspects from target absent photo-spread line-ups. There were two conditions of administrator knowledge in this experiment; half of the administrators did not know the identity of the suspect (double-blind condition). And half of them were informed by the experimenter who the suspect was (single-blind

condition). Each administrator conducted simultaneous and sequential line-ups. The results of this study confirmed the researchers' hypothesis; sequential line-ups in the single-blind condition were more likely to result in false identifications.

The Phillips et al., (1999) study included one further variable, the presence of an observer during the line-up procedure. The researchers hypothesised that the presence of a silent observer during a line-up procedure would negate the effect of investigator bias. However, the reverse was found, more false identifications were recorded when the observer was present. The researchers explained this turn-around of expectations by suggesting that the presence of the observer (who was also the experimenter) provided extra incentive for the administrator to perform well and produce the expected result.

Haw and Fisher (2004) argue that whilst double-blind procedures may eliminate the cues from the administrator as to the identity of the suspect, it does not eliminate the perception of the witness to changes in the administrators' demeanour, posture, or expression. They argue that there will always be "biasing factors" (p. 1107) when there is an interaction between the administrator and the witness. They therefore conducted an experiment which limited the amount of contact between the administrator and the witness. In the high contact condition the administrator sat close to or across from the witness whilst administering the line-up procedure. In the low contact condition the administrator sat behind the witness out of their eyesight whilst the witness administered the line-up procedure themselves. The administrator stayed in the room to ensure the correct procedure was followed. Both target-absent and target-present line-ups were conducted in both a sequential and simultaneous format. A number of interesting results emerged from this study, firstly witnesses were more likely to make a choice from the line-up in the high contact condition, and secondly in the target-absent condition witnesses were more likely to select the target-substitute, thereby, making a false identification.

Haw and Fisher (2004) therefore conclude that reducing the contact between the administrator and the witness is particularly useful as false identifications were reduced, without adversely affecting the number of correct identifications. However, they also highlight an apparent contradiction in their research, and one that has been found in many of the studies reviewed above also. Although double-blind procedures and limiting contact

reduced false identifications there was no comparable increase in correct identifications when the target was present in the high contact and single-blind conditions. The authors cautiously suggest that the presence of the perpetrator serves in "outshining" (p. 1110) the effect of the administrator.

The administrator effect has been shown not only to influence the witness's identification, but also their confidence in their identification. In an experiment which manipulated administrator knowledge of the location of the suspect in a target absent line-up, Garrioch and Brimacombe (2001), found that witnesses were more confident in their identifications if the administrator of their line-up knew the location of the suspect. Even though the administrators were overtly informed that they should not give any feedback to the witnesses, the video-taped interactions between the administrators and witnesses revealed verbal behaviours (the questioning tone when repeating the witnesses choice) and non-verbal behaviours (maintaining eye contact with the witness and smiling), which the authors claim explains the increased confidence ratings.

Although the above literature provides some compelling evidence for investigator bias, there are also a number of studies which did not find this effect. As Greathouse and Kovera (2009) point out, the effect of investigator bias is not "robust" (p. 72), with inconsistent results emerging from the data. In an attempt to provide some clarity in this area Greathouse and Kovera (2009) conducted a large study which manipulated target presence (target present vs. target absent), administrator knowledge (single-blind vs. double-blind), line-up presentation (simultaneous vs. sequential), and line-up instruction (biased vs. unbiased). They found that witnesses were more susceptible to the effects of administrator knowledge when other biasing factors (simultaneous line-ups and biased instructions) were present. In these conditions witnesses were more likely to make an identification in the single-blind administration regardless of the absence or presence of the culprit.

Greathouse and Kovera (2009) also examined the interaction between the line-up administrator and the witness by video-recording the procedure. They observed that the behaviour of the administrator changed according to their knowledge of the identity of the suspect. The blind observers found that single-blind administrators displayed certain behaviours at greater rates than double-blind administrators. In particular, they were more

likely to tell the witness to examine the line-up carefully, or to look again at the line-up, and in some cases they also told the witness they knew the location of the suspect. The observers also noted non-verbal behaviours which were more prevalent with single-blind administrators, for example they were more likely to remove the picture of the suspect more slowly if the witness failed to identify it. Furthermore, as the line-up administrator is in control of the procedure, they can control the duration of the line-up. They can end the procedure quickly if the witness picks the suspect, or they can prolong the procedure if a filler is chosen, giving the witness chance to change their mind (Wells & Quinlivan, 2009).

Indeed a number of other suggestive behaviours by line-up administrators have been identified. For example, verbal utterances during the line-up, or non-verbal behaviour, such as nodding, head shaking, smiling, and leaning forward (Buckhout, 1974), have been argued to influence the witness as to whom they should choose. Although verbal behaviours are more discernible in the experimental procedure, Garrioch and Brimacombe (2001) stress that non-verbal behaviours can be just as effective as verbal behaviours in their influence. Indeed the research conducted by Rosenthal (1976), reviewed in Chapter 1, has found that experimenters expectancies can be communicated through both kinesic and paralinguistic channels. They are most effective though when they are communicated through both of these channels.

From their research Greathouse and Kovera (2009) tentatively claim evidence for demand characteristics and hypothesis leaking in line-ups, and stress the importance of double-blind line-up procedures. However, they also concede the need for continued investigation in this area. In particular they stress the need for a concerted effort to investigate the effects of administrator knowledge with different moderating variables such as the presentation method of the line-up. Furthermore, they call for an exploration of how the strength of the witness's memory trace interacts with administrator effects.

3.2. Line-up good practice and the need for double blind procedures

As we have seen, research has highlighted the potential for bias in police line-up procedures. In response to this, researchers have compiled a number of recommendations of good practice when conducting line-ups (Wells, 1988). Indeed one of the most vociferous proponents in this field Gary Wells was responsible for drafting good practice

recommendations for the American Psychology/Law Society (Wells et al., 1998). Wells and colleagues devised four rules for reducing the risk of false identifications from a line-up.

1. *The person who conducts the line-up should not be aware of which member of the line-up is the suspect.*

This rule addresses the issue of investigator bias. As we have seen in the section above, whether intentional or not, the line-up administrator can influence who the witness selects from the line-up. Wells et al., (1998) therefore advocate the use of double-blind line-up procedures, where the person conducting the line-up is not involved with the case and does not know who the suspect is.

2*. Eyewitnesses should be explicitly told that the person in question might not be in the line-up, and they should not feel they have to make a selection from the line-up. Furthermore, eyewitnesses should be told that the person administering the line-up does not know which person the suspect is.*

This rule addresses the issue of instruction bias and follows from the research cited above that witnesses are less likely to make a false identification if they have been warned that the suspect may not be in the line-up.

3*. Based on the eyewitness's description of the suspect the suspect should not stand out in the line-up as different to the foils.*

This rule addresses foil bias and states in order for the line-up to be fair and unbiased all members of that line-up should match the description given by the witness. As Wells et al., (1994) point out, this does not mean that all foils should look exactly the same as the suspect, rather they should match the description of the suspect, and there should be some variation in the line-up, but not enough for the suspect to stand out.

4. *A statement of the confidence of the eyewitness that they have picked the right person should be taken at the time of the line-up and before any feedback on their selection.*

This rule addresses the issue of suggestibility and the over-inflation of confidence that a witness may have in their identification if they receive confirming feedback. Whether the feedback specifically tells them that they have selected the person the police suspect is

responsible (Wells & Bradfield, 1998), or they are told they have picked the same person as another witness (Luus & Wells, 1994). Over-inflation of confidence is a serious issue in eyewitness testimony, as eyewitnesses are often asked how confident they are in their identification in court. If they have received confirming feedback their confidence will be inflated even if their identification is false.

Wells et al., (1998) stress that whilst the four recommendations have minimal financial costs for the police; the benefits of following these procedures far outweigh these costs. They do concede that Rule 1, that a person blind to the location of the suspect should conduct the line-up, involves the most effort on the part of the law enforcement official to implement. However, the authors maintain that the implementation of Rule 1 is possible as there is no extensive training needed to conduct line-ups, therefore, any person in the department whether a police officer or a civilian could conduct the line-up. The authors furthermore, do not regard the above recommendations as exhaustive. For example they also recommend the use of sequential line-ups over simultaneous line-ups (although only in conjunction with double-blind procedures) and the video-recording of line-up procedures.

The police however continue to resist these recommendations particularly the double-blind line-up procedure. In particular police officers point to the financial costs and organisational effort of ensuring there is an administrator who is not aware of the identity of the suspect available. They also argue that it is beneficial to have a trained police officer who is familiar with the details of the investigation to oversee the line-up in order to "recognise relevant information during the identification procedure" (Mecklenburg, 2006: p. 57). Officers from the Illinois Pilot Program report that they received complaints from witnesses and victims because of the delay whilst waiting for a blind administrator. Moreover, they stress that the double-blind procedure is the antithesis of the ethos of collaboration and information sharing, and can have a detrimental effect on the rapport established between an investigating officer and a vulnerable witness.

Despite the objection to some of the research recommendations, legislation has begun to implement a number of the suggested procedures. In England and Wales eyewitness identification procedures are governed by the Police and Criminal Evidence Act (PACE) 1984. Code D of the Act, which was revised in 2011, addresses some of the issues of bias reviewed

above (Home Office, 2010). The code states that a line-up must contain at least eight foils for one suspect, furthermore, the age, general appearance, and position in life of the foils must resemble the suspect. The suspect must not stand out from the foils; any unusual physical features should be where possible concealed or replicated on the foils. Witnesses should be informed that the culprit may not be present in the line-up, and they should indicate if they cannot make a positive identification. The code also prescribes the use of video identification procedures wherever possible, unless a live line-up is more appropriate. However, the sequential format of the video identification in the Code differs to the procedure advocated by Wells et al., (1998) as the Code allows the witness to view the sequential line-up more than once, and the witness is not required to make a decision about each person before they see the next one (Roberts, 2009). Whilst the Code states that the person conducting the line-up should not be involved in the case, it does not stipulate that the administrator should be blind to the location of the suspect. A number of jurisdictions in the United States have recently begun to implement double-blind line-ups, although the vast majority still do not (Greathouse & Kovera, 2009).

3.3. Chapter summary

In this chapter we have seen through the line-up as an experiment analogy and the vast amount of research that this methodology has spurned, that a line-up administrator can bias a line-up procedure in a number of ways. They can provide bias through their instructions to the witness, through their post-identification feedback to the witness, how the line-up is presented, and the selection of line-up fillers. Moreover, from the research conducted by Rosenthal and colleagues it is clear that the line-up procedure is also susceptible to experimenter expectancy effects. The knowledge of the location of the suspect in the line-up and the expectation that the eyewitness will identify the suspect can bias the witness's response. The mediation of experimenter's expectancy is usually unconscious and is a very subtle process that involves minute kinesic and paralinguistic communications. Therefore, research is still attempting to uncover the channels that communicate experimenter expectancies.

The criminal justice system has taken note of the psychological research conducted. Legislation has implemented some of the recommendations that have been advanced.

However, many law enforcement agencies still resist the blind administrator recommendation, which is perhaps the most important. This may be because it is the recommendation that lay people have the most trouble understanding. The processes involved are so minute and people generally feel that they can control their speech and behaviours; therefore the criminal justice system does not give adequate weight to the dangers of administrator bias. Mecklenberg (2006) reported that the general feeling amongst the police officers in the Illinois Pilot Program was to dismiss the effect of their knowledge of the location of the suspect; “a witness who can identify the offender can do so under either procedure” (p. vi).

Chapter 4: Human Memory and Eyewitness Identification

In Chapter 3 the various ways that an experimenter can bias an identification procedure were reviewed. In this chapter we examine how the bias exhibited by experimenters can be exacerbated by the witness's memory of the event and the identification procedure. Memory, and the ability to identify a person from that memory is dependent on a large number of factors, some of which will now be discussed.

4.1. The memory process and its fallibilities

It is now generally understood that there are three stages in the memory process, perception, storage, and retrieval. Research has found that errors can occur at each of these three stages (Ainsworth, 2000). The first of these stages, perception involves the acquisition and encoding of the information. A number of errors can occur at this stage, firstly perception is subjective to the individual, and secondly it is selective as not all information can be encoded. Finally, perception is effected by an individual's prejudices and stereotypes (Duncan, 1976). In a criminal situation specifically, the high level of arousal, caused by witnessing a traumatic event such as a violent crime, can disrupt the encoding of the memory (Deffenbacher, 1983).

The length of the retention interval, the time between encoding the information and retrieving it, can affect the accuracy of the memory retrieved. Indeed over 100 years ago Ebbinghaus (1885, cited in Wells & Quinlivan, 2009) was reporting on the 'forgetting curve' which states that forgetting occurs most quickly just after the event. Research has determined that the accuracy of a memory declines over time (Deffenbacher, Carr, & Leu, 1981), and therefore culprit identification rates also decline as the length of the retention interval increases (Shapiro, & Penrod, 1986). This decline has been shown to occur quickly, one or two days after the event (Tollestrup, Turtle, & Yuille, 1994), and has been reported as substantial after seven days (Behrman, & Davey, 2001). As Wells and Quinlivan (2009) point out, the reliability of the 'forgetting curve' has been substantiated; however, the time frame of forgetting can vary according to the type of information to be remembered. They suggest that a name can be forgotten after a few minutes whereas a telephone number can be remembered for years. Remembering, they argue, is facilitated by rehearsal of the memory.

Whilst memories can fade over time, they can also be significantly transformed by new information which is received after perception of the original memory. In an oft cited study, Loftus and Palmer (1974) found that recall can be affected by verbal information given to the eyewitnesses. After viewing a film of a car accident, participants were asked to judge the speed of the cars, however, the questions to the participants varied. Some participants were asked how fast the cars were going when they *hit* each other, other questions used *smashed, collided, contacted,* or *bumped* to describe the accident. The witnesses estimates of the speed of the car were faster when the word 'smashed' was used compared to when the word 'contacted' was used. Indeed, Loftus has consistently found that witness's memories are particularly malleable when presented with post-event information, and most notably, witnesses are unaware that their memory has changed even if they are deliberately misled (Weingardt, Toland, & Loftus, 1994). It is unknown whether the original memory is changed by the new information, or a new memory is created which then consumes the original (Brigham, Wasserman, & Meissner, 1999). However, studies in this area indicate that once the new information has been presented to the witness, the original memory is lost and cannot be retrieved (Fruzzetti, Toland, Teller, & Loftus, 1992). Furthermore, research suggests that witnesses are more susceptible to the damaging effects of post-event information when there is a longer time interval between the event and the introduction of the information (Loftus, Miller, & Burns, 1978).

In a similar vein to memory transformation by new information, numerous studies have also found that memory is susceptible to the 'misinformation effect' where misleading information is introduced to the witness. In one such study (Stark, Okado, & Loftus, 2010) participants were shown photographs depicting the theft of a wallet, where the thief placed the wallet in his jacket pocket. Participants later listened to narratives of the photographs which included misleading information, including that the thief had put the wallet in a pocket of his trousers. When asked about the photographs a number of the participants reported seeing the thief put the wallet in his trouser pocket from the photographs.

Research has attempted to understand why some people are susceptible to the misinformation effect. Loftus (2005) has suggested that misinformation can take hold in a person's memory if the person is forgetful or has a poor memory. When considering personality and cognition, a number of factors have been found to correlate with the ability

to resist or not resist misinformation. A study conducted by Zhu et al. (2010a; 2010b) found that intelligence, perception, memory, and face judgement all correlate negatively with false memories following exposure to misinformation. When personality characteristics were considered, negative correlations were also found between false memory and depression, fear of negative evaluation, novelty seeking, and negative coping. Positive correlations though were found for persistence, self-directedness, and active coping strategies. Subjects were found to be particularly susceptible to the misinformation effect if they had low cognitive abilities combined with certain personality characteristics, including high self-directedness, high reward dependence, high cooperativeness, low harm avoidance, or low fear of negative evaluation.

The final stage of the memory process is retrieval. At this stage eyewitnesses can be particularly affected by the stress of the situation, whether attending an identification parade or giving evidence in court, which can affect the recall of the memory. However, the method of retrieving the information is also important for its accuracy. Indeed, it has been suggested that if eyewitnesses are asked to give a "full and complete" description of the suspect they are more likely to inaccurately fill-in gaps in their memory in order to give the full and complete description (Meissner, cited in Brigham et al., 1999).

Research has also suggested that eyewitnesses are susceptible to the unconscious transference phenomenon whilst retrieving memories. Extensively studied by Loftus and colleagues, the term unconscious transference relates to the bystander effect, where a witness mistakenly identifies a person familiar to them as the suspect. In real-life example a railway ticket agent, who was the victim of an armed robbery, identified his assailant from a line-up. The man he identified though had a cast-iron alibi, and could not have committed the robbery. The witness stated that the man looked familiar to him, and it transpired that the man had been a customer of the witnesses on a number of occasions. The witness had transferred the familiarity of the man's face from being a customer to the robber. Early empirical research of unconscious transference found that 60% of subjects mistakenly identified an innocent bystander (Loftus, 1976).

4.2. Eyewitness identification

The importance of eyewitness identifications to the criminal justice system cannot be overstated. The police and judicial system rely heavily on eyewitnesses to accurately and objectively recall what they witnessed, and then confidently identify the culprit, regardless of the amount of time that has passed. In court, jurors find eyewitness testimony to be persuasive and of high evidential value (Loftus, 1996). However, it has been argued that the expectations placed on the human memory by the criminal justice system are improbable (Ainsworth, 1998). This view would appear to be substantiated by the number of wrongful convictions due to eyewitness testimony that have come to light since the implementation of DNA testing.

In recent years cases of wrongful convictions have received widespread attention and publicity. Indeed there are a number of high profile cases of wrongful conviction due to faulty eyewitness testimony, including the case of Ronald Cotton who was convicted and spent over ten years in prison for rape, on the basis of a single identification from a photographic line-up. In this case the victim was convinced that she had correctly identified her attacker. When asked if she recognised the man who was the actual perpetrator based on DNA evidence, she stated she had never seen him before (Wells & Quinlivan, 2009). Research conducted by Wells et al., (1998), has found that 90% of 40 trials of wrongful convictions involved eyewitness testimony, and some of these innocent people were identified by more than one witness. In America to date the Innocence Project has secured the exoneration of 303 individuals through DNA testing, who had been wrongly convicted. It is stated that for 75% of those exonerated, faulty eyewitness testimony was the cause of the wrongful conviction (The Innocence Project, 2013). However, 303 exonerations must be seen as the tip of the iceberg of wrongful convictions as it is estimated that only 5-10% of criminal cases include DNA evidence (The Innocence Project, 2013). Nevertheless, eyewitness testimony has long been, and continues to be deemed as very strong evidence of a persons' guilt. Indeed, uncorroborated evidence from one eyewitness can still be sufficient to secure a conviction.

4.2.1. The fallibility of eyewitness identification

The fallibility of eyewitness identification has been recognised for a number of years, indeed in 1925 the Home Office provided guidelines for the collection of identification evidence. In response to a number of high profile cases of miscarriages of justice the Devlin Committee was convened in the 1970's to consider the role of mistaken identifications in these miscarriages. The Committee found that in approximately 260 cases the defendant was convicted when the only evidence against them was eyewitness testimony. Furthermore, in half of these cases the eyewitness evidence consisted of just one eyewitness identification. The recommendations of the Devlin report (1976) led to a judgement in *R. v. Turnball.* Where identification is disputed the judge must caution the jury that a confident witness is not necessarily accurate, and when considering the reliability of the identification the circumstances of that identification must be considered. However, it was not until 1984 that the U.K. government produced legislation in this area in the form of the Police and Criminal Evidence Act (PACE) (Valentine & Heaton, 1999) which provides a code of practice for eyewitness identification procedures (Home Office, 2010).

In a meta-analysis of eyewitness accuracy in natural settings, for example, a convenience store, Cutler and Penrod (1995) found that correct identifications ranged from 34% - 48%, whilst false identifications ranged from 34% - 38%, indicating that eyewitnesses are as likely to be correct as incorrect. In a review of real-life line-ups which included over 1500 witnesses, Wright and McDaid (1996) found that almost 20% of the witnesses picked a foil from the line-up. However, the unreliability of the human memory does not deter the criminal justice system from assuming that memories are accurately perceived, then stored in the brain waiting to be recalled, much like a video-camera (Ainsworth, 2000). Indeed, research suggests that jurors are poor assessors of eyewitness credibility, as they can be unaware that eyewitnesses can be inaccurate. In fact they have been found to overestimate the abilities of witnesses to both retain memories and produce accurate identifications (Cutler & Penrod, 1995). Furthermore, research has shown that people are unaware of the bias inherent in many of the identification procedures conducted by the police. In particular people do not consider a non-blind line-up procedure to be less reliable or more biased than a double-blind procedure (Wright et al., 2010).

It is precisely the over-reliance on eyewitness identifications by the criminal justice system that has led to the large body of psychological research that we have today. Researchers have attempted, under experimental conditions, to determine why eyewitnesses make mistakes in their identifications, whether accurate identifications can be discriminated from inaccurate identifications, and what measures can be introduced to improve the reliability of eyewitness identifications. The vast majority of this research can be classified as laboratory studies where volunteers (usually undergraduate students) view a staged event and then are asked to identify the suspect from a line-up under varying conditions. A number of variables have been studied and these variables can be characterised as belonging to one of two groups. Firstly there are estimator variables, which are those variables which are not under the control of the criminal justice system, including, lighting conditions, credibility of the witness, confidence of the witness, and the presence of a weapon. The second group of variables are system variables, which are (or could be) under the control of the criminal justice system, and include how a line-up is conducted, how questions are asked of the witness, and the level of support given to the witness (Wells, 1978). Many of the system variables were discussed in detail in Chapter 3.

In an area as large as eyewitness identifications it is impossible to adequately review the entirety of the research conducted. However, an attempt is made here to review some of the more pertinent studies, in order to highlight some of the factors that can affect the reliability of eyewitness identifications. Starting with estimator variables, it is necessary to divide this group into two sub-groups; witness characteristics and event characteristics.

4.3. Estimator variables – Witness characteristics

The research comparing male and female witnesses is contradictory and complex; the overall impression of the literature is that males and females cannot be distinguished by their accuracy. A meta-analysis conducted by Shapiro and Penrod (1986) suggested that as females were more likely than males to make an identification, they were slightly more likely to make both correct and false identifications. A study by Shaw and Skolnick (1994) found evidence for an own-gender bias, female witnesses were more successful at identifying female targets, whilst male witnesses were more successful at identifying male targets. The same own-gender bias was found in a replication by the authors (Shaw &

Skolnick, 1999), however, this effect was found to be dependent on the object the culprit was seen carrying at the time of the crime. When the culprit was viewed not carrying an object, or when the object was an expected item (a book) the own-gender bias was found. However, when the culprit carried a weapon or an unusual or salient item (a child's toy or a stethoscope) the eyewitnesses were more accurate in their identifications of culprits of the opposite sex. To explain this contradiction Shaw and Skolnick (1999) suggest that when carrying an interesting object a witness attends to the culprit more closely when they are of the opposite gender because they are more interesting and attractive. However, they suggest that the interesting object only serves to distract the witness from attending to the culprit when the culprit is of the same gender.

Perhaps one of the most widely known concepts in eyewitness research is the 'own race bias' which states that witnesses are better at recognising members of their own race than members of other races (Anthony, Cooper, & Mullen, 1992). This finding has been validated as robust in a meta-analysis conducted by Meissner and Brigham (2001), which considered data from 39 research articles and almost 5,000 participants over a period of 25 years. The meta-analysis found what the authors have termed a "mirror effect" (p. 15), whereby witnesses were more likely to correctly identify, and less likely to mistakenly identify members of their own race. Whilst the finding of an own race bias is consistent, the reasons for such findings have been less consistent. Studies have suggested that there may be a number of reasons for the own race bias, including racial attitudes and prejudice, physiognomic homogeneity of faces, low interracial contact, perceptual learning, and cognitive processes (see Meissner & Brigham, 2001 for a review).

A large amount of research has been conducted looking specifically at the age of the witness and how this affects the reliability of their memory. Overall the vast majority of this research would seem to suggest that witnesses at either end of the age spectrum are the least reliable (Pozzulo & Lindsay, 1998; Wells & Olson, 2003). In particular, post-event information and suggestive questioning are most damaging to children and the elderly (Yarmey, 1996). Young adults have been found to be significantly superior to older adults when recalling perpetrator characteristics, environmental details, and details of actions and events, for both free recall and cued recall (Yarmey, Jones, & Rashid, 1984). Furthermore, older adults are less descriptive about the perpetrator (Brimacombe, Quinton, Nance, &

Garrioch, 1997), they are more likely to falsely recognise a face (Searcy, Bartlett, & Memon, 1997), they are more open to source monitoring deficits, and are more confident than younger adults (Mitchell, Johnson, & Mather, 2003). In a series of studies conducted by Memon, Gabbert, and Hope (2004) they found that older adults were more susceptible to the 'mug-shot exposure effect', they were significantly more likely to make false identifications from both the mug-shot album and the photo line-up. Furthermore, when they split the older adults into two groups, young-old (under 68) and old-old (over 69) they found a significant difference in line-up performance between the two groups, with the old-old group making more false choices from the target-absent line-up.

When considering child witnesses, a review of studies has suggested that as children get older the number of correct identifications increases. Children aged three to five years only make correct identifications at slightly above chance level. By the age of eight children are correct between 50% and 58% of the time. By the age of eleven this has risen to between 60% and 70%, and by age fourteen up to 80% of witnesses are accurate (Chance & Goldstein, 1984). In an ecologically valid experiment, groups of children of differing ages were unwitting eyewitnesses to a staged theft. When asked to identify the suspect from a six-person photo-array, older students were significantly more accurate than the younger children (Brigham, Van Verst, & Bothwell, 1986).

It is generally assumed that people with learning disabilities are unreliable witnesses (Green, 2001), due to the belief that a learning disability results in impairment at each stage of the memory process (Kebbell & Wagstaff, 1999). However, research conducted in this area suggests that the criminal justice system exacerbates many of the memory problems experienced by witnesses with learning disabilities (Kebbell & Hatton, 1999). It is argued that if the criminal justice system were to increase the number of open-ended questions, facilitating free recall, then witnesses with learning disabilities would be more reliable and credible. As with witnesses with learning disabilities witnesses with mental illness are not taken seriously, their credibility and reliability are questioned (Karras, McCarron, Gray, & Ardasinski, 2006). Research has suggested that there is a link between mental illness, most notably depression, and memory impairment. In a meta-analysis Burt, Zember and Niederehe (1995) found a relationship between depression, schizophrenia and memory impairment, but they found no such relationship in those subjects with substance abuse

issues or anxiety disorders. An early study by Howells (1938) seemed to indicate a relationship between low levels of intelligence and poor face recognition accuracy. However, research conducted since has found a paucity of evidence for this relationship (Brown, Deffenbacher, & Sturgill, 1977).

Early laboratory research suggested that there was no difference in the accuracy of victims and witnesses (Hosch & Cooper, 1982), however, a subsequent study by Tollestrup et al. (1994), using actual criminal cases, found that victims of robbery were more likely to identify the suspect than witnesses of robbery. This effect though seems dependent on the type of crime, as they also found that victims of fraud had lower levels of correct identifications than witnesses of robbery. Researchers explain this in terms of the emotional arousal experienced (Yuille, 1993). This arousal will be greater in those who directly experience an emotional event, compared to those who only witness it. However, Behrman and Davey (2001) argue that witnesses can still experience high emotions, and correctly identify at the same rate as victims, even though they were not directly involved in the crime.

4.4. Estimator variables – Event characteristics

The amount of time the witness views the suspect and the distance from which the witness views the suspect both have an impact on the quality of the memory that the witness will have. Research has suggested that up to a distance of 25 feet, face perception and face identification is accurate. This steadily decreases until 150 feet, where face perception and identification is zero (Loftus & Harley, 2005). If people are aware of the identity of a person, they state that they can clearly identify those people even from several hundred feet away. In what has been termed a 'visual hindsight illusion' people fill in the details that they think they can see. However, when they do not know the identity of the person, and therefore cannot fill in the details, they cannot identify that person past 150 feet (Harley, Carlsen, & Loftus, 2004). A meta-analysis found that there is a relatively weak relationship between the duration of the exposure to the face and the accuracy of the identification (Shapiro & Penrod, 1986). Furthermore, witnesses reliably over-estimate the time they were exposed to the suspect's face (Shiffman & Bobko, 1975); whilst under-estimating the time that the witnesses face is occluded (Wells & Murray, 1983).

Whilst the criminal justice system considers the time spent looking at the suspects face as equalling the time spent attending to the suspects face, researchers have suggested that this is not the case. Instead they have drawn attention to the type of attention and the type of processing that the witness undertakes. The human memory is limited in its capacity; witnesses are therefore restricted to what they can attend to. Processing all of the features of a person's face (specific processing) takes time and uses mental capacity, whilst processing the face globally saves time and capacity. However, each type of processing has its own advantages; global or holistic processing of the face results in increased facial recognition (Patterson & Baddeley, 1977), whereas attending to the individual features increases the ability to reconstruct the face (Wells & Hryciw, 1984). Due to the capacity of memory, research has suggested that an eyewitness who attends to the peripheral aspects of a crime, may be able to give a detailed description of the event, but they will be less accurate at identifying the suspect from a line-up (Cutler et al., 1987a).

A meta-analysis conducted by Steblay (1992) purported to confirm the long-held belief among researchers that the presence of a weapon distracts the witness from attending to the perpetrator, resulting in a decrease in the number of correct identifications. In studies which track witnesses eye movements, researchers indeed found that a witness's visual attention is drawn to a weapon, away from other aspects, including the culprit's face (Loftus, Loftus, & Messo, 1987). However, in a review of actual criminal cases Behrman and Davey (2001) reported the opposite, suspect identifications increased when a weapon was present, although this is in contrast to another real-world study which did find evidence (albeit only slightly significant evidence) of the weapon effect (Tollestrup et al., 1994). The stress experienced by a witness during an event can also affect their perception and recall of an event. A certain amount of stress can actually improve a witness' perception and encoding of the event. However, when stress levels become too high, such as when there is violence or a weapon present, the stress can negatively affect memory.

In a review of the research literature Clark and Godfrey (2009) argue that the literature concerning factors that can affect memory accuracy, including exposure duration, stress, and retention interval, is counterintuitive and controversial. Clark and Godfrey (2009) reviewed published data, and compared longer exposures, lower stress, and shorter

retention intervals (better memory conditions), with shorter exposures, higher stress, and longer retention intervals (worse memory conditions). They found that in the worse memory conditions the correct identification rate decreased, however, there was little effect on the false identification rate, and the innocence risk increased (the probability that the suspect is innocent, given that the suspect was identified). Results that they conclude are "self-contradictory" (p. 26).

4.5. System variables – How line-ups are conducted

As noted above, system variables concern those factors affecting an identification procedure which are or can be under the control of the criminal justice system. The procedure for conducting line-ups varies by country to country and even within countries. There are a number of different identification procedures, each with their own positive and negative features. In mug-shot inspections witnesses are shown the mug-shot files which contain photos of convicted criminals. The benefits of this procedure include reducing the stress on the witness as they do not need to physically see the perpetrator. However, there are some problems with this approach, namely that only those who have been previously convicted of an offence are contained within the mug-shot files. Research has also found that the more photos a witness views the more likely they are to make a false identification (Ellis, Shepherd, Flin, & Davies, 1989).

Show-ups are the simplest form of identification procedure, and involve the police showing the witness one person who they believe is the perpetrator. Research suggests that show-ups are particularly popular with criminal justice personnel in the USA, with studies reporting a show-up use rate of between 30% and 77% (Gonzalez, Ellsworth, & Pembroke, 1993; McQuiston & Malpass, 2001, cited in Steblay, Dysart, Fulero, & Lindsay, 2003). However, this procedure has also been found to produce the most false identifications (Kassin, Ellsworth, & Smith, 1989; Yarmey, Yarmey, & Yarmey, 1996), and has been described as inherently suggestive (Lindsay & Wells, 1980; Phillips et al., 1999).

Despite the popularity of show-ups and the concerns over the reliability of this method, there is a dearth of research in this area. When attempting to conduct a meta-analysis comparing line-ups and show-ups, researchers could only find eight articles to include

(Steblay et al., 2003). Regardless of the small sample size, the results of the meta-analysis were contradictory to the accepted view that show-ups produce more false identifications. For target present procedures, line-ups and show-ups had comparable correct identification rates, whilst show-ups also had greater correct rejection rates in target present procedures. Furthermore, when foil choices were removed from the analysis of false identifications, these rates were again comparable for both line-ups and show-ups. Nevertheless, Steblay et al. (2003) argue that the show-up can still produce a "dangerous error" (p. 532) whereby an innocent suspect is identified because of their resemblance to the perpetrator. The chance of this occurring in a line-up is mitigated by the presence of foils. A fact that the authors suggest balances the performance of line-ups and show-ups.

Studies comparing line-ups and show-ups over a time-delay (Yarmey et al., 1996; Dekle, 1997) have found that line-ups and show-ups conducted immediately after the event had comparable false identification rates, however correct identification rates for show-ups were considerably higher than for line-ups. Both line-ups and show-ups showed sharp drops in correct identification rates over a 30 minute, 2 hour, and 24 hour delay, although the show-up maintained higher rates of correct identification throughout. Nevertheless, the show-ups also exhibited a large increase in false identification rates compared to line-ups over the time delay, at the 24 hour point correct and false identification rates for show-ups had converged.

In contrast to show-ups, photo-spread identifications involve showing the witness an array of photographs, containing foils and the suspect, and have been found to be preferable to both mug-shots and show-ups. However, they also have their own problems. Firstly, witnesses who saw a live person may not recognise them in a static photograph, and secondly witnesses may feel pressured to pick someone (Ainsworth, 2000). Identification parades are similar to photo-spread identifications, but witnesses see the suspect amongst a number of foils. They are also susceptible to the same problems as photo-spread identifications.

In order to reduce the costs of conducting identification parades for the criminal justice system, researchers have suggested the use of video identifications. Valentine and Heaton (1999) argue that a video line-up can provide a fairer and less biased option to the standard

photo line-up or identification parade. In particular they suggest that a large database of volunteers could be compiled which would allow fair line-ups to be constructed quickly, without having to ensure that the witness, suspect, and suitable foils are all available at the same time and place. Thereby also reducing the number of cancelled identification procedures. Furthermore, video line-ups are less intimidating for the witness, resulting in less anxiety, which it is argued improves the likelihood of the witness correctly identifying the suspect (Ainsworth & King, 1988).

In order to determine the reliability of identifications from video line-ups Valentine and Heaton (1999) compared them to traditional photo line-ups using a mock witness paradigm. They argued that in a fair line-up 11% of mock witnesses should identify the subject by chance. However, more mock witnesses identified the suspect from photo line-ups (25%) compared to video line-ups (15%). The authors conclude that video line-ups are therefore fairer. In subsequent research conducted by Valentine, Darling, and Memon (2007), the use of moving video images (members of the line-up face the camera, turn to show the right profile, then turn to show the left profile, then turn back to face the camera), was compared to static images. The authors hypothesised that the witness would be better able to identify the culprit from a moving image which showed the face from different angles. The results however did not conform to the hypothesis; the correct identification rates from culprit-present moving and static video line-ups were indistinguishable. Nevertheless, the moving video line-ups do result in a decrease in false identifications from a culprit-absent line-up.

4.6. The determinants of the accuracy of a line-up identification

A number of further factors, which cannot be described as either estimator or system variables, have been found to relate to the accuracy of an identification from a line-up. Researchers have attempted to determine whether the confidence of the witness in their identification and how quickly a witness makes the identification, can inform the accuracy of their identification. In court proceedings it is imperative that witnesses appear credible and confident in their testimony. There is a widespread belief amongst criminal justice officials that witness confidence and witness accuracy are linked (Wells, Malpass, Lindsay, Fisher, Turtle, & Fulero, 2000). Indeed the Supreme Court of the United States in 1972 stated that eyewitness accuracy could be indicated by eyewitness confidence (Howitt, 2006).

Traditionally however researchers have maintained that overall the relationship between confidence and accuracy is weak (Sporer, Penrod, Read, & Cutler, 1995). This relationship is not straight-forward though; indeed research has found that older adults are more confident in their identifications even when they are incorrect (Dodson & Krueger, 2006). However, when research has limited its analysis to the confidence of accurate identifications, and when confidence is measured immediately after the identification, a considerable correlation between confidence and accuracy has been reported (Behrman, & Davey, 2001; Brewer, & Wells, 2006; Sporer et al., 1995).

A theoretical basis for a strong correlation between confidence and accuracy has also been suggested. Sauer, Brewer, Zweck, and Weber (2010) highlight that the evidence for an accurate identification, such as optimal viewing conditions and a short retention interval, is the same evidence used for a confidence rating, therefore, there should be a correlation between the two. As the memory of the event is said to be responsible for both the accuracy and the confidence of the identification, and as the accuracy of an identification decreases with a greater retention interval, Sauer et al. (2010) have studied the effects of retention interval length and confidence. Using an immediate identification condition and a delayed identification condition (with retention intervals ranging from 20 to 50 days), they found that as expected, accuracy was greater in the immediate identification condition. There was a relationship between accuracy and confidence in both of the conditions, however, participants in the delay condition were more overconfident than participants in the immediate condition.

As has been reviewed above (Chapter 3), witness confidence is particularly malleable, especially in those witnesses who make mistaken identifications and receive confirmatory feedback (Bradfield, Wells, & Olson, 2002; Wells, & Bradfield, 1998). Witness confidence can also be affected by repeated testing of the memory. In particular, repeatedly questioning a witness about an inaccuracy in their testimony increases their confidence in said testimony (Shaw, 1996; Shaw & McClure, 1996).

The time taken for a witness to make an identification from a line-up is argued to be a better predictor of the witness's accuracy than their confidence. The finding that witnesses who make an identification in less time are more accurate has been reported in a number of

studies (Dunning & Perretta, 2002; Smith, Lindsay, & Pryke, 2000; Sporer, 1993). Indeed Smith et al. (2000) found 69% of their witnesses to be accurate within 15 seconds. Denning and Perretta (2002) found their witnesses to be almost 90% accurate when they made their identifications in less than 10-12 seconds. Witnesses who made an identification after this were only accurate 50% of the time. However, a review of the 10-12 second rule has suggested that whilst identification speed is a valid predictor of identification accuracy, it is not necessarily within the 10-15 seconds. Furthermore, combining identification speed and confidence is a better predictor of identification accuracy than considering these aspects separately (Weber, Brewer, Wells, Semmler, & Keast, 2004).

4.7. Chapter summary

In this chapter we have discussed the fallibilities of human memory and eyewitness testimony. Memory is limited in its capacity; people can only attend to a certain amount of information. The information that is attended to is then susceptible to being forgotten or transformed, and the procedure used to retrieve the memory can also have detrimental effects on the reliability of that memory. Despite the fallibility of memory the criminal justice system places great stock in the testimony of eyewitnesses.

Many of the studies described in Chapter 3 measure the bias of an identification procedure by the number of correct and incorrect identifications after a staged event. There is therefore a fundamental problem with assessing whether the line-up procedure is biased when a poor memory of the event or a characteristic of the witness may confound the influence of the administrator. In Chapter 6 the rationale for removing the memory component in the current research is discussed.

Chapter 5: The Background to the Present Research: A real-life example of a questionable identification from a line-up procedure?

5.1. The case of Al Megrahi

On the evening of the 21st of December 1988, Pan Am flight 103, a Boeing 747, exploded over the Scottish border town of Lockerbie. The 259 passengers and crew on board were killed along with 11 people from the town of Lockerbie. Almost two years after the bombing, following an extensive multi-jurisdictional investigation, Scottish and American authorities charged Abdelbaset Ali Mohmad Al Megrahi and Al Amin Khalifa Fhimah with the murder of 270 people through the bombing of flight 103. Both men, who were of Libyan descent, were believed to be members of the Libyan Intelligence Services. However, it was to be a further eight years after these indictments before the two suspects were surrendered by the Libyan government to U.N. officials to stand trial.

The trials of Al Megrahi and Khalifa Fhimah, under Scottish law, began on the 3rd of May 2000 at a former military base in the Netherlands. Subsequently Al Megrahi was convicted and sentenced to life in prison; his co-defendant however, was acquitted. Much of the evidence against Al Megrahi centred on a small shop in Malta and the eyewitness testimony of the shopkeeper Mr Anthony Gauci. Despite a number of unsuccessful appeals against the conviction and sentence, Al Megrahi remained in prison until August 2009 when he was released on compassionate grounds. Although he maintained his innocence until his death in May 2012, his conviction was not quashed and he remains to date the only person to be convicted, and to have served time in prison, for the bombing of Pan-Am flight 103.

Almost as soon as the bomb exploded innumerable conspiracy theories were advanced as to who was responsible for the attack. Against a complex backdrop of international political intrigue and conspiracy theories, many doubted Al Megrahi's and indeed the Libyans involvement in the bombing. Al Megrahi and the Libyan government maintained their innocence, and this appeared to have been vindicated in 2007, when, at the culmination of a four year investigation, the Scottish Criminal Cases Review Commission (SCCRC) stated that "a miscarriage of justice may have occurred" and granted Al Megrahi leave to appeal against his conviction for a second time. Nevertheless, before the appeal could be heard, it was abandoned in August 2009, and amid much controversy Al Megrahi was released on

compassionate grounds. Therefore, only some of the evidence contained within the appeal was heard before the appeal court. However, due to a number of miscarriage of justice campaigners, most notably the Justice for Megrahi Committee, much of the evidence that was to form the basis of the appeal has become public knowledge.

There were a number of grounds of appeal proposed by Al Megrahi's defence team. Grounds 1 and 2 were argued in court before the appeal was dropped. They maintained that the circumstantial evidence on which Al Megrahi was convicted was inherently weak and therefore no reasonable jury could have returned a guilty verdict. In particular the identification evidence of Anthony Gauci, the date of the purchase of the clothing from Gauci's shop, and the inference that the suitcase containing the bomb was ingested at Malta were challenged as unreasonable inferences. Grounds 3 of the appeal were due to be heard before the court in November 2009. These grounds maintained that the appellant had been denied a fair trial, specifically, in relation to how the identification evidence from Anthony Gauci was obtained, and a failing by the Crown to disclose information about Anthony Gauci and his identification evidence. There were other grounds of appeal that at the time the appeal was dropped had not been finalised, but dealt with concerns about forensic evidence and defective representation.

5.2. The eyewitness identification evidence against Al Megrahi

It is evident that the identification evidence supplied by Anthony Gauci formed a crucial part of the prosecution's case against Al Megrahi. Indeed, the trial court found the identification evidence to be "entirely reliable" (Trial Court Opinion [12] and [67]), and found Mr Gauci himself to be a "credible and careful witness" (Trial Court Opinion [69]). Nevertheless, the trial court also expressed views that Mr Gauci's evidence was "not unequivocal" and "not absolute" (Trial Court Opinion [88] and [89]).

The importance of Anthony Gauci's identification evidence proceeds from the fact that clothing from 'Marys House', the shop owned by Anthony Gauci in Sliema, Malta, was found in the suitcase containing the Improvised Explosive Device (IED) which destroyed the Boeing 747. Anthony Gauci was first interviewed by police in September 1989 and over the next eleven years participated in a number of interviews and identification procedures,

culminating in his courtroom identification of Al Megrahi in 2000. In addition to the identification of Al Megrahi, Anthony Gauci also provided the police with details of the items of clothing bought, the date that the items were bought, and the movements of the suspect on the day he bought the clothing. Nevertheless, as with his identification evidence, there are a number of discrepancies with the statements made by Mr Gauci.

Police officers investigating the Lockerbie bombing originally visited Mr Gauci's shop on the 1st of September 1989. On this visit they showed Mr Gauci a sample of cloth and photographs of "blast damaged" clothing. Mr Gauci told police that he remembered a customer in the winter of 1988 that stood out in his memory because he bought a random assortment of clothing, and did not seem to care about the sizes. Mr Gauci stated that the man bought three sets of pyjamas, a tweed type jacket, brown trousers, lighter trousers, a blue baby-gro with a sheep's face on the front, a red and black tartan cardigan, and a black umbrella. In a subsequent interview (13th September 1989) Mr Gauci stated "I cannot remember anything else that this man bought in my shop". On a later visit Mr Gauci was shown a blast damaged fragment of shirt with a SLALOM label, on this occasion Mr Gauci remained adamant that the man did not buy any shirts. However, in a statement a year later (10th September 1990) Mr Gauci recollected that he had actually sold the man two shirts.

In his first statement Mr Gauci gave a very clear account of the man's movements on the day he bought the clothing. He stated that the man entered the shop at 6:30pm, bought a number of items, paid 56 Maltese pounds in cash then left the store stating he had other shops to visit and would return for his purchases. Mr Gauci recollects that it was raining and the man opened his umbrella as he left the shop. Mr Gauci states that the man returned 15 minutes later, he took the parcels, saying he had a taxi waiting, left the shop, turning left and walking up the street. Mr Gauci recalls seeing a white Mercedes taxi waiting.

However, in subsequent statements Mr Gauci's recollection of events that day changed. In his first statement he recalls the man took his purchases to a waiting white Mercedes taxi. In his evidence though Mr Gauci states that he carried the purchases to the taxi for the man, (Day 31, page 4752, line 15). Furthermore, initially Mr Gauci recalls that he was working alone in the shop when the man came in. However, in the Crown precognition (18th March

1999 and 25th August 1999), he stated that his brother was present in the shop as the man returned to collect his parcels.

Furthermore, from the outset Mr Gauci was not certain of the date that the purchase took place. In his initial statement he thought that it had been a week day in late November 1988. In a later statement (19th September 1989) when trying to pinpoint the date of the transaction Mr Gauci stated that Christmas lights were not up at the time the man bought the clothing. However, when giving evidence at trial he reversed his original statement and stated that there were Christmas lights, thereby claiming that the man had visited the shop in the middle of December (Day 31, page 4739).

5.2.1. Timeline of identification evidence

In the years between the Lockerbie bombing and the trial in 2000 Anthony Gauci was interviewed a number of times and made several statements about the identification of the man who bought the clothing from him. The timeline of this evidence will now be reviewed.

- 1st September 1989 – As well as detailing the items sold, Mr Gauci provided a description of the man who bought the clothing.
- 8th September 1989 – Mr Gauci was shown a photo-spread containing 23 photographs, no identification was made from this photo-spread.
- 13th September 1989 – Mr Gauci constructed an artist's impression and facial composite of the man who bought the clothing. He also states that he had seen the man about three months ago in a bar in Malta.
- 14th September 1989 – Mr Gauci identifies Mohamed Salam from a photo-spread as "similar" to the man who bought the clothing.
- 26th September 1989 – In a statement to the police Mr Gauci states that the person who bought the clothing in 1988 visited the shop again the day before (25th September 1989) and bought four elasticised girls dresses.
- 2nd October 1989 – Mr Gauci viewed a freeze frame image of Abu Talb and described him as "similar" to the man who bought the clothing.

- 6th December 1989 – Mr Gauci viewed a photo-spread of 12 photographs (one of which was Abu Talb) but failed to identify anyone.
- 5th March 1990 – Mr Gauci stated that Abu Talb "may" have been the man who bought the clothing after viewing his picture in a newspaper.
- 10th September 1990 – Mr Gauci viewed a photo-spread of an unspecified number of photographs (one of which was Abu Talb). Mr Gauci failed to definitively identify anyone, but identified three other men (Salem Mohd Abdel Hady Taha, Khalil, and Ayad Salama Hussein Mustafa Abueweiner) as "similar" to the man who bought the clothing.
- 15th February 1991 – Mr Gauci made a qualified identification of Al Megrahi for the first time, after picking his photograph out of a 12 photograph photo-spread. Mr Gauci stated that he was "similar" to the man who bought the clothing if he was a bit older.
- December 1998 – Mr Gauci viewed a photograph of Al Megrahi in a magazine article (*Focus*) which named him as the person responsible for the Lockerbie bombing.
- 28th February 1999 – Mr Gauci viewed a photograph of Al Megrahi in a magazine article (*It Torca*) which named him as the person responsible for the Lockerbie bombing.
- 13th April 1999 – Mr Gauci attended an identity parade at Kamp Zeist where he makes a qualified identification of Al Megrahi.
- July 2000 – Mr Gauci identifies Al Megrahi from the dock during the trial.

5.3. Expert witness reports on the identification evidence of Anthony Gauci

For the 2009 appeal against his conviction, Al Megrahi's defence team commissioned a number of expert witness reports to review the evidence given by Anthony Gauci. Three of these reports compiled by Professor Valentine, Professor Clarke, and Professor Canter (with the assistance of Dr Youngs and Dr Hammond) are reviewed here.

Each of the reports highlights serious inconsistencies with the evidence provided by Anthony Gauci, and discusses a number of reasons why Mr Gauci's evidence should be treated with caution. These reasons can be grouped under two headings, firstly the fallibility of Mr Gauci's memory after a time delay, and secondly the procedural impropriety

employed by the police when conducting line-ups, photo-spreads, and during questioning of Mr Gauci.

5.3.1. The fallibility of Mr Gauci's memory

As has already been stated, memory is susceptible to decay over time. Unless active practising techniques are employed, research has shown that there is a rapid forgetting of the information after the event (Wixted & Ebbesen, 1997). It must be remembered that nearly 10 months had passed between the man entering the shop and Mr Gauci being asked by the police to recall the events and the person who bought the clothing. A further 17 months passed before Mr Gauci identified Mr Megrahi for the first time. Although research has not examined the extent of forgetting over a 27 month period Shepherd, Ellis, and Davies (1982), found that after an interval of 11 months the correct identification rate was 11%, and therefore no better than chance. Furthermore, it is widely accepted that people are more likely to remember events that are unusual or uncommon in their experience. Low saliency events; those that occur on a daily basis, are much less likely to be remembered, particularly after a time delay. Canter (2009), Clark (2008), and Valentine (2008), all maintain that for a shopkeeper such as Mr Gauci, selling clothes must be an everyday occurrence, and therefore a low saliency event.

The time delay and the low saliency of the event both concur with the view of Canter (2009), that Mr Gauci did not have a clear memory of the event or the person, which explains why Mr Gauci's evidence changed and was so often contradictory. This may also explain why Mr Gauci's identifications of Mr Megrahi were qualified and never definitive. On the first occasion that Mr Gauci identified Mr Megrahi he stated that he "resembled" the man who bought the clothing. At the line-up parade at Camp Zeist Mr Gauci again qualified his identification with the following statements "I wasn't sure but the one who looked most like him was number 5," and "I'm not 100 percent sure." Furthermore, throughout the course of the police investigation, Mr Gauci made a number of similar qualified identifications of other men. Prior to the line-up of February 1991 (in which Mr Gauci identified Mr Megrahi), Mr Gauci picked out three other men who were "similar" to the man who bought the clothing.

5.3.2. Procedural impropriety

Professor Valentine highlights that procedural impropriety was employed by investigating officers from the beginning of their contact with Mr Gauci (Valentine, 2008). From very early in the investigation (from at least the 26th September 1989) Mr Gauci realised that he was being questioned in relation to the Lockerbie bombing as he was shown fragments of blast damaged clothing. Indeed, Canter (2009) argues that this constitutes a poor interview procedure which pervades throughout the police investigation. Instead of allowing Mr Gauci to provide an open and detailed account of the clothing he remembers selling to the man, the police, by showing him the fabric, lead the witness and imply that this evidence is important. Thereby the cue of the fabric allows Mr Gauci to reconstruct a memory of selling shirts, which he had previously consistently denied doing.

In the eleven years between the event of selling the clothes and giving evidence at the trial of Al Megrahi, Mr Gauci was interviewed many times and subjected to repeat questioning. Research has suggested that repeated questioning can result in increased confidence in answers given, even if those answers are incorrect (Odinot, Wolters, & Lavender, 2009). Moreover, as Canter (2009) observes, the knowledge that his evidence was imperative to the investigation coupled with the repeated interviewing may have led Mr Gauci to believe that he needed to add more detail to his evidence in order to help the police, even though some of this evidence may not have been accurate.

All three reports cite problems with the line-up procedures as casting considerable doubt on the reliability of Mr Gauci's identification evidence. In particular it is argued that the photographs used in the line-up of the 15th February 1991 (the first time Mr Gauci identified Mr Megrahi) may have been biased against Mr Megrahi. Valentine (2008) contends that the quality of the picture of Mr Megrahi is poor compared to the others and therefore, stands out. Whilst Clark (2008) argues that there are serious issues with the composition of both the photo line-up of February 1991 and the line-up parade of April 1999, resulting in a bias against Mr Megrahi. In the photo line-up Mr Megrahi stands out as the closest in age and ethnicity to Mr Gauci's description. Furthermore, in the line-up parade, nine of the eleven fillers were on average thirteen years younger than Mr Megrahi. Clark (2008) maintains that based on age, eight of the eleven fillers should have been excluded. Of the three remaining

fillers, one had a Dutch name (could thereby be eliminated due to his ethnicity), and one was 5'3", therefore shorter than Mr Megrahi, who was also shorter than 6-foot or more description given by Mr Gauci in 1989.

It is clear from the documentation of the photo line-up and the line-up parade that a "blind" procedure was not carried out at either time. In particular during the 1991 photo line-up Valentine (2008) points out that there were four police officers present during the procedure who knew the location of the suspect in the line-up. Furthermore, the statements taken from this line-up procedure show that Mr Gauci was prompted by one of the police officers. After viewing the photo line-up, which is included in Appendix 1, Mr Gauci stated that all of the men in the line-up were younger than the man who had bought the clothing. DCI Bell told Mr Gauci to look at the photographs carefully and try to allow for any age differences. This serves as a very strong cue to the witness that he is expected to make an identification and that the suspect is present (Valentine, 2008). Research in this area suggests that prompting a witness during a line-up can result in an increase in mistaken identifications, particularly when the line-up composition is biased (Clark and Tunnicliff, 2001).

In the intervening years between Mr Gauci's "identification" of Mr Megrahi and the trial, whilst the authorities attempted to extradite Mr Megrahi and Khalifa Fhimah, the media published many stories and pictures of Mr Megrahi as the Lockerbie bomber. Indeed two of these publications (*Focus* and *It Torca*) were viewed by Mr Gauci a matter of months before he identified Mr Megrahi at the trial. This exposure to outside information, according to Valentine (2008) has allowed Mr Gauci to "learn" Mr Megrahi's appearance (pg. 50). This outside information would also serve to increase Mr Gauci's confidence that he had identified the correct person, which therefore raises further questions as to the validity of the identification at Kamp Zeist in 1999.

Mr Gauci made a final identification of Mr Megrahi, from the dock during the trial. However, it is argued that dock identifications are highly suggestive and should be treated with extreme caution for their lack of reliability. A review of six studies comparing show-ups (which is effectively what a dock identification is) and line-ups, have found that the show-up both increases the false identification rate whilst also slightly decreasing the correct

identification rate (Clark & Godfrey, 2009). Moreover, as Mr Gauci had been shown the Focus magazine article with Mr Megrahi's picture, naming him as the Lockerbie bomber, only minutes before, it is highly unlikely that Mr Gauci would not identify Mr Megrahi from the dock.

5.4. Chapter summary

Now twenty five years after the event, and with the only person convicted of the Lockerbie bombing deceased, we will perhaps never know who was actually responsible. We will moreover perhaps never know whether Mr Gauci truly recognised Mr Megrahi as the person he sold the clothing to. However, what can be said with some certainty is that the line-up identification of Mr Megrahi by Mr Gauci was flawed and biased towards the identification of Mr Megrahi as the culprit. Of particular concern in this instance is the fact that police officers who knew the identity and location of the suspect administered the line-up to Mr Gauci.

In addition to the possibility that Mr Gauci attended to the expectancy of the line-up administrator, could it also be argued that the procedural impropriety displayed by the police led Mr Gauci to attend to the demand characteristics of the situation? A number of aspects of Mr Guaci's evidence suggest that this was the case. Firstly, Mr Gauci continued to be a part of the investigation and to be repeatedly questioned and interviewed by authorities even though it was against the express wish of his family. Secondly, it is clear from an early stage that Mr Gauci was aware his evidence was an important part of a high profile international criminal investigation. Thirdly, Mr Gauci changed aspects of his testimony possibly to coincide with cues he received from the police officers. This all indicates that Mr Gauci was anxious to appear as the 'good eyewitness.'

The review of the literature in Chapters 1 to 3 has suggested that individuals are inherently susceptible to bias. Firstly, it has been highlighted that experimenters can influence the responses of their participants through their expectations. Secondly, subjects can bias their responses with their attention to the demand characteristics of the procedure, and their susceptibility to the influence of the experimenter. Thirdly, in addition to influencing an eyewitness identification with their knowledge of the location of the suspect, line-up administrators have also been shown to influence the identifications from their line-ups

through their instructions and comments to the subject, the foils chosen, and the presentation method of the line-up. Finally, from Chapter 4 it is clear that an eyewitness' memory can be imperfect; a range of estimator and system variables can affect the quality of the memory. Therefore, the presence of a memory may confound the influence of the experimenter. In Chapter 6 then, a case will be made for examining experimenter expectancy effects without the memory component.

Chapter 6: Isolating the Experimenter Expectancy Effect: Removing the memory component

6.1. The antecedents of the research

In order to highlight the impact of conducting a non-blind line-up procedure Professor Canter devised an experiment to examine the effect of an informed administrator on respondent choice. This experiment was detailed in the report prepared by Canter and colleagues for Mr Megrahi's appeal. Using the photographs from the line-up of the 15th February 1991, two groups of participants were asked to guess which person from the line-up was a terrorist. In one group the person who conducted the line-up was informed of the position of Mr Megrahi; in the second group they were not informed. The results of this small experiment are particularly striking, out of the 56 participants over a quarter picked out Mr Megrahi, and all of those participants were in the informed condition (Canter, 2009). Therefore, even without a memory of Mr Megrahi, participants identified him as a terrorist at a much higher rate than chance when the administrator of the line-up was informed of his location.

To explore this phenomenon further and to improve on the validity of the first study a second, larger experiment was devised. In addition to a larger sample size, the researchers also included a comparison condition whereby a different target was chosen. To conduct the experiment ten facilitators were recruited, six facilitators were assigned to condition A, and four to condition B. Each facilitator was required to recruit two administrators. In this experiment there were four experimental conditions:

- AI – Three facilitators and six administrators. Both the facilitator and the administrator were informed that the suspect was in location number 8.
- AU – Three facilitators and six administrators. The facilitator was informed that the suspect was in location number 8, but the administrator was not informed.
- BI – Two facilitators and four administrators. Both the facilitator and the administrator were informed that the suspect was in location number 4.
- BU – Two facilitators and four administrators. The facilitator was informed that the suspect was in location number 4, but the administrator was not informed.

Each administrator recruited twenty participants. The sample of 400 participants was shown the same photo-spread from the first experiment. The administrators asked each participant to identify the man they thought was most likely to be the man convicted for the Lockerbie bombing. The results of this study show the same effect as the first study, what the authors have termed a "fundamental administrator effect" (Canter, Youngs & Hammond, 2009). Although the effect is not as strong as the first experiment, the results are still compelling. In condition AI 25% of the participants identified target eight, compared to 10.83% in the uniformed condition (AU). Similar results were found for condition B. In the Informed condition (BI) 26.25% identified target four, compared to 12.50% in the uninformed condition.

6.2. The rationale for removing the memory component

Despite the significant implications of investigator bias on the reliability of eyewitness identifications, there has been a paucity of research in this specific area (Greathouse & Kovera, 2009; Haw & Fisher, 2004; Phillips et al., 1999). As Canter, Youngs, and Hammond (2013) stress, the research that has been conducted to date, instead of providing clarity, has only served to muddy the water. Studies have combined different factors of line-up administration, such as different levels of contact between the administrator and witness, or different levels of motivation between the administrators. Furthermore, when devising the experiments detailed above, Canter et al., (2009) argued that the ability of an administrator to influence a witness depends partly on the susceptibility of the witness to that influence. In an eyewitness identification procedure, the memory for the event may confound the administrator's influence. It is suggested that the strength of the memory trace is an important factor in choosing rates from a line-up, if a witness has a poor memory of the suspect they will be unwilling to choose from a line-up. The poor memory trace may then make the witness more or less susceptible to the administrator's influence (Russano, Dickinson, Cass, Kovera, & Cutler, cited in Russano, Dickinson, Greathouse, & Kovera, 2006). The differences in the variables in the studies and the confounding variable of the memory for the suspect has resulted in, in the small number of studies reported, a variation in the magnitude of the administrator effect and therefore an un-clear picture of administrator influence.

Therefore, in order to examine the true extent of the administrator effect, it is necessary to remove the memory component. Canter proposed the novel experiment detailed above, which differs from the traditional mock witness paradigms. In such paradigms mock witnesses are given a description of the perpetrator and then asked to select from a line-up who they think the culprit is. Mock witness paradigms are used to test the validity of the construction of the line-up. If all members of the line-up are selected equally then the line-up is well constructed, however, if one member of the line-up is selected more often than chance then there is a bias in the construction of the line-up. There is an important distinction between mock witness experiments and the research conducted by Canter reviewed here, and the present research. Mock witness experiments provide 'witnesses' with some form of information about the features of the suspect, whereas the approach proposed by Canter asks participants to make a selection with no such information or background knowledge.

6.3. The aims of the present research

From the review of the literature in the preceding chapters we have seen that experimenters bias their data with their expectations and influence. Subjects also bias their responses with their attention to the demand characteristics of the experiment and their susceptibility to the influence of the experimenter. Line-up administrators can produce biased line-up procedures through their knowledge of the location of the suspect and their expectancy that the witness will identify the suspect. Moreover, bias in a line-up procedure is exacerbated and mediated by the memory of the event and the identification procedure. We have also seen in a real-life example what can happen when biased identification procedures are employed.

This research however, does not aim to replicate the research conducted on the effects of investigator bias in a line-up situation (Greathouse & Kovera, 2009; Haw & Fisher, 2004; Phillips et al., 1999). This research is not a standard memory paradigm experiment, there is not a staged event for participants to view and then identify the suspect with varying conditions of administrator knowledge. Instead this research examines experimenter expectancy effects in a similar way to Rosenthal's original person perception photo-rating task. In those experiments when experimenters were led to expect ratings of photos in a

particular direction, they indeed obtained ratings from their subjects in that direction. In this research administrators are led to expect the participant to pick a particular person from the line-up. Consequently, if participants pick the person expected by the experimenter more often than the other members of the line-up, this provides evidence for an experimenter expectancy effect. In Rosenthal's original experiments, experimenters influenced participants to rate photographs as more or less successful. In this experiment, if there is an experimenter expectancy effect present, they will be influencing the participant to pick a particular photo. Therefore, this research does link to the research concerning biased line-up procedures, as this is the same type of influence that a line-up administrator, who knows the identity of the suspect, can exert on an eyewitness.

As the participants in this research are being influenced to do different things, then the processes underlying the influence highlighted by Rosenthal may also be different. It is argued that the influence and processes proposed by Rosenthal in his classroom studies and his studies with rats is fundamentally different to the processes involved in a procedure where a respondent is requested to make a choice by an administrator who is informed of the correct response. Indeed it is argued that the four factor (climate, input, output, and feedback) theory of expectancy effects proposed by Rosenthal (1994) cannot be extended to explain administrator effects in a line-up procedure. In the classroom studies teachers had an extended period of interaction with their pupils, during which they can subtly modify their behaviour, by creating a warmer socio-emotional climate, and giving more input and personal feedback to students they expect to excel. The line-up procedure however is a short interaction. The line-up administrator does not have the time to influence the participant by altering the climate of the interaction. It is therefore the aim of this research to expand upon the research conducted to date by Canter and colleagues and to explore in greater depth the process involved in transmitting the experimenter expectancy effect during a short interaction between the administrator and the participant.

To achieve this, the research will focus on the interaction between the administrator and the participant. Using an experimental design the knowledge of the location of the suspect will be varied amongst line-up administrators. Participants will then make a choice from the line-up. In order to further the understanding of the process of the influence, administrators and participants will complete the Fundamental Interpersonal Relations Orientation –

Behaviour questionnaire (Schutz, 1958), a full description of which is given in Chapter 7. Suffice to say at this point that this research will attempt to uncover the interpersonal behaviour that may account for experimenter expectancy effects. Furthermore, as has been shown in the review of the literature concerning expectancy effects, demand characteristics, and biased line-up procedures, experimenters can make seemingly innocuous statements to their subjects, which nonetheless have an effect on the responses of those subjects. This research will therefore take an interest in the verbal interaction between the experimenter and the subject during the experiment.

6.3.1. The interpersonal relations factor

As part of the body of research that Rosenthal and his colleagues have conducted in the area of experimenter expectancy effects, they have studied the interaction between the experimenter and the subject, in order to determine behaviours that are associated with greater expectancy effects. The data for these analyses has come from accounts of the experimenter's behaviour from their subjects, and observations of experimenter behaviour from recordings of the interaction between the experimenter and the subject. Rosenthal (1976) has suggested that experimenter expectancy effects are communicated through kinesic and paralinguistic means. From direct observations of the experimenter by the subject, those experimenters with the greatest expectancy effects were judged to display subtle movements of the legs and head (kinesic communication), and speak in an expressive tone of voice (paralinguistic communication). This communication is aided by the interpersonal style of the experimenter. Experimenters who achieved the greatest expectancy effects were judged to have a number of interpersonal traits; they tended to be more professional and business-like with their subjects, but also more interested, enthusiastic, relaxed and personable than experimenters who were not as influential.

Similar results were found when observers of the interaction between the experimenter and the subject were asked to rate the experimenter (from 1 to 10) on five variables; dominance, professionalism, friendliness, likeability, and activity. Those experimenters who the observers judged as more professional, more dominant, less hyperactive, and more likeable, demonstrated greater expectancy effects. Rosenthal (1976) suggests that observers view experimenters as less professional and less in control if they are hyperactive. The

professionalism, dominance, and likeability traits identified by the observers relate to the characteristics professional, business-like, and personable identified by the subjects. Overall then it would appear that experimenters exert the most influence over their subjects when they are likeable professionals, who are in charge of the situation, and relaxed in their movements. As Harris and Rosenthal (1985) point out in their meta-analysis, there are numerous behaviours which can mediate expectancy, some of which are yet to be discovered. This research will therefore consider how the mediating variable of interpersonal behaviour, as measured by the FIRO-B questionnaire, can aid the communication of expectancy effects.

6.3.2. Research hypotheses

This research is predominantly exploratory in nature. With regard to the link between interpersonal relations and experimenter expectancy effects, the FIRO-B questionnaire has not been used as the test instrument in this area before. Therefore, fixed hypotheses about the results of the FIRO-B will not be made, although the research will be looking for differences between participants in their FIRO-B scores. In particular the research will look for differences in the FIRO-B scores for those who have been subject to the administrator effect, compared to those who have not. This research will also examine whether some administrators are more likely to influence participants than others, and whether some participants are more susceptible to influence than others. Regarding the experimenter expectancy effects, more concrete hypotheses can be proposed. It is therefore hypothesised that;

- Those participants in the informed condition will identify the target more often than those participants in the uninformed condition and the control condition.
- Those participants in the uninformed condition will identify the target more often than those in the control condition.

Examining the verbal interaction between the experimenter and the participant is again exploratory. However, if the presence of an experimenter expectancy effect is established then it would be hypothesised that some evidence of this would be apparent in the verbal interaction between the Administrator and the participant.

Chapter 7: Methodology

7.1. A methodology for studying administrator effects

There are a number of methodological issues to consider when conducting experimental research. The psychological experimental model has been borrowed from the natural and physical sciences. In eyewitness research this model has allowed researchers to empirically test aspects of the line-up. This experimental paradigm has also given researchers control over which aspects of the line-up to manipulate. Control of extraneous factors is the precise reason why the experimental procedure has been used in this instance. This research is interested in the effect of administrator knowledge of the location of the suspect. It is therefore important that participants undergo the same experimental procedure, where only the administrator's knowledge of the location of the suspect is varied. Controlling for extraneous variables is not the only advantage of the experimental procedure. The random assignment of participants into the different experimental conditions, alongside the control of the variables, allows researchers to infer a causal relationship from a correlation between two variables (Carlsmith et al., 1976).

Despite the benefits of experimental research there are concerns about the ecological validity of such research. One area of contention is the over-use of students as subjects in these experiments. Indeed, research of human behaviour has been described as "the science of the behaviour of sophomores" (McNemar, 1946, p. 333), due to the overrepresentation of undergraduate students as subjects in psychological research. However, reviews in this area (Pozzulo, 2006; Bartlett & Memon, 2006) have found that students tend to perform better than other members of the public, and therefore underestimate the observed effects, such as suggestibility. In this research, although a proportion of the participants are students, they are not exclusively so.

Notwithstanding student subjects, concern has been raised generally at the representativeness of research conducted with the "volunteer subject" (Rosnow & Rosenthal, 1997, p.89). Precisely what motivates people to volunteer for an experiment could affect the responses that they give or the behaviour that they exhibit. There is also the possibility that the volunteer subject differs in some fundamental way from the population to which the research wishes to generalise to. Whilst in a strict sense the participants in this

research did not volunteer to be part of the experiment, they did agree when asked to take part. We do not know how many people were asked by the administrators and who refused to take part, and whether these individuals would have behaved or responded in a fundamentally different way to the present sample. Therefore, caution must be taken when generalising the results to a wider population.

A criticism often levelled at the type of eyewitness research reviewed above is that the experimental situation cannot adequately (for valid ethical reasons) replicate the stress experienced by actual eyewitnesses. Although, when research has put the witness under stress (Morgan et al., 2004) they have found that stress proves a hindrance when encoding memories. Finally critics argue that the experimental line-up cannot reproduce the real-world significance or consequences of taking part in line-up procedure. Archival research of real cases has shown, nevertheless, that even with serious crimes eyewitnesses still identify a known innocent filler on average 30% of the time (Wright, & Skagerberg, 2007). Moreover, as explained in Chapter 6, this study forms part of a group of new and novel experiments which removes the memory component and attempts to "disentangle" (Canter et al., 2013, p. 85) the experimental expectancy effect from the memory process. Therefore, as the participant is not being tested on their memory of an event those concerns are not pertinent to this research.

7.1.1. Ethics and confidentiality

In compliance with university policy, the author submitted a proposal of the research to the Human and Health Sciences School Research Ethics Panel (SREP). SREP approved the research as meeting their ethical guidelines for conducting research with human participants (see Appendix 2 for the SREP approval certificate). The main ethical issues concerning this research were ensuring the confidentiality and anonymity of the people involved, and providing support to participants who may experience distress as a result of taking part in the study. Firstly the research guaranteed the anonymity of the individuals conducting the line-ups. The identity of each Facilitator and Administrator was known only by the author, and they were given a corresponding code for the analysis. The anonymity of the participants was also maintained by ensuring that they did not provide any identifiable information to the administrator conducting the experiment. Each participant was given a

code which delineated who their administrator was and which condition they were in. Secondly, to ensure the confidentiality of the data, hard-copies of completed questionnaires and answer sheets were stored in a locked cupboard that only the author had access to. The recordings of the experiments were transcribed and the original recordings deleted, and electronic data was stored on a password protected computer.

Although the Lockerbie bombing occurred over twenty years ago, as it involved a terrorist attack, it is possible that the research may inadvertently trigger emotional or psychological distress for the participant. In order to mitigate this distress the consent form (Appendix 3) emphasised that the participant was under no obligation to take part in this experiment, and if they felt uncomfortable they could withdraw their participation. The telephone number and e-mail address of the university counselling service, along with the authors e-mail address was provided on the consent form in case the participant wished to seek help, or ask questions about the research.

7.2. The recruitment of Facilitators and Administrators

In the first stage of the experiment the experimenter (author) randomly recruited six people to act as Facilitators (four females and two males). All Facilitators were told that the aim of the experiment was to research different types of line-up. The Facilitators were divided into three groups (informed, informed of alternative, and control) and were briefed separately by the experimenter. Three of the Facilitators were randomly assigned into group A or the informed group, two of the Facilitators were assigned to group B or the informed of alternative group and one Facilitator was assigned to group C or the control group. Each Facilitator was asked to recruit as many people as possible to act as administrators to conduct the experiments.

There were no specific guidelines for the recruitment of administrators; they were just required to be over the age of 16. The Facilitators recruited differing numbers of administrators, in total sixteen administrators were recruited (twelve females and four males). The three Facilitators in the informed condition (A) recruited seven administrators between them, with Facilitator one recruiting four administrators, Facilitator two recruiting two administrators, and Facilitator three recruiting one administrator. In the informed of alternative condition (B) Facilitator four recruited five administrators and Facilitator five

recruited one administrator. Finally in the control condition (C) Facilitator six recruited three administrators.

There were two reasons for using Facilitators to recruit administrators, who in turn recruited participants. Firstly, this procedure allowed the true rationale for the study to remain concealed from the administrators, as the Facilitators who briefed them were themselves unaware. Furthermore, using separate Facilitators for each condition (informed, informed of alternative, and control), rather than the author who knew the correct location of the suspect, allowed for firstly, a true double-blind control condition where both the Facilitator and the administrators were uninformed of the location of the suspect. Secondly, this allowed for the informed of alternative condition. In this condition the Facilitators and administrators were informed that the suspect was number four in the line-up, which is incorrect. If the author had recruited the administrators in this condition their knowledge of the correct location of the suspect may have confounded the communication of the incorrect location.

The second reason for employing Facilitators was to investigate whether they communicated their knowledge of who the suspect was to administrators who were supposed to be blind to the location of the suspect. By doing this the study becomes more akin to line-ups conducted in real world settings. In Chapter 3 the financial and practical objections of the police to conducting double-blind line-ups were discussed. However, even if these objections can be overcome and a blind administrator conducts the line-up, the possibility that the administrator may have been briefed by an investigator who is not blind to the identity of the suspect has not been considered. This study, by using Facilitators, therefore aims to investigate the implication of a non-blind investigator asking a blind investigator to conduct a line-up.

7.3. The sample

The participants of the study consist of a convenience sample of 526 people from the general population. Again there were no specific requirements for the recruitment of participants; administrators were asked to recruit people over the age of sixteen, and to try to recruit an equal number of males and females from as diverse a background as possible. Table 7.3.1 shows the number of participants recruited by each administrator.

Table 7.3.1: The number of participants recruited by each administrator

Administrator	Number of participants	Percentage %
1	38	7.2
2	30	5.7
3	32	6.1
4	40	7.6
5	40	7.6
6	5	1.0
7	18	3.4
8	32	6.1
9	40	7.6
10	40	7.6
11	40	7.6
12	39	7.4
13	22	4.2
14	40	7.6
15	30	5.7
16	40	7.6
Total	526	100

7.3.1. General and social characteristics

Gender and Age

The aim of the research was to include a representative sample of the general population, by recruiting male and female participants in roughly equal numbers, and by sampling a wide age range of people. The sample of 526 participants consists of 245 (46.6%) male participants, and 278 (52.9%) female participants (three participants failed to record their gender). The aim of the research was also to sample a wide range of ages. The age range of the participants in the sample is 72 years, with ages ranging from 16 years to 88 years. The mean age of the sample is 31.74 years (Standard Deviation = 14.47 years), with a median age of 25 years.

Education and Occupation

The participants were asked about their highest level of education and their occupation. Table 7.3.1.1 shows the level of education of the participants. For 28 participants (5.3%) there is no data on the highest level of education. The table shows that approximately 90%

of the participants had gained some formal educational qualifications, with the majority (32.7%) of participants gaining GCSE's or equivalent. Only 1% of the sample had gained no formal qualifications.

Table 7.3.1.1: The educational attainment of the participants

Educational attainment	Number of participants	Percentage %
No Qualifications	5	1.00
GCSE/O-Level/NVQ	172	32.70
A-Level/BTEC	120	22.80
Degree/Diploma	157	29.80
Masters/PhD	44	8.40
Total	498	94.70

Note: 28 participants failed to indicate their educational attainment

The Standard Occupational Classification (2010) from the Office for National Statistics was used to categorise the occupations of the participants. Table 7.3.1.2 shows the occupational categories of the sample, twelve participants did not indicate their occupation, and three of the occupations could not be categorised. The majority of the sample, 30.6%, are students. A further 28.3% of the sample are grouped in the first three categories, Managers/Directors/Senior Officials, Professionals, and Associate Professionals and Technical occupations. The rest of the sample are quite evenly distributed through the elementary occupations, sales and customer service occupations, administrative and secretarial, and skilled trade occupations. A small percentage of the sample consists of caring, leisure and other service occupations, process, plant and machine operatives, unemployed participants, and retired participants.

Table 7.3.1.2: The occupational categories of the participants

Occupational category	Number of participants	Percentage %
Managers/Directors/Senior Officials	20	3.8
Professionals	80	15.2
Associate Professionals and Technical	49	9.3
Administrative and Secretarial	30	5.7
Skilled Trade	29	5.5
Caring/Leisure and Other Services	20	3.8
Sales and Customer Service	41	7.8
Process/Plant and Machine Operatives	3	0.6
Elementary Occupations	46	8.7
Student	161	30.6
Unemployed	13	2.5
Retired	19	3.6
Total	511	97.1

7.4. The procedure

As stated above, all of the Facilitators were misinformed as to the true rationale for the experiment. In their individual briefings they were informed that the study was to research different types of line-up administration. The three groups of Facilitators also received slightly different instructions. The Facilitators in the informed condition (A) were informed of the correct location of the target suspect, photograph number eight in the line-up. The Facilitators in the informed of alternative condition (B) were informed that the target suspect was number four in the line-up. The Facilitator in the control condition (C) was not informed of the location of the suspect. The Facilitators were supplied with instructions according to their condition.

7.4.1. Instructions for Facilitators in Condition A

You are taking part in an experiment examining the effects of type of line-up on identifications made from a police photographic line-up. What we would like you to do is enlist 'line-up administrators' – people to help you conduct this experiment. To half of these, please give the instructions for condition 'A1', and to the other half give the instructions for condition 'A2'. They will each ask a number of 'respondents' to pick the

person that they think is most likely to be the man convicted of the Lockerbie bombing from a photographic line-up consisting of 12 pictures. They should do this using the instructions for respondents given below, so please make sure these are available to them. The individual who was actually held responsible for the bombing is the person in photo number 8, although only the line-up administrator in condition A1 will know this. In condition A2 the administrator will be blind to the position of the target within the line-up. Each administrator will conduct both simultaneous and sequential line-ups and will therefore need both copies of the photographic line-up to use, and the sheet for recording responses and participant information.

7.4.2. Instructions for Facilitators in Condition B

You are taking part in an experiment examining the effects of type of line-up on identifications made from a police photographic line-up. What we would like you to do is enlist 'line-up administrators' – people to help you conduct this experiment. To half of these, please give the instructions for condition 'B1', and to the other half give the instructions for condition 'B2'. They will each ask a number of 'respondents' to pick the person that they think is most likely to be the man convicted of the Lockerbie bombing from a photographic line-up consisting of 12 pictures. They should do this using the instructions for respondents given below, so please make sure these are available to them. The administrator in condition B2 will not have any prior beliefs about who might be the 'suspect' – the person convicted of the bombing – but in condition B1 the administrator will be told that it is the individual in position 4 in the line-up, even though it's not. Each administrator will conduct both simultaneous and sequential line-ups and will therefore need both copies of the photographic line-up to use, and the sheet for recording responses and participant information.

7.4.3. Instructions for Facilitators in Condition C

You are taking part in an experiment examining the effects of type of line-up on identifications made from a police photographic line-up. What we would like you to do is enlist 'line-up administrators' – people to help you conduct this experiment. To these administrators, please give the instructions for condition 'C1'. They will each ask a number of 'respondents' to pick the person that they think is most likely to be the man convicted of

the Lockerbie bombing from a photographic line-up consisting of 12 pictures. They should do this using the instructions for respondents given below, so please make sure these are available to them. Each administrator will conduct both simultaneous and sequential line-ups and will therefore need both copies of the photographic line-up to use, and the sheet for recording responses and participant information.

7.4.4. Instructions for the line-up administrators

Facilitators were given the instructions for line-up administrators for their group and were asked to recruit and brief the administrators individually. They were instructed to provide each administrator with a copy of the instructions and a copy of the guidelines for conducting the experiment as detailed below.

Instructions for Line-Up Administrator in Condition A1:

You have been provided with two photographic line-ups, line-up X1 and line-up Y1. Both line-ups have also been provided in a sequential format. The line-ups consist of photographs of 12 men, one of which is the man believed to be responsible for the Lockerbie bombing. This is the man in position number 8. What we would like you to do is to ask different people (preferably an equal split of males and females) to look at the pictures and decide which of the people they think is most likely to be the man who did it, using the instructions specified. Please record the age, gender, occupation, and educational achievement of each respondent and the number they pick as the most likely culprit on the sheet provided. Each respondent must make a selection. Hopefully most people should correctly identify the suspect from the line-up.

Instructions for Line-Up Administrator in Condition A2:

You have been provided with two photographic line-ups, line-up X1 and line-up Y1. Both line-ups have also been provided in a sequential format. The line-ups consist of photographs of 12 men, one of which is the man believed to be responsible for the Lockerbie bombing. What we would like you to do is to ask different people (preferably an equal split of males and females) to look at the pictures and decide which of the people they think is most likely to be the man who did it, using the instructions specified. Please record the age, gender, occupation, and educational achievement of each respondent and the number that they

pick as the most likely culprit on the sheet provided. Each respondent must make a selection. Hopefully most people should correctly identify the suspect from the line-up.

Instructions Given to Line-Up Administrators in Condition B1:

You have been provided with two photographic line-ups, line-up X1 and line-up Y1. Both line-ups have also been provided in a sequential format. The line-ups consist of photographs of 12 men, one of which is the man believed to be responsible for the Lockerbie bombing. This is the man in position number 4. What we would like you to do is to ask different people (preferably an equal split of males and females) to look at the pictures and decide which of the people they think is most likely to be the man who did it, using the instructions specified. Please record the age, gender, occupation, and educational achievement of each respondent and the number they pick as the most likely culprit on the sheet provided. Each respondent must make a selection. Hopefully most people should correctly identify the suspect from the line-up.

Instructions Given to Line-Up Administrators in Condition B2:

You have been provided with two photographic line-ups, line-up X1 and line-up Y1. Both line-ups have also been provided in a sequential format. The line-ups consist of photographs of 12 men, one of which is the man believed to be responsible for the Lockerbie bombing. What we would like you to do is to ask different people (preferably an equal split of males and females) to look at the pictures and decide which of the people they think is most likely to be the man who did it, using the instructions specified. Please record the age, gender, occupation, and educational achievement of each respondent and the number that they pick as the most likely culprit on the sheet provided. Each respondent must make a selection. Hopefully most people should correctly identify the suspect from the line-up.

Instructions Given to Line-Up Administrators in Condition C1:

You have been provided with two photographic line-ups, line-up X1 and line-up Y1. Both line-ups have also been provided in a sequential format. The line-ups consist of photographs of 12 men, one of which is the man believed to be responsible for the Lockerbie bombing. What we would like you to do is to ask different people (preferably an equal split of males and females) to look at the pictures and decide which of the people they think is most likely

to be the man who did it, using the instructions specified. Please record the age, gender, occupation, and educational achievement of each respondent and the number that they pick as the most likely culprit on the sheet provided. Each respondent must make a selection. Hopefully most people should correctly identify the suspect from the line-up.

7.4.5. Guidelines for conducting the experiment

All administrators were provided with a copy of the guidelines detailed below.

Procedure for conducting the line-ups:

All line-ups must be recorded with a dicta-phone. You need to give each participant a unique code so that their answer sheet can be matched with their recording. The participant code needs to be written on the answer sheet and recorded on the Dictaphone.

Condition 1

Twenty witnesses (10 male/10 female). Show half (10) of the witnesses a simultaneous line-up and half (10) of the witnesses a sequential line-up using photo array X1 (simultaneous) and X2 (sequential). Record their personal details and their choice on the sheet provided. Then give the witnesses the FIRO.

In this condition you need to record whether the witness took part in a simultaneous or a sequential line-up.

Condition 2

Twenty witnesses (10 male/10 female)

1. Show ten witnesses a simultaneous line-up using photo array X1. Record their personal details and their choice on the sheet provided. Then give them the FIRO. Then show them a sequential line-up using photo array Y2. Record their choice again.

2. Show ten witnesses a sequential line-up using photo array X2. Record their personal details and their choice on the sheet provided. Then give them the FIRO. Then show them a simultaneous line-up using photo array Y1. Record their choice again.

In this condition you need to record whether witnesses took part in option 1 or 2.

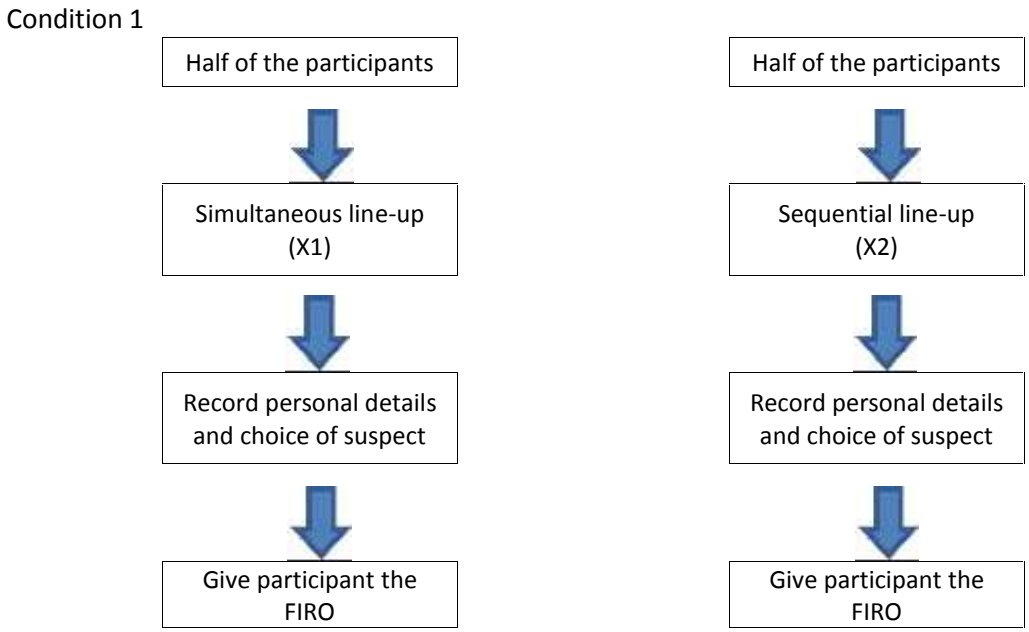
Procedure for conducting simultaneous line-ups

A simultaneous line-up is where the witness sees all the suspects at once. These are the X1 and Y1 photo-spreads. In this condition you need to place the photo-spread in front of the witness. The witness can have as much time as they need to identify the suspect.

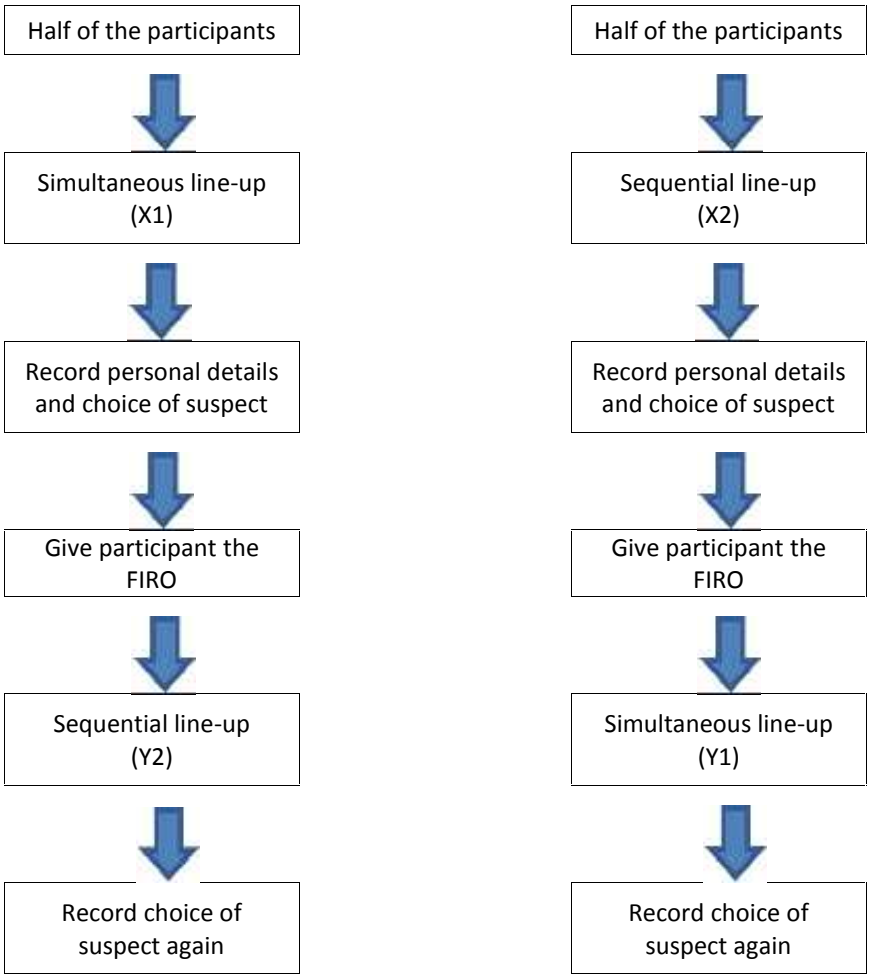
Procedure for conducting sequential line-ups

A sequential line-up is where the witness sees one suspect at a time. These are the X2 and Y2 photo-spreads. In this condition you need to place one photo at a time in front of the witness. The witness can have as much time as they need to look at each photo. They must make a decision on each photo whether that person is the suspect. If they choose a photo as that of the suspect they should not see the rest of the photos. If by the end of the photos they have not chosen a suspect you can show them the sequential photos again in the same order. It does not matter which order you present the photos, any random order is fine, but change the order for each witness. Figure 7.4.5.a. shows a pictorial representation of Conditions 1 and 2 of the experiment.

Figure 7.4.5.a: A pictorial representation of conditions 1 and 2



Condition 2



During their individual briefings the Facilitator ensured that the administrator understood how to conduct the line-ups, and the administrator had the opportunity to ask the Facilitator any questions they had about the procedure. Once the Facilitator was satisfied that the administrator understood how to conduct the line-ups the Facilitator provided them with a bundle of materials (detailed below) and the administrator could begin to recruit participants.

In order to obtain a diverse sample of participants, administrators were permitted to recruit participants from both within and outside the university. Line-ups were therefore conducted in a variety of settings, including university classrooms, business offices and administrators own homes. Despite the variety of settings administrators were told to conduct the line-ups privately, in a quiet room, with one participant at a time.

7.5. The materials

The Facilitators were firstly provided with their instructions corresponding to their allocated group. Once they had read through and understood the instructions they were given a bundle of materials to give to the administrators. Included in this bundle was a consent form for administrators to administer to participants (Appendix 3). The consent form explains that the researcher is conducting a study on the effects of different types of line-ups on eyewitness testimony, thus concealing the actual rationale for the research. The consent form briefly indicates that the participant will be required to look at different line-ups and also fill in a personality questionnaire. The consent form also indicates that the line-up procedure will be audio-recorded. The anonymity of the participants is guaranteed, the participants are asked to tick three boxes if they agree to take part in the experiment, consent for their responses to be used in further analysis and research, and agree for the experiment to be recorded. The consent form gives details of organisations and support services for participants to contact if they feel they have been affected by any issues raised in the study, and finally, the researchers details are given in order for the participants to ask any questions about the research.

Also included in the bundle of materials were the instructions and guidelines for administrators detailed above, and two photo-arrays. The first photo-array (Appendix 4) was the photo-line-up used in the Al Megrahi case in 1991, with Mr Megrahi appearing in position eight of the photo-array. This photo-array was provided in both a simultaneous (X1) and a sequential format (X2). The X2 photo-array was constructed by cutting out the photos from the X1 photo-array and backing them with cardboard to ensure they were the same size in both types of line-up. The second photo-array (Appendix 5) was also from the Al Megrahi case, but did not include a photo of Mr Megrahi, and therefore could be termed a target-absent line-up. This photo-array was also provided in a simultaneous (Y1) format and a sequential (Y2) format, with the same procedure utilised for the sequential photos.

An answer sheet was devised in order to record the participants' responses (Appendix 6). Participants were asked their age, gender, occupation, and highest level of education. Participants were then also asked if they had ever been asked to identify someone from a police line-up. Dependent on the condition of the experiment the answer sheet has the

space required to record the identifications of the participants. Finally the participant is asked to describe how confident they are in their identification on a five point Likert scale (1 = Very unconfident, 5 = Very confident).

7.5.1. The FIRO-B Questionnaire

The bundle of materials also contained the FIRO-B questionnaire (Schutz, 1958; Appendix 7). The Fundamental Interpersonal Relations Orientation (FIRO) is a theory of interpersonal personality developed by Schutz in the 1950's, originally for use in a military setting. Since then the theory has been revised, and expanded, and has been used in diverse settings including public schools (Schutz, 1976), family therapy (Doherty & Colangelo, 1984), and organisations (Schutz, 1994). During the extensive revision of the original theory a number of instruments within the FIRO family were developed. In particular the Elements of Awareness instruments which since 1984 have included a revised version of the FIRO-B, Element B: Behaviour, but also included were Element F: Feelings; Element S: Self; Element W: Work Relations; Element C: Close Relations; Element P: Parental Relationships; and Element O: Organizational Climate (Schutz, 1992; 1994).

In this instance it is the FIRO-Element B instrument that was utilised (from now on referred to as FIRO-B). Schutz (1992) states that the original FIRO-B was designed to identify individuals who were compatible with one another and who would work well together, by predicting the interaction between two individuals. However, the FIRO-B instrument has been taken up by others and has been widely used in training, human relations and personality arenas. The FIRO-B questionnaire gathers insights into how an individual's needs for inclusion, control, and affection can shape their interactions with others. The FIRO theory is based on the facet framework originally developed by Guttman (1959). In the original theory Schutz proposed three *Form* domains for describing interpersonal personality, Control, Affection, and Inclusion. During the revision of the FIRO-B Affection became known as Openness, as it was felt that Affection related more to a feeling not a behaviour. Schutz (1992) describes the Control scale as measuring the amount of control people desire to have over other people, indicating that some people enjoy control and responsibility over other people, whilst others actively seek situations where they have no control or responsibility. The Openness scale measures the extent to which people desire to

be open with others, with some preferring relationships where they confide in people, whilst others prefer impersonal relationships. Finally, the Inclusion scale measures how much contact people seek with others, with some people seeking to be part of a group, whilst others seeking more solitary pursuits.

In conjunction with these three *Form* domains Schutz (1992) described a *Mode* facet with two levels, *Expressed "*what I do toward you" and *Received* "what I get from you" (p. 918)*,* which describes a persons' tendency to convey *Form* behaviour or obtain *Form* behaviour from another person. Finally there was an *Experience* facet with two levels, *Wanted* and *Perceived,* which describes whether a person actually perceives these experiences, or wants to experience these behaviours. A visual representation of Schutz's model can be achieved with the use of a mapping sentence (Borg & Shye, 1995) which can be seen in Figure 7.5.1.a.

Figure 7.5.1.a: Mapping sentence for Schutz's (1992) theory of interpersonal relations

The extent to which a person (p) agrees with statements about their

A. *Experience*	B. *Mode*	C. *Form*	
{1. Wanting }	{1. To Express }	{1. Control }	
{ }	{ }	{2. Openness }	in interactions with others
{2. Perceived }	{2. To Receive }	{3. Inclusion }	

1. Disagree
2
3
4
5
6. Agree

Where p is from a population of people P, who are not artificially selected with respect to the facets

The mapping sentence above indicates how the facets can combine to form a structuple (Borg & Lingoes, 1987), or a template for the formation of the instruments questions. The combination of the facets provides 12 (2 experience x 2 mode x 3 form) possible templates of questions for the FIRO-B. Schutz (1958) developed nine questions for each of these structuples resulting in a questionnaire consisting of 108 items.

In the present study the research was particularly interested in the actual experience of the participants and how they perceive themselves. Therefore the *Wanted* domain of the *Experience* facet was excluded, resulting in an instrument consisting of a more manageable and expeditious 54 items. The mapping sentence for the reduced FIRO-B with only the perceived experiences can be seen in Figure 7.5.1.b.

Figure 7.5.1.b: Revised mapping sentence for Schutz's (1992) theory of interpersonal relations, perceived experiences only

The extent to which a person (p) agrees with statements about their

A. *Experience*	B. *Mode*	C. *Form*	
{1. Perceived }	{1. To Express }	{1. Control }	
{ }	{ }	{2. Openness }	in interactions with others
{ }	{2. To Receive }	{3. Inclusion }	

1. Disagree
2
3
4
5
6. Agree

Where p is from a population of people P, who are not artificially selected with respect to the facets

The above mapping sentence results in six scales;

- Expressed Control – I control people
- Expressed Openness – I am open with people
- Expressed Inclusion – I include people
- Received Control – People control me
- Received Openness – People are open with me
- Received Inclusion – People include me

Examples of the items constructed from the above mapping sentence include item 1. *I seek out people to be with* (A1, B1, C3), item 2. *People decide what to do when we are together*

(A1, B2, C1), and item 15. *I am more comfortable when people do not get too close* (A1, B1, C2). Table 7.5.1.1 provides a list of the items from the FIRO-B which constitutes each scale.

Table 7.5.1.1: The item numbers of the FIRO-B for each scale

Scale	Item Numbers
Received Inclusion	4, 10, 16, 22, 28, 34, 40, 46, 52
Expressed Inclusion	1, 7, 13, 19, 25, 31, 37, 43, 49
Received Control	2, 8, 14, 20, 26, 32, 38, 44, 50
Expressed Control	5, 11, 17, 23, 29, 35, 41, 47, 53
Received Openness	6, 12, 18, 24, 30, 36, 42, 48, 54
Expressed Openness	3, 9, 15, 21, 27, 33, 39, 45, 51

The received inclusion scale describes an individual's tendency to obtain involvement and attention from others, an example item of this scale is: *People invite me to do things.* The expressed inclusion scale describes an individual's tendency to seek involvement and attention from others, an example item of this scale is: *I join social groups.* The received control scale describes an individual's tendency to be dominated by others, an example item of this scale is: *People decide what to do when we are together.* The expressed control scale describes an individual's tendency to dominate others, an example item of this scale is: *I am the dominant person when I am with people.* The received openness scale describes an individual's tendency to have intimacy and emotion from others, an example item of this scale is: *My close friends tell me their real feelings.* The expressed openness scale describes an individual's tendency to be intimate and emotional with others, an example item of this scale is: *I am totally honest with my close friends.* The FIRO-B questionnaire is measured on a 6-point Likert scale (1 = Disagree, 6 = Agree). Eight of the Openness items (15, 18, 21, 27, 30, 33, 42, 45,) are reverse coded.

Gender and age norms for the FIRO-B have been postulated by Schutz (1992). Female respondents have been found to score significantly higher on the received inclusion and received openness scales. They are also higher, although not significantly, on the expressed inclusion scale. For the expressed control and received control scales, males and females have not been found to differ significantly. For the age of the respondent, Schutz (1992) has found that the scores for the sociability factors, inclusion and openness, are higher for

younger respondents. Scores for expressed control appear to reduce with the age of the respondent, whereas scores for received control appear to increase with age.

Despite the popularity and widespread use of the FIRO-B questionnaire, some have questioned its construct validity, most notably in a series of papers by Hurley (1990; 1991; 1992). In contrast to Schutz's three domains of *Inclusion, Control,* and *Openness* research has suggested that *Inclusion* and *Openness* are not distinct components. Indeed, studies examining the FIRO-B have suggested that there are in fact two dimensions. Mahoney and Stasson (2005) have labelled the two dimensions Dominance (Schutz's Control) and Socio-Emotional Affect (Schutz's Inclusion and Openness combined). Mahoney and Stasson (2005) concur with the findings of Macrosson (2000) who labelled the two domains Control and Nurturance.

It should be noted however, that the research conducted above examined Schutz's original FIRO-B questionnaire, not the updated Element-B questionnaire. Furthermore, they used the coding framework originally developed by Schutz, and the scores of the six subscales even though there is no psychometric basis for the coding framework or the subscales. As described above the FIRO-B questionnaire was designed using a facet framework; however the studies that have questioned the validity of the FIRO-B have utilised a factor analysis approach. It is argued that in order to adequately examine the facet structure of the FIRO-B a facet analysis approach is required. Research that has utilised this type of analysis of the FIRO-B has been conducted by researchers of the International Research Centre for Investigative Psychology (IRCIP, 2010). In order to do this the reduced FIRO-B (the 54 items constituting the *Perceived* experience not the *Wanted* experience) was given to 186 respondents.

The completed FIRO-B questionnaires were analysed using a form of Multi-Dimensional Scaling (MDS), Smallest Space Analysis (SSA-I) which was developed by Guttman and Lingoes (Guttman 1968; Lingoes, 1973). SSA is now a well-established technique that has been used in diverse settings; it has particularly been applied to studying criminality by Professor Canter and colleagues. MDS and in particular SSA has been increasingly utilised in a variety of studies concerning different aspects of criminality, including arson (Canter & Fritzon,

1998), murder (Salfati & Canter, 1999), sexual offences (Häkkänen, Lindlöf, & Santilla, 2004), and stalking (Groves, Salfati, & Elliot, 2004).

The SSA programme correlates every variable with every other variable, the rank order of these correlations then create a triangular matrix. Whilst SSA utilises similar mathematical calculations as factor analysis and cluster analysis, the correlation coefficients of the SSA form a spatial representation in a statistical geometric space that is visual, and therefore more easily interpretable. The correlation coefficient plots each variable in the SSA space; those variables that are more closely inter-correlated are closer together in the SSA space. Therefore, those variables most likely to co-occur together are closer together in the plot, whilst those variables that do not co-occur are further away from each other. Under this assumption the thematic structure of a scale can be analysed by examining those variables which are grouped together on the SSA plot.

The coefficient of alienation (Borg & Lingoes, 1987) indicates how closely the rank orders of the distances between the points in the spatial representation relate to the rank orders of the correlations between the variables. The smaller the coefficient of alienation the better the fit between the derived SSA configuration and the correlation matrix from which it is derived. Zero is considered to be a perfect fit, a coefficient smaller than 0.15 is considered a good fit while a coefficient between 0.15 and 0.20 is considered a reasonable fit (Guttman, 1968). The particular advantage of using SSA as the form of MDS is that SSA is a non-metric procedure, it is therefore less sensitive to high or low absolute frequencies as it uses the rank order of the co-occurrences, for example, the relative differences between the variables rather than the absolute values (Canter and Youngs, 2009).

The study conducted by IRCIP researchers (2010) used SSA to examine the FIRO-B scale items and to delineate the thematic structures of the scale. It is hypothesised that those variables of the FIRO-B scale that measure the same structure will be closely correlated and therefore appear in the same area of the SSA plot, the closer that two points are to each other on the plot the more likely that participants will have provided similar answers to those questions. The SSA plot found that items of the *Inclusion* and *Control* facets each occupied distinct separate areas of the plot, however, the items of the *Openness* facet are distributed throughout the plot and do not occupy a distinct area. The results of this SSA

therefore indicate that, in concurrence with the research cited above, there are indeed two *Form* facets of the FIRO-B. It is argued that Schutz's *Openness* and *Inclusion* facets actually form one *Personal / Social Inclusion* facet, which Mahoney and Stasson (2005) labelled Socio-Emotional Affect, and Macrosson (2000) labelled Nurturance.

However, the facet analysis has allowed further insight into the thematic structure of the FIRO-B that was not possible with factor analysis. The facet analysis indicates that there is indeed a distinct *Mode* facet (*Expressed* and *Received*) that acts upon both the *Inclusion* and *Control* facets. Furthermore, the facet analysis also suggests that *Openness*, rather than just being an aspect of *Inclusion*, is actually a mediator of received behaviours. From this facet analysis the IRCIP researchers have suggested four interpersonal relations tendencies comprised of two facets:

- Expressed Personal / Social Inclusion
- Expressed Control
- Received Personal / Social Inclusion
- Received Control

Table 7.5.1.2: The item numbers of the revised FIRO-B structure

Scale	Item Numbers
Expressed Personal / Social Inclusion	1, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 46, 49, 52
Expressed Control	5, 11, 17, 23, 29, 35, 41, 43, 47, 53
Received Personal / Social Inclusion	3, 4, 6, 9, 12, 24, 36, 39, 48, 51, 54
Received Control	2, 8, 14, 15, 18, 20, 21, 26, 27, 30, 32, 33, 38, 42, 44, 45, 50

A number of items have changed position, below these changes are detailed.

- Expressed Personal / Social Inclusion – includes all of the original expressed inclusion items except item 43, and now includes items 10, 16, 22, 28, 34, 40, 46, 52 of the received inclusion scale.

- Expressed Control – includes all original expressed control items plus item 43 from the expressed inclusion scale.
- Received Personal / Social Inclusion – includes item 4 from the received inclusion scale, items 6, 12, 24, 36, 48, 54 from the received openness scale, and items 3, 9, 39, 51 from the expressed openness scale.
- Received Control – includes all of the original received control items, and also includes items 18, 30, 42 from the received openness scale, and items 15, 21, 27, 33, 45 from the expressed openness scale.

The question of the structure of the FIRO-B will be returned to in Chapter 10, where the structure will be explored with the FIRO-B data from the current study.

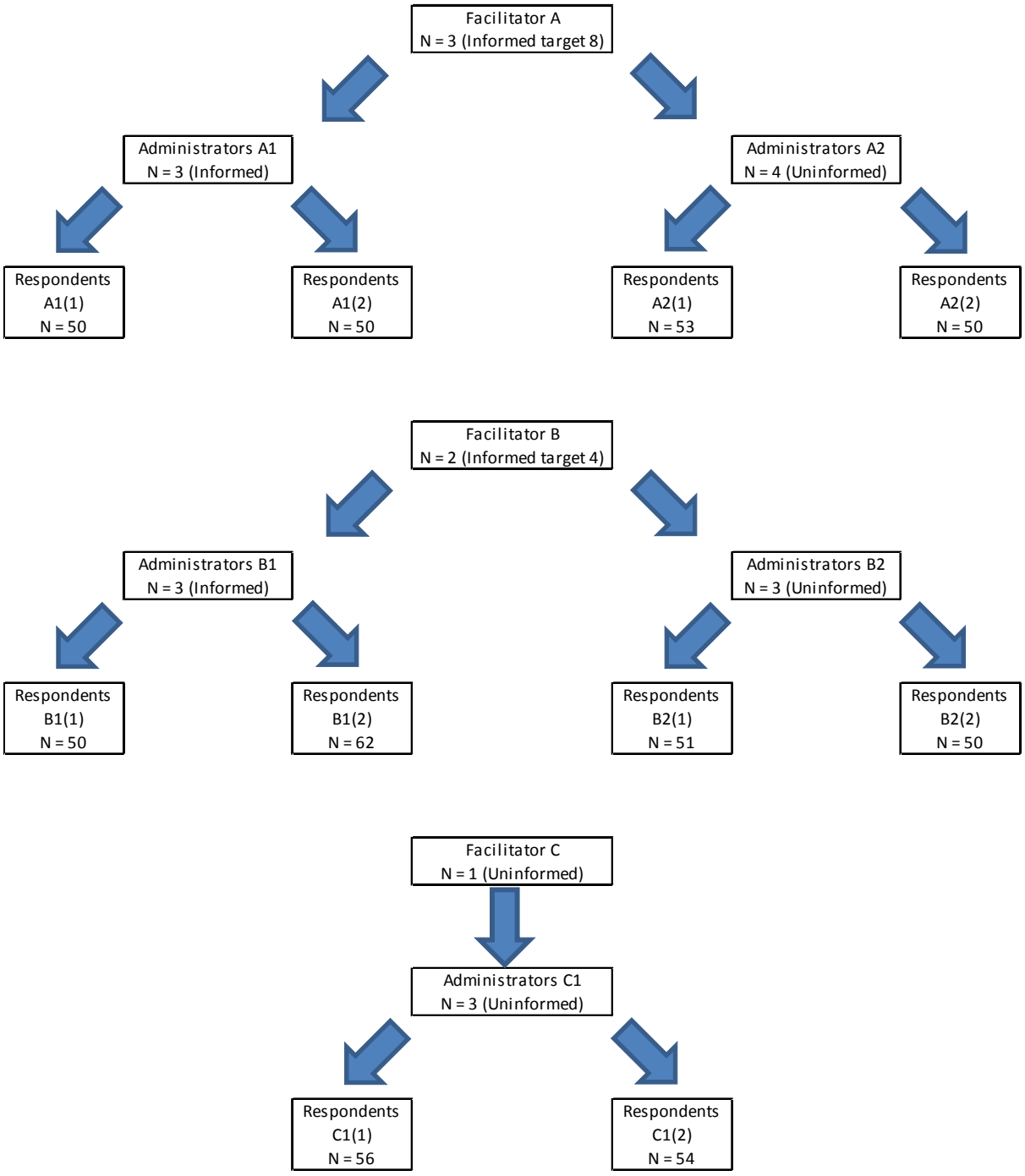
7.6. The design

There were five Facilitator/Administrator conditions which were labelled as follows:

- A1 – both the facilitator and administrator were informed of the correct location of the suspect in the line-up ("informed").
- A2 – the facilitator was informed of the correct location of the suspect, but the administrator was not ("blind").
- B1 – both the facilitator and administrator were informed of the alternative location of the suspect in the line-up ("alternative informed").
- B2- the facilitator was informed of the alternative location of the suspect, but the administrator was not ("blind").
- C1 – both the facilitator and administrator were not informed of the location of the suspect in the line-up ("double-blind").

Within the five Facilitator/Administrator conditions there were also two conditions of the experiment, therefore there were ten groups of respondents. Figure 7.6.a. shows the ten groups and the numbers of participants in each group.

Figure 7.6.a: The ten experimental groups and the number of participants in each group



This research utilises the actual photographic line-up containing Mr Megrahi's picture. Therefore, there is the possibility that instead of choosing Mr Megrahi from the line-up because of some expectancy or influence on the part of the administrator, the participants will choose Mr Megrahi because they recognise him. However, it is highly unlikely that the

participants in the study will have seen the picture used in the line-up before. Furthermore, even if the participants had seen recent pictures of Mr Megrahi in the media, it is unlikely that the participants would recognise the 25 year old photograph. Nevertheless, the design of the study will indicate whether the participants do indeed recognise Mr Megrahi. In conditions B and C Mr Megrahi's picture is present in the line-up, although the administrators are unaware that he is photograph number 8. If the participants in these conditions recognise Mr Megrahi we would expect high numbers of identifications of Mr Megrahi from these conditions. Furthermore, in Condition 2 (the target absent line-up) Mr Megrahi's photograph is not present. If photograph number 8 in group A of Condition 2 is chosen more often it will indicate that the expectancy of the administrator is influencing the participant to choose number 8.

In the proceeding chapters the analysis of the data collected will be conducted. In Chapter 8 the demographic data of the participants for each of the five administrator/participant conditions will be analysed, and the homogeneity of the participants in these conditions will be assessed. Chapter 9 will explore whether the frequencies of identifications indicate an experimenter expectancy effect. Although not strictly an identification, as there is no memory component to identify from, the term identification will be used for ease of understanding. In Chapter 10 the results of the FIRO-B questionnaire for both the participant and the administrator will be analysed. Finally in Chapter 11 the transcripts of the recordings of the experimental procedure will be analysed.

In order to conduct the analyses, distinctions will need to be made between the different groups of participants. For some analyses the five administrator/participant conditions (A1, A2, B1, B2, and C1) will be used. For some analyses a distinction between informed and uninformed participants will apply, whereby conditions A1 and B1 (informed) will be compared to conditions A2, B2, and C1 (uninformed). For other analyses a distinction will be made between the informed group (A1 and B1), the uninformed group (A2 and B2), and the control group (C1).

Analysis Section 1 – Examining the Sample

Chapter 8:

It is the aim of this chapter to document the demographic features (gender, age, occupational categories, and educational attainment) of the sample in each of the respondent conditions. The confidence ratings of the participants will also be analysed, and finally the homogeneity of the sample will be assessed.

8.1. The age and gender of the participants in each of the experimental conditions

Table 8.1.1 shows the percentage of male and female participants in each of the five conditions. It is clear from this table that there is a fairly even split of male and female participants in each of the conditions. There are proportionally more female participants in condition A1.

Table 8.1.1: Percentage and frequencies of male and female participants in each condition

Respondent condition	% Males	% Females	Total
A1	40.00% (40)	60.00% (60)	100% (100)
A2	49.51% (51)	50.49% (52)	100% (103)
B1	49.11% (55)	50.89% (57)	100% (112)
B2	47.52% (48)	49.50% (50)	97.02% (98)*
C1	46.36% (51)	53.64% (59)	100% (110)
Total	46.58% (245)	52.85% (278)	99.43% (523)

* 3 participants in this condition failed to record their gender

Table 8.1.2 shows the age range and mean ages of the participants in each of the five conditions. There is a good spread of ages in all of the conditions, with all of the means around 30 years of age. Condition A1 has younger participants than the rest of the sample with the lowest mean age, whilst condition B1 has older participants with the highest mean age.

Table 8.1.2: Age range and mean ages of participants in each condition

Respondent condition	Minimum age	Maximum age	Mean age (Std dev.)
A1	16	62	26.79 (13.75)
A2	16	88	30.65 (14.53)
B1	16	75	36.76 (14.72)
B2	16	63	33.29 (14.96)
C1	16	72	30.74 (12.66)

8.2. The occupational categories and educational attainment of the participants in each of the experimental conditions

Table 8.2.1 shows the percentage of participants from each occupational category for each of the five conditions. For conditions A1, A2, B1, and C1 the majority of the participants are students. However, for condition B2 there are more professionals, associate professionals, and sales and customer service workers than students. There is a high proportion of elementary workers in condition A1, and professionals in condition C1. Overall though there is a good spread of participants from different occupations in each of the five conditions.

Table 8.2.2 shows the level of educational attainment for participants in each of the five conditions. It is clear that participants in condition A1 are clustered more towards the lower end of the educational attainment table, with the vast majority obtaining GCSE's or NVQ's and relatively few obtaining degrees or higher degrees. This could be explained by the lower age of participants in this condition (mean age = 26.79 years). Condition B1 has a disproportionate number of participants attaining a higher degree. This again could be explained by the higher age of the participants in this condition (mean age = 36.76 years).

Table 8.2.1: Percentage of participants from each occupational category in each condition

	Respondent condition				
Occupational category	A1	A2	B1	B2	C1
Managers/ Directors/ Senior Officials	~	3.9% (4)	6.3% (7)	8.9% (9)	~
Professionals	5% (5)	5.8% (6)	20.5% (23)	16.8% (17)	26.4% (29)
Associate Professionals and Technical	5% (5)	9.7% (10)	7.1% (8)	16.8% (17)	8.2% (9)
Admin and Secretarial	7% (7)	3.9% (4)	7.1% (8)	3% (3)	7.3%(8)
Skilled Trade	5% (5)	8.7% (9)	4.5% (5)	5.9% (6)	3.6% (4)
Caring/Leisure and Other Services	1% (1)	2.9% (3)	2.7% (3)	7.9% (8)	4.5% (5)
Sales and Customer Service	2% (2)	7.8% (8)	8% (9)	14.9% (15)	6.4% (7)
Process/ Plant and Machine Operatives	1% (1)	1% (1)	0.9% (1)	~	~
Elementary Occupations	26% (26)	8.7% (9)	3.6% (4)	4% (4)	2.7% (3)
Student	40% (40)	40.8% (42)	24.1% (27)	12.9% (13)	35.5% (39)
Unemployed	5% (5)	1.9% (2)	1.8% (2)	2% (2)	1.8% (2)
Retired	~	4.9% (5)	4.5% (5)	5% (5)	3.6% (4)

Table 8.2.2: Levels of educational attainment for participants in each condition

	Educational attainment				
Respondent condition	No Qualifications	GCSE/ O-Level/ NVQ	A-Level/ BTEC	Degree/ Diploma	Masters/ PhD
A1	~	55% (55)	22% (22)	12% (12)	4% (4)
A2	4.9% (5)	33% (34)	23.3% (24)	35.9% (37)	2.9% (3)
B1	~	23.2% (26)	14.3% (16)	31.3% (35)	22.3% (25)
B2	~	33.7% (34)	26.7% (27)	25.7% (26)	9.9% (10)
C1	~	20.9% (23)	28.2% (31)	42.7% (47)	1.8% (2)

8.3. The confidence ratings of the participants

A measure of an eyewitness's confidence is a standard measure in eyewitness research and in actual line-up identifications conducted by the police. Although the current research is not a standard memory paradigm experiment, the author was interested to explore whether levels of confidence differed according to the experimental condition. Immediately following the identification each participant was asked to rate on a scale of 1 to 5 (1 = Very Unconfident, 5 = Very Confident) how confident they were in their identification. As can be seen from Table 8.3.1 over half of the participants (55.10%) were unconfident or very unconfident in their identifications, only a small percentage (19.6%), were confident or very confident in their identifications. Low confidence ratings are to be expected in a procedure where the participant has no memory of the event.

Table 8.3.1: Confidence ratings for all participants

	Frequency	Percentage %
Very Unconfident	150	28.50
Unconfident	140	26.60
Neither Confident or Unconfident	128	24.30
Confident	90	17.10
Very Confident	13	2.50
Total	521	99%

Note: Five participants failed to give a confidence rating

Table 8.3.2 provides a breakdown of the confidence ratings for the participants in each condition, and Table 8.3.3 provides the mean scores for each condition. It is the aim of this section to compare the confidence ratings of the participants in each of the respondent conditions. Previous research suggests (Garrioch & Brimacombe, 2001) that those in the informed condition should be more confident in their responses. It is argued that the administrator may exhibit subtle cues indicating that they are aware of the identity of the suspect, thereby increasing the confidence of the witness that they have correctly identified the suspect.

Table 8.3.2: Confidence ratings for each condition as a percentage (frequencies in parentheses)

	Confidence ratings				
Respondent condition	Very Unconfident	Unconfident	Neither Confident or Unconfident	Confident	Very Confident
A1	25% (25)	22% (22)	30% (30)	22% (22)	1% (1)
A2	14.6% (15)	28.2% (29)	35.9% (37)	17.5% (18)	1.9% (2)
B1	45.5% (51)	24.1% (27)	14.3% (16)	9.8% (11)	3.6% (4)
B2	32.7% (33)	34.7% (35)	18.8% (19)	9.9% (10)	4% (4)
C1	23.6% (26)	24.5% (27)	23.6% (26)	26.4% (29)	1.8% (2)

From Table 8.3.3 we can see that the mean confidence ratings are very low for each of the five conditions, with all of the means below the median point for the scale. Conditions A2 and B1 reported the highest and lowest mean scores respectively.

Table 8.3.3: Mean confidence ratings for each condition

Respondent condition	Mean confidence rating	Standard deviation
A1	2.52	1.12
A2	2.63	1.01
B1	1.99	1.17
B2	2.18	1.12
C1	2.58	1.17

An independent samples t-test was conducted to compare the confidence ratings for the informed conditions (groups A1 and B1), and the uninformed conditions (groups A2, B2, and C1). There was a significant difference in the scores for the informed group (M = 2.24, SD = 1.17) and the uninformed group (M = 2.47, SD = 1.12), t (519) = -2.20, p = .03 (two-tailed). The magnitude of the differences in the means (mean difference = -.22, 95% CI: -.42 to -.02) was very small (eta squared = .009). Although there was a significant difference in the mean scores for the informed and uniformed groups, the results were not in the expected direction. The uninformed group reported higher confidence ratings than the informed group which is contrary to the result expected from previous research (Garrioch & Brimacombe, 2001).

8.4. The homogeneity of the sample

One of the most important differences between the participants in this experiment is the distinction between those participants whose line-up Administrators were informed of the location of the suspect in the line-up, and those whose Administrators were uninformed. Analysis was conducted to determine whether there are any significant differences between the participants in the informed condition and the participants in the uninformed condition. As we have seen above the uninformed group were significantly more confident in their identifications than the informed group. The variables in the analysis in this section were gender, age, occupation, and educational attainment. The first analysis conducted examined whether there was any difference in the number of male and female participants in the informed and uninformed conditions. Table 8.4.1 shows the number of male and female participants in the informed and uninformed conditions.

Table 8.4.1: The percentage of male and female participants in the informed and uninformed conditions (frequencies in parentheses)

	Male	Female	Total
Informed	44.8% (95)	55.2% (117)	100% (212)
Uninformed	48.2% (150)	51.8% (161)	100% (311)
Total	46.85% (245)	53.15% (278)	100% (523)

Note: three participants did not indicate their gender

A Chi-square test for independence (with Yates Continuity Correction) was conducted. It indicated no significant difference in the number of males and females in the informed and uninformed conditions, χ^2 (1, n = 523) = .46, p = .50, phi = -.03 indicating a very small effect size.

In order to determine whether there is an equal spread of ages in both the informed and the uninformed conditions, the ages of the participants were combined into eight groups. Table 8.4.2 shows the number of participants in each age group for each condition.

Table 8.4.2: The percentage of participants in each age group for the informed and uninformed conditions (frequencies in parentheses)

Age groups	Informed	Uninformed
16-19	28.3% (60)	12.4% (39)
20-29	30.2% (64)	49.4% (155)
30-39	7.5% (16)	12.7% (40)
40-49	16.5% (35)	10.5% (33)
50-59	13.2% (28)	9.2% (29)
60-69	3.3% (7)	4.1% (13)
70-79	0.9% (2)	1.3% (4)
80-89	~	0.3% (1)
Total	212	314

A Chi-square test for independence was conducted; however the minimum cell frequency assumption was violated. The normality testing of the age data was conducted, it indicated that the participant aged 88 years was an outlying variable. Therefore, three age groupings were constituted, those aged 16-29 were grouped as young, those aged 30-49 were grouped as middle, and those aged 50-88 were grouped as older. Table 8.4.3 shows the number of participants in each of the three new age groups for each condition.

Table 8.4.3: The percentage of participants in each of the age groupings for the informed and uninformed conditions (frequencies in parentheses)

Age groups	Informed	Uninformed
Young	58.50% (124)	61.80% (194)
Middle	24.10% (51)	23.20% (73)
Older	17.50% (37)	15.00% (47)
Total	212	314

The Chi-square test for independence was conducted on the three age groups, It indicated no significant difference between the age groups in the informed and uninformed conditions, χ^2 (2, n = 526) = .75, p = .69, Cramer's V = .04, indicating a small effect size.

Analysis was then conducted to determine whether there is an equal spread of occupational categories in both the informed and the uninformed conditions. Table 8.4.4 shows the percentage of participants in each occupational category for the informed and uninformed conditions.

Table 8.4.4: Percentage of participants in each occupational category for the informed and uninformed conditions (frequencies in parentheses)

Occupational Category	Informed	Uninformed
Managers/ Directors/ Senior Officials	3.3% (7)	4.1% (13)
Professionals	13.2% (28)	16.6% (52)
Associate Professionals and Technical	6.1% (13)	11.5% (36)
Admin and Secretarial	7.1% (15)	4.8% (15)
Skilled Trade	4.7% (10)	6.1% (19)
Caring/ Leisure and Other Services	1.9% (4)	5.1% (16)
Sales and Customer Service	5.2% (11)	9.6% (30)
Process/ Plant and Machine Operatives	0.9% (2)	0.3% (1)
Elementary Occupations	14.2% (30)	5.1% (16)
Student	31.6% (67)	29.9% (94)
Unemployed	3.3% (7)	1.9% (6)
Retired	2.4% (5)	4.5% (14)
Total	93.9% (199)	99.4% (312)

Note: 13 participants in the informed condition and 2 in the uninformed condition did not indicate their occupation

A Chi-square test for independence was conducted, the minimum cell frequency assumption was again violated, however according to Field (2005) a violation below 20% is acceptable for larger tables. In this instance the violation was 8.3%. The Chi-square test indicated a significant difference in the spread of occupational categories for the informed and uninformed conditions, χ^2 (11, n = 511) = 28.46, p = .003, Cramer's V = .24, indicating a small effect size.

Analysis was also conducted to determine whether there is an equal spread of educational attainment in both the informed and the uninformed conditions. Table 8.4.5 shows the percentage of participants in each educational attainment category for the informed and uninformed conditions.

Table 8.4.5: Percentage of participants in each educational category for the informed and uninformed conditions (frequencies in parentheses)

Educational attainment	Informed	Uninformed
No qualifications	~	1.6% (5)
GCSE/O-Level/NVQ	38.2% (81)	29% (91)
A-Level / BTEC	17.9% (38)	26.1% (82)
Degree / Diploma	22.2% (47)	35% (110)
Masters / PhD	13.7% (29)	4.8% (15)
Total	92% (212)	96.5% (303)

Note: 17 participants in the informed condition and 11 participants in the uninformed condition failed to record their educational attainment

A Chi-square test for independence was conducted; the minimum cell frequency assumption was again violated, as the violation was 20% Fishers exact test statistic was reported. The test indicated a significant difference in the spread of educational attainment for the informed and uninformed conditions, $p < .001$, Cramer's V = .24, indicating a small effect size.

In summary the informed and uninformed conditions do not differ significantly in their number of male and female participants, or in their spread of participant ages. Therefore, it cannot be said that differences in identification rates between the two groups are due to differing numbers of males and females or different age groups. However, the participants of the two groups did differ significantly in their occupations and educational attainment. Examination of the occupations and educational attainment of the two groups finds that most of the differences between the informed and uninformed groups are small. The largest difference in occupations is for Elementary workers who are more prevalent in the informed condition. Whilst for educational attainment the most substantial difference between the two groups is for those who obtained a Degree or Diploma who are more prevalent in the uninformed condition. Although the difference in occupations and educational attainment between the two groups is significant it is difficult to determine how these differences would affect the line-up identifications.

Analysis Section 2 – Examining the Extent of the Administrator Effects

Chapter 9:

This chapter will examine the line-up identification data obtained in order to determine whether administrators have affected the outcome of the line-ups they conducted. In condition A1 both the Facilitator and the administrator were informed that the target occupied position 8 in the line-up, whilst in B1 they were informed that the target occupied position 4 in the line-up. A1 and B1 therefore constitute the informed condition. If administrator effects are in evidence then it would be expected to find higher incidences of identifications of line-up member 8 in condition A1 and line-up member 4 in condition B1. In conditions A2 and B2 the Facilitators were informed that the target occupied position 8 and 4 respectively, although the administrators in these conditions were not informed. Conditions A2 and B2 therefore constitute the uninformed condition. If the knowledge of the location of the target has transferred from the Facilitator to the administrator, higher incidences of identifications of line-up members 8 and 4 would be expected, although to a lesser extent than in the informed condition. Finally, condition C1 constitutes the control group. In this condition both the Facilitator and the administrator were not informed of the location of the target in the line-up. In this condition then higher incidences of identifications of line-up member 8 and 4 would not be expected. Indeed in this condition we would expect to see an equal spread of identifications for all of the members of the line-up.

9.1. Line-up member selection frequency for the five conditions

From Table 9.1.1 a number of interesting results emerge. Firstly, in condition A1, 20% of the participants identified target number 8 from the line-up. The next most frequent identification was of target number 7 at 12%. Therefore, as expected, in the informed condition the target was identified more often than other members of the line-up. The second informed condition, B1, provides similar results. Target number 4 was identified by 14.30% of the participants in the condition, which although is lower than the 20% in condition A1 is still the joint highest frequency of identifications.

Table 9.1.1: Percentage of line-up member selections for the five conditions (frequencies in parentheses)

	Experimental condition				
Line-up member	A1	A2	B1	B2	C1
1	4.00% (4)	4.90% (5)	1.80% (2)	5.00% (5)	2.70% (3)
2	9.00% (9)	5.80% (6)	8.00% (9)	5.00% (5)	6.40% (7)
3	10.00% (10)	9.70% (10)	14.30% (16)	9.90% (10)	8.20% (9)
4	11.00% (11)	11.70% (12)	14.30% (16)	11.90% (12)	15.50% (17)
5	8.00% (8)	8.70% (9)	5.40% (6)	8.90% (9)	12.70% (14)
6	6.00% (6)	4.90% (5)	8.00% (9)	8.90% (9)	10.00% (11)
7	12.00% (12)	21.40% (22)	10.70% (12)	9.90% (10)	7.30% (8)
8	20.00% (20)	14.60% (15)	5.40% (6)	5.00% (5)	8.20% (9)
9	7.00% (7)	4.90% (5)	7.10% (8)	9.90% (10)	9.10% (10)
10	3.00% (3)	3.90% (4)	5.40% (6)	5.00% (5)	~
11	4.00% (4)	3.90% (4)	3.60% (4)	7.90% (8)	6.40% (7)
12	6.00% (6)	5.80% (6)	11.60% (13)	12.90% (13)	12.70% (14)
	100% (100)	100% (103)	95.60% (107)	100% (101)	99.10% (109)

For the uninformed conditions (A2 and B2) there are again high incidences of identification for the two targets (8 and 4). In condition A2 14.60% of participants identified target number 8, whilst in condition B2 11.90% identified target number 4. These percentages represent the second most frequent identifications in these conditions with slightly more participants identifying number 7 in condition A2 and number 12 in condition B2. For the control condition, C1, there is, as expected a good spread of the frequencies of identifications for all line-up members, apart from line-up member number 10 who was not identified by any of the participants in this condition. The highest frequency of identification (15.50%) is for number 4, although frequencies of identification for number 5 and number 12 were also high (12.70%).

Table 9.1.2: The percentage of identification responses (frequencies in parentheses) for the two targets and the line-up fillers for all conditions

Administrator Condition	Line-up Member No 8 Selected	Line-up Member No 4 Selected	Other Line-up Member Selected
Condition A1	20% (20)	11% (11)	69% (69)
Condition A2	14.6% (15)	11.7% (12)	73.7% (76)
Condition B1	5.4% (6)	14.3% (16)	80.3% (85)
Condition B2	5% (5)	11.9% (12)	83.1% (84)
Condition C1	8.2% (9)	15.5% (17)	76.3% (84)
Total	10.56% (55)	13.05% (68)	76.39% (398)

To summarise, from a general examination of the above data there does appear to be an effect of administrator knowledge on the frequency of identifications obtained. The effect appears most strongly in condition A1, and is more diluted or subtle in condition B1. As we would expect the effect is less discernible in the uninformed conditions (A2 and B2), although identifications of the targets are still the second most frequent identifications.

In this novel experiment participants were not the standard eyewitnesses used in the majority of memory and eyewitness identification research. For these line-ups the participants had not viewed a mock crime or a staged event, they therefore had no memory of the target. In the control condition both the Facilitator and the administrator were unaware of the location of the suspect. In the absence of administrator effects we would expect that each member of the line-up in the control condition would have an equal chance of being identified. Table 9.1.3. shows the comparison of observed identifications against identifications expected by chance for each line-up member in the control condition.

Table 9.1.3: Observed and expected frequencies of line-up member identification in condition C1

Line-up Member No.	Observed N	Expected N	Residual
1	3	9.17	-6.17
2	7	9.17	-2.17
3	9	9.17	-0.17
4	17	9.17	7.83
5	14	9.17	4.83
6	11	9.17	1.83
7	8	9.17	-1.17
8	9	9.17	-0.17
9	10	9.17	0.83
10	7	9.17	-2.17
11	14	9.17	4.83
12	1	9.17	-8.17
Total	110		

χ^2 (11, N = 110) = 24.84, $p < .05$

A series of Chi-square goodness of fit tests were conducted to examine whether the observed identifications in the two informed conditions (A1 and B1) and the two uninformed conditions (A2 and B2) differed significantly from the identifications observed in the control condition (C1). Tables 9.1.4 to 9.1.7 show the comparison of observed identifications against identifications observed in the control condition (adjusted for the number of identifications) for each line-up member in the two informed conditions and the two uniformed conditions, with the chi-square results at the foot of each table.

Table 9.1.4: Observed and expected frequencies of line-up member identification in condition A1

Line-up Member No.	Observed N	Expected N	Residual
1	4	2.17	1.83
2	9	6.17	2.83
3	10	8.17	1.83
4	11	16.17	-5.17
5	8	13.17	-5.17
6	6	10.17	-4.17
7	12	7.17	4.83
8	20	8.17	11.83
9	7	9.17	-2.17
10	3	6.17	-3.17
11	4	13.17	-9.17
12	6	0.17	5.83
Total	100		

χ^2 (11, N = 100) = 237.58, $p < .001$

Table 9.1.5: Observed and expected frequencies of line-up member identification in condition A2

Line-up Member No.	Observed N	Expected N	Residual
1	5	2.42	2.58
2	6	6.42	-0.42
3	10	8.42	1.58
4	12	16.42	-4.42
5	9	13.42	-4.42
6	5	10.42	-5.42
7	22	7.42	14.58
8	15	8.42	6.58
9	5	9.42	-4.42
10	4	6.42	-2.42
11	4	13.42	-9.42
12	6	0.42	5.58
Total	103		

χ^2 (11, N = 103) = 126.11, $p < .001$

Table 9.1.6: Observed and expected frequencies of line-up member identification in condition B1

Line-up Member No.	Observed N	Expected N	Residual
1	2	2.75	-0.75
2	9	6.75	2.25
3	16	8.75	7.25
4	16	16.75	-0.75
5	6	13.75	-7.75
6	9	10.75	-1.75
7	12	7.75	4.25
8	6	8.75	-2.75
9	8	9.75	-1.75
10	6	6.75	-0.75
11	4	13.75	-9.75
12	13	0.75	12.25
Total	107		

Five participants in this condition failed to make an identification

χ^2 (11, N = 107) = 222.24, p < .001

Table 9.1.7: Observed and expected frequencies of line-up member identification in condition B2

Line-up Member No.	Observed N	Expected N	Residual
1	5	2.25	2.75
2	5	6.25	-1.25
3	10	8.25	1.75
4	12	16.25	-4.25
5	9	13.25	-4.25
6	9	10.25	-1.25
7	10	7.25	2.75
8	5	8.25	-3.25
9	10	9.25	0.75
10	5	6.25	-1.25
11	8	13.25	-5.25
12	13	0.25	12.75
Total	101		

χ^2 (11, N = 101) = 661.57, p < .001

The chi-square tests show that there is a significant difference in the number of observed identifications in the two informed conditions and the two uninformed conditions, and the identifications observed in the control condition. To further explore the significant differences the observed number of target and non-target identifications in conditions A1, A2, B1, and B2 are compared to the observed number of target identifications in the control condition. To reiterate a target identification in conditions A1 and A2 is the identification of number 8, whilst in B1 and B2 it is the identification of number 4. This comparison can be seen in Table 9.1.8.

Table 9.1.8: Comparison of observed frequencies of target and non-target line-up member identification for conditions A1, A2, B1, and B2, with observed frequencies from the control condition

		Observed N	Control condition N	Residual
A1	Yes	20	8.17	11.83
	No	80	91.83	-11.83
A2	Yes	15	8.42	6.58
	No	88	94.58	-6.58
B1	Yes	16	16.75	-0.75
	No	91	90.25	0.75
B2	Yes	12	16.25	-4.25
	No	89	84.75	4.25

Four chi-square goodness of fit tests were conducted. In order to control for a Type I error a Bonferroni correction was applied, resulting in an alpha level of .01. The chi-square goodness of fit tests indicate that there is a significant difference in the number of participants identifying target number 8 in condition A1 (χ^2 (1, n = 100) = 18.65, $p < .001$) than the frequency of identifications of number 8 in the control condition. However there were no significant differences in the number of participants identifying target number 8 in condition A2 (χ^2 (1, n = 103) = 5.60, $p = .02$) or the number identifying target number 4 in conditions B1 (χ^2 (1, n = 107) = .04, $p > .05$) and B2 (χ^2 (1, n = 101) = 1.33, $p > .05$) than the frequency of identifications of number 8 and number 4 in the control condition.

9.2. Line-up member selection frequency for type of line-up

Table 9.2.1 shows the percentage of identifications of each member of the line-up for both simultaneous and sequential line-ups. It is clear from this table that marginally more simultaneous line-ups than sequential line-ups were conducted. Further examination of the data in this table shows that the percentage of identifications for each of the line-up members is comparable for both types of line-up. The only difference is for line-up member 12 who was identified more often from a simultaneous line-up. For the two targets of the line-up, members 4 and 8, the table shows that these two line-up members were identified from the simultaneous line-up as often they were identified from the sequential line-up.

Table 9.2.1: Percentage of line-up member selections by type of line-up (frequencies in parentheses)

	Type of line-up	
Line-up member	Simultaneous	Sequential
1	2.9% (8)	4.4% (11)
2	6.5% (18)	7.2% (18)
3	9.8% (27)	11.2% (28)
4	13.4% (37)	12.4% (31)
5	8.3% (23)	9.2% (23)
6	7.2% (20)	8% (20)
7	11.6% (32)	12.8% (32)
8	10.9% (30)	10% (25)
9	7.6% (21)	7.6% (19)
10	2.9% (8)	4% (10)
11	5.4% (15)	4.8% (12)
12	12.3% (34)	7.2% (18)
Total	98.9% (273)	98.8% (247)

Note: For 6 participants type of line-up was not indicated

9.3. Analysis of the participants who made target identifications in the informed group

In this study conditions A1 and B1 constitute the informed group, with conditions A2, B2, forming the uninformed group and condition C1 forming the control group. In condition A1 those who selected number eight were judged as making a target identification, whereas in condition B1 those who selected number four were judged as making a target identification.

Out of 212 participants in the informed condition, 36 (16.98%) individuals "identified" the target, below the particulars of those thirty six participants are analysed.

The 36 participants were comprised of 17 (47.22%) male participants and 19 (52.78%) female participants. Their ages ranged from 17 to 75 years with a mean age of 31.11 years (SD = 16.82). This group of participants is therefore very homogenous in terms of age and gender with the whole sample. Table 9.3.1 shows the percentage of male and female participants who correctly identified the target split between conditions A1 and B1.

Table 9.3.1: Percentage of male and female participants making target identifications in the informed condition (frequencies in parentheses)

Condition	Male	Female
A1	27.78% (10)	27.78% (10)
B1	19.44% (7)	25.00% (9)
Total	47.22% (17)	52.78% (19)

Table 9.3.2 shows the age group of the participants making a target identification in the informed condition, split between conditions A1 and B1. We can see from this table that the majority (two thirds) of the participants who made a target identification were aged between 16 and 29. The remaining third of the participants were aged 40 to 79.

Table 9.3.2: Percentage of participants making target identifications for each age group in the informed condition (frequencies in parentheses)

Condition	16-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89
A1	30.56% (11)	13.89% (5)	~	5.56% (2)	5.56% (2)	~	~	~
B1	2.78% (1)	19.44% (7)	~	8.33% (3)	8.33% (3)	2.78% (1)	2.78% (1)	~
Total	33.33% (12)	33.33% (12)	~	13.89% (5)	13.89% (5)	2.78% (1)	2.78% (1)	~

Table 9.3.3 shows the percentage of target identifications in the informed condition for each type of line-up. There was no difference in the type of line-up on the number of target identifications in the informed condition. Exactly half of the 36 participants identified the

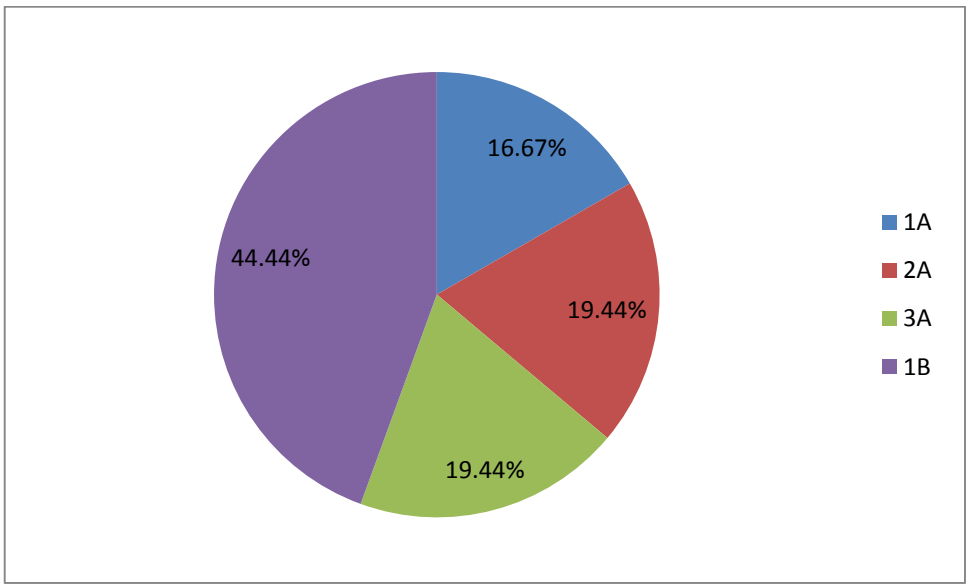
target from a simultaneous line-up. This indicates that participants were just as likely to be influenced in a simultaneous line-up as they were in a sequential line-up; therefore, the superiority of the sequential line-up has not been supported in this instance.

Table 9.3.3: Percentage of target identifications by type of line-up in the informed condition (frequencies in parentheses)

Condition	Simultaneous	Sequential
A1	25.00% (9)	30.56% (11)
B1	25.00% (9)	19.44% (7)
Total	50.00% (18)	50.00% (18)

For the informed condition there were four Facilitators; three Facilitators for condition A1 (1A, 2A, and 3A) and one for condition B1 (1B). Facilitator 1A and 1B are female, facilitators 2A and 3A are male. Figure 9.3.a shows the percentage of target identifications in the informed condition that each facilitator is responsible for. For condition B1 there was only one facilitator, therefore they were responsible for all of the target identifications in condition B1. For condition A1 we can see that Facilitator 1A was responsible for 16.67% of the target identifications, whilst Facilitators 2A and 3A were responsible for slightly more target identifications, 19.44% each.

Figure 9.3.a: Percentage of correct identifications by Facilitator in the informed condition



In conditions A1 and B1 there were six administrators; two females and one male in each condition. Figure 9.3.b shows the percentage of correct identifications for each administrator in the informed condition. We can see that Administrator 8 was responsible for a large proportion of the correct identifications, however, Administrators 9 and 10 were responsible for only a small proportion of the correct identifications. We can also see from Figure 9.3.c that female administrators were responsible for more correct identifications than male administrators.

Figure 9.3.b: Percentage of target identifications by Administrator in the informed condition

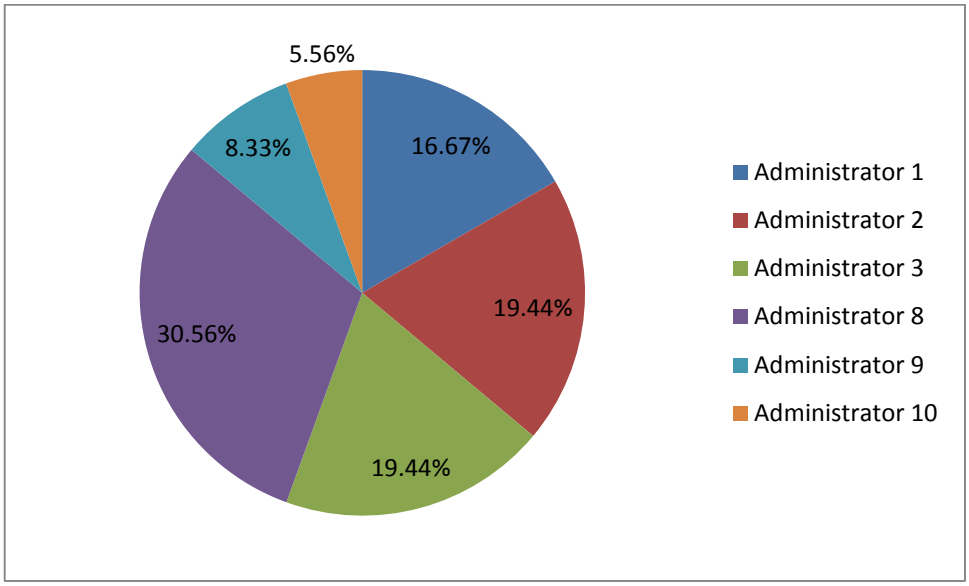
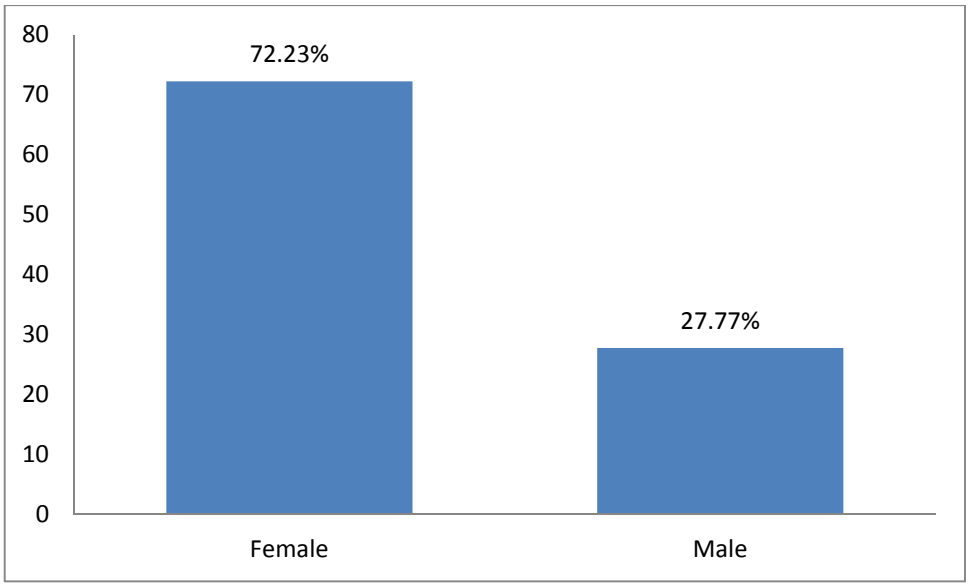
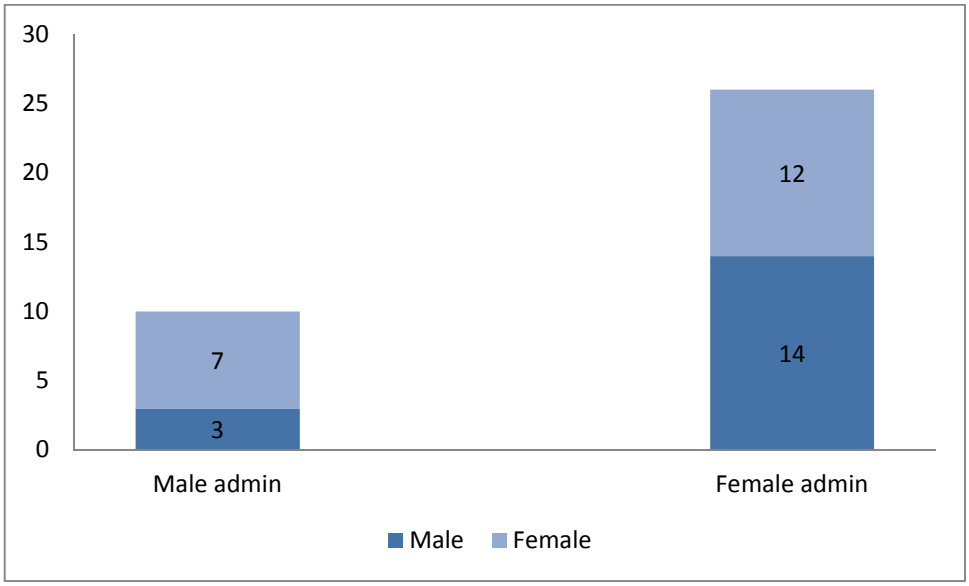


Figure 9.3.c: Percentage of target identifications by male and female Administrators in the informed condition



In line with previous research we have found that the female participants in the informed condition were more susceptible to influence than the male participants. Surprisingly though Figure 9.3.d shows that female administrators were more influential with male participants.

Figure 9.3.d: Number of target identifications made by male and female participants in the informed condition by the gender of the Administrator



9.4. Analysis of the participants who made target identifications in the uninformed group

Conditions A2 and B2 form the uninformed group, in condition A2 those who selected number eight were judged as making a target identification, whereas in condition B2 those who selected number four were judged as making a target identification. Of the 204 participants in the uninformed condition 27 (13.24%) correctly "identified" the target, below the particulars of those twenty seven participants are analysed.

The 27 participants were comprised of 10 (37.04%) male participants and 16 (59.26%) female participants. Their ages ranged from 16 to 63 years with a mean age of 29.15 years (SD = 13.60). This group of participants is fairly homogenous to the whole sample, although female participants are marginally overrepresented in this group. The range of ages is also 25 years lower than the whole sample, and the mean age is marginally lower. Table 9.4.1 shows the percentage of male and female participants who identified the target split between conditions A2 and B2.

Table 9.4.1: Percentage of male and female participants making target identifications in the uninformed condition (frequencies in parentheses)

Condition	Male	Female
A2	23.08% (6)	34.62% (9)
B2	15.38% (4)	26.92% (7)
Total	38.46% (10)	61.54% (16)

Note: one participant who made a target identification in B2 did not indicate their gender

Table 9.4.2 shows the age group of the participants making a target identification in the uninformed condition, split between conditions A2 and B2. We can see from this table that the majority (over 90%) of the participants who made a target identification were aged between 16 and 49. Only two participants were older than this, and were aged between 60 and 69.

Table 9.4.2: Percentage of participants making target identifications for each age group in the uninformed condition (frequencies in parentheses)

Condition	16-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89
A2	14.81% (4)	25.93% (7)	11.11% (3)	3.70% (1)	~	~	~	~
B2	7.41% (2)	18.52% (5)	3.70% (1)	7.41% (2)	~	7.41% (2)	~	~
Total	22.22% (6)	44.44% (12)	14.81% (4)	11.11% (3)	~	7.41% (2)		

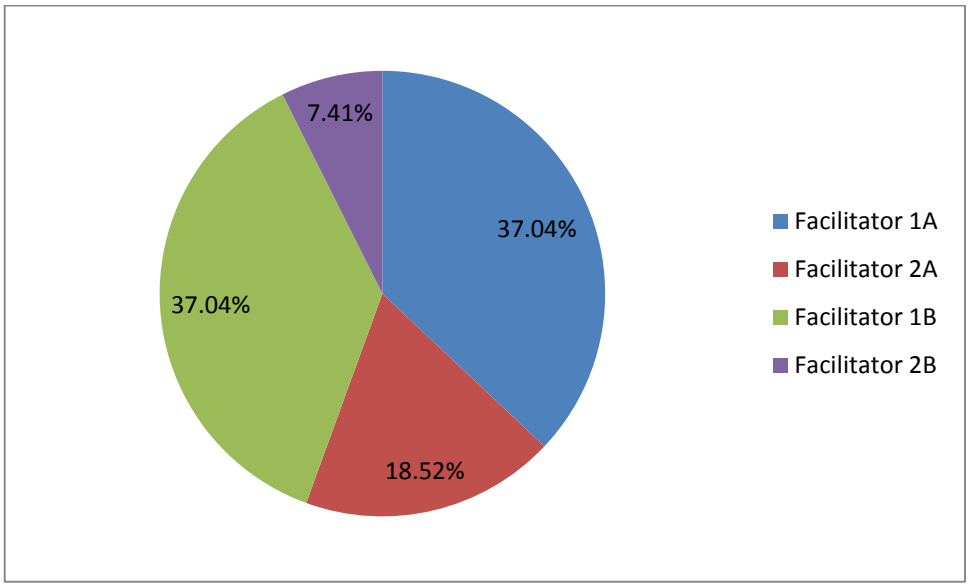
Table 9.4.3 shows the percentage of target identifications in the uninformed condition for each type of line-up. For the uninformed condition there was a slight difference in the type of line-up on the number of target identifications. Just over half of the 27 participants identified the target from a sequential line-up. This indicates that participants were slightly more likely to be influenced in a sequential line-up than they were in a simultaneous line-up; therefore, the superiority of the sequential line-up has again not been supported in this instance.

Table 9.4.3: Percentage of target identifications by type of line-up in the uninformed condition (frequencies in parentheses)

Condition	Simultaneous	Sequential
A2	29.63% (8)	25.93% (7)
B2	14.81% (4)	29.63% (8)
Total	44.44% (12)	55.56% (15)

For the uninformed condition there were four Facilitators; two Facilitators for condition A2 (1A, and 2A) and two for condition B2 (1B and 2B). Facilitators 1A, 1B, and 2B are female, facilitator 2A is male. Figure 9.4.a shows the percentage of target identifications in the uninformed condition that each facilitator is responsible for. For conditions A2 and B2 we can see that Facilitators 1A and 1B were responsible for over a third (37.04%) of the target identifications each. Facilitators 2A was responsible for nearly 20% of the identifications, however, Facilitator 2B was only responsible for 7.41%.

Figure 9.4.a: Percentage of target identifications by Facilitator in the uninformed condition



In conditions A2 and B2 there were seven administrators; two females and two males in condition A2, and three females in condition B2. Figure 9.4.b shows the percentage of target identifications for each administrator in the uninformed condition. We can see that administrators 4, 5, 7, 11, and 12 were responsible for a relatively equal proportion of the

target identifications. Administrator 6 was not responsible for any of the target identifications, whilst administrator 13 was responsible for only a small proportion of the target identifications. Figure 9.4.c shows that the female administrators were responsible for the vast majority of target identifications in the uninformed condition.

Figure 9.4.b: Percentage of target identifications by Administrator in the uninformed condition

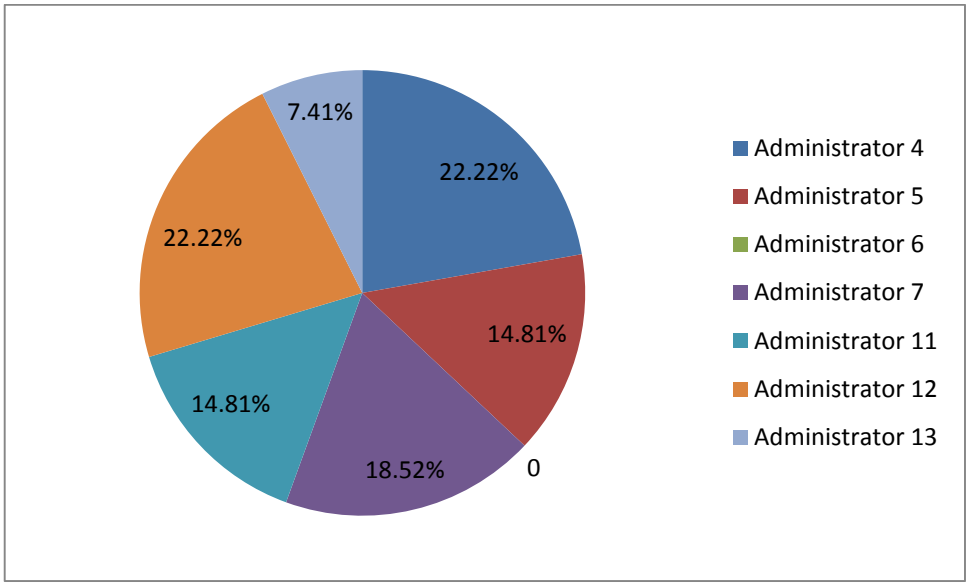
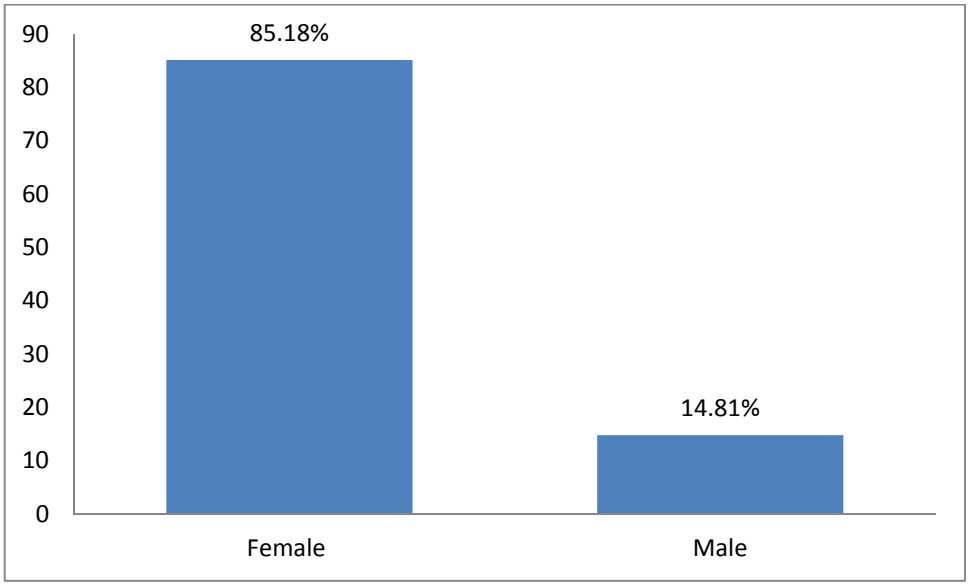
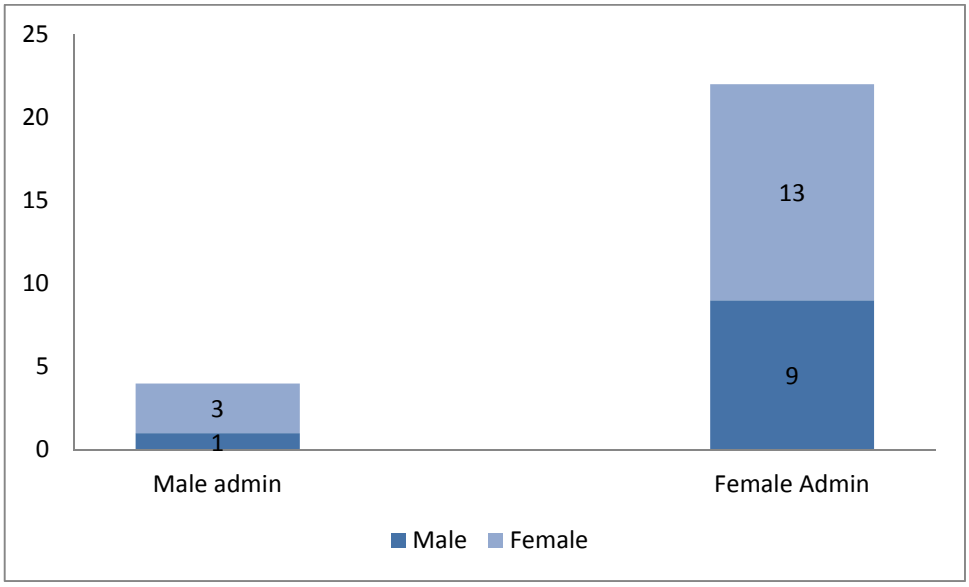


Figure 9.4.c: Percentage of target identifications by male and female Administrators in the uninformed condition



As with the informed condition, female participants were more susceptible to influence by a male administrator than male participants. However, contrary to the informed condition and in-line with previous research, female administrators in the uninformed condition were more influential with female participants.

Figure 9.4.d: Number of target identifications made by male and female participants in the uninformed condition by the gender of the Administrator



9.5. Target absent line-up analysis

As detailed in Chapter 7 there were two conditions in the study. Administrators were asked to conduct Condition 1 with half of their participants, whereby they conducted one line-up, either simultaneous or sequential, with their participants. In Condition 2 the same procedure as Condition 1 was followed, however, after completing the FIRO, participants in Condition 2 were asked to make a second identification, this time from a target absent line-up. The sub-sample of Condition 2 consists of 266 participants, 122 (45.9%) males and 141 (53%) females (3 participants failed to record their gender). The participants in Condition 2 range in age from 16 years to 88 years. The mean age of the sample is 31.42 years (Standard Deviation = 14.61 years), with a median age of 24.50 years.

The aim of Condition 2 was to examine whether the identifications in the first condition were due to the expectancy effect of the experimenter or were confounded by the presence of Mr Megrahi's photo in the line-up. The frequencies of identifications in Condition 2,

which can be seen in Table 9.5.1, indicate that there is indeed evidence for expectancy effects in Condition 2. Although the effect is not as strong as in the first condition, 16% of participants in A1 identify target number 8 as the Lockerbie bomber, whilst 12% of participants identified number 8 in condition A2, which is the second highest frequency in that group. The effect is stronger for condition B1 as 21.05% of the participants in this condition identified target number 4 from the line-up. This effect continues into condition B2 where the highest proportion of participants (14%) also identified target number 4.

Table 9.5.1: Percentage of line-up member selections in Condition 2 for each condition (frequencies in parentheses)

Line-up Member No.	A1	A2	B1	B2	C1	Total
1	4% (2)	8% (4)	3.51% (2)	8% (4)	3.7% (2)	5.36% (14)
2	10% (5)	10% (5)	8.77% (5)	8% (4)	1.9% (1)	7.66% (20)
3	12% (6)	8% (4)	10.53% (6)	8% (4)	13% (7)	10.34% (27)
4	14% (7)	4% (2)	21.05%(12)	14% (7)	5.6% (3)	11.88% (31)
5	10% (5)	18% (9)	7.02% (4)	10% (5)	9.3% (5)	10.73% (28)
6	~	6% (3)	7.02% (4)	6% (3)	5.6% (3)	4.98% (13)
7	8% (4)	6% (3)	1.75% (1)	6% (3)	7.4% (4)	5.75% (15)
8	16% (8)	12% (6)	~	8% (4)	5.6% (3)	8.05% (21)
9	8% (4)	8% (4)	5.26% (3)	8% (4)	18.5% (10)	9.58% (25)
10	8% (4)	12% (6)	14.04% (8)	8% (4)	13% (7)	11.11% (29)
11	8% (4)	8% (4)	15.79% (9)	8% (4)	7.4% (4)	9.58% (25)
12	2% (1)	~	5.26% (3)	8% (4)	9.3% (5)	4.98% (13)
Total	100% (50)	100% (50)	100% (57)	100% (50)	100% (54)	100% (261)

Note: five participants from B1 failed to make a selection

Table 9.5.2 shows the percentage of identifications of the target in each group of the informed and uninformed conditions. Twenty participants identified 4 or 8 in the informed condition, whilst thirteen participants identified 4 or 8 in the uninformed condition.

Table 9.5.2: Percentage of target identifications in the informed and uninformed conditions in Condition 2 (frequencies in parentheses)

Condition	Informed	Uninformed
A1	40% (8)	~
B1	60% (12)	~
A2	~	46.15% (6)
B2	~	53.85% (7)
Total	100% (20)	100% (13)

It is possible however, that participants in Condition 2 simply chose the same number as in Condition 1. There were 37 participants from the informed and uninformed groups who identified either target number 4 or 8 from the line-up, then took part in the target-absent line-up in Condition 2. Of those 37 participants, 5 (13.51%) made the same identification in both conditions (one from A1, A2, and B2, and two from B1). Those 5 participants may not have remembered the number they chose from the first line-up, but even if those participants are removed, there is still evidence of expectancy effects in Condition 2. A Chi-square test for independence (with Yates Continuity Correction) indicated no significant association between correct or incorrect identifications in the first line-up and correct or incorrect identifications in the second line-up, χ^2 (1, n = 212) = .02, p > .05, phi = -.03.

9.6. Summary of the main findings of this chapter

When the administrator of the line-up was informed of the location of the target, participants selected those targets with greater frequencies than participants whose administrator was not informed of the location of the target. When the Facilitator was informed of the location of the target, but the administrator was not informed, participant selections of the targets were the second most frequent selections. These target identifications were also more frequent than the target identifications when the administrator was uninformed of the location of the targets. In condition A1, when the administrator was informed of the location of the target, the frequencies of target selection were significantly greater than the frequencies of target selection in the control condition.

Condition 2 of the study provides further evidence that the knowledge of the location of the target affects the frequency of the selection of the target. When the administrator was

informed of the location of the target, the target was the most frequent selection in that condition. Furthermore, the target was selected more frequently when the administrator was informed of the location than when the administrator was not informed. Participants also selected the target more frequently in the uninformed condition of Condition 2, where the Facilitator but not the administrator was informed of the location of the target, than participants in the control condition.

Analysis Section 3 – Analysis of the FIRO-B Questionnaire

Chapter 10:

10.1. Analysis of the structure of the FIRO-B questionnaire for this sample

As discussed in Chapter 7, the internal structure of the FIRO-B questionnaire has been queried. In order to determine the structure of the FIRO-B and how to analyse the data in this instance, Principal Components Analysis (PCA) was conducted. Prior to conducting PCA the suitability of the data for this procedure was assessed. Firstly inspection of the correlation matrix revealed many coefficients of .3 and above. Secondly the Kaiser-Meyer-Olkin value was .92, which according to Kaiser (1974) is a superb value, and which highly exceeds the recommended value of .6. Finally, Bartlett's Test of Sphericity (Bartlett, 1954) reached statistical significance ($p < .001$); therefore, the factorability of the correlation matrix was supported.

The Principal Components Analysis indicated the presence of nine components with eigenvalues exceeding 1. Table 10.1.1 shows the variance explained by each component.

Table 10.1.1: The variance explained by each component

Component	Variance explained %	Cumulative %
1	25.47	25.47
2	9.71	35.18
3	8.30	43.47
4	5.49	48.96
5	4.43	53.40
6	4.25	57.65
7	2.52	60.17
8	2.30	62.47
9	2.19	64.66

However, an inspection of the screeplot, which can be viewed in Appendix 8, revealed a clear break after the sixth component. Using Catell's (1966) scree test, it was therefore decided to retain six components for further analysis. It has been argued that both Kaiser's eigenvalue criterion and Catell's scree test overestimate the number of components to extract (Zwick & Velicer, 1986). Therefore, Horn's (1965) parallel analysis was also conducted, which compares the eigenvalues of the data with eigenvalues from a data set of

the same size which is randomly generated. The Monte Carlo PCA for parallel analysis statistical programme (Watkins, 2000) was utilised. The programme will generate a specified number, in this case 100, random data sets of the same size, the average eigenvalues of the 100 randomly generated samples are then calculated. The average eigenvalues of the randomly generated data are then compared to the actual data; if the actual eigenvalue is larger than the eigenvalue from the randomly generated data then the factor is retained, if it is smaller it is rejected. As Table 10.1.2 shows, only the eigenvalues from the PCA for the first six factors were larger than the criterion values from the parallel analysis, therefore the parallel analysis supported the decision to extract six components.

Table 10.1.2: Comparison of eigenvalues from PCA and criterion values from parallel analysis

Component	Actual eigenvalue from PCA	Criterion value from parallel analysis
1	13.75	1.69
2	5.24	1.63
3	4.48	1.58
4	2.96	1.54
5	2.39	1.50
6	2.30	1.47
7	1.36	1.43
8	1.24	1.40
9	1.19	1.37

The six component solution explained a total of 57.65% of the variance (Table 10.1.1.). In order to aid the interpretation of the components the data was rotated. As there is a theoretical reason to assume that interpersonal relationship factors are related, it is also assumed that the factors of the FIRO-B are related, therefore oblique rotation, specifically oblimin rotation was conducted. Appendix 9 shows the pattern and structure matrix for the PCA (with major loadings for each item in bold). As can be seen from this table the rotated solution does not produce a simple structure, with many variables loading substantially onto more than one component. Although the structure is not simple there is a discernible pattern that has emerged. Firstly we can see that 16 items have loaded onto the first component, these 16 items include all of the received inclusion items and seven of the expressed inclusion items. The second component is made up exclusively of the expressed

control items, whilst component three is made up of only received control items. Component four contains only three items, one expressed openness item, and two expressed inclusion items. The fifth component contains four expressed openness items, whilst component six contains the remaining four expressed openness items and all of the received openness items. We can see from the correlation matrix (Table 10.1.3.) that there is a very weak correlation between all of the factors apart from the relationship between factor 1 and factor 6 which demonstrates a moderate positive correlation (Cohen, 1988).

Table 10.1.3: Component correlation matrix for the six factor solution

Component	1	2	3	4	5	6
1	1.000	.183	-.005	-.051	.083	-.453
2	.183	1.000	.015	-.019	-.094	-.047
3	-.005	.015	1.000	-.034	-.109	.036
4	-.051	-.019	-.034	1.000	-.013	.049
5	.083	-.094	-.109	-.013	1.000	-.094
6	-.453	-.047	.036	.049	-.094	1.000

The above analysis shares some commonality with the work conducted by IRCIP (2010) on the structure of the FIRO-B questionnaire detailed in Chapter 7. As in that example the *Mode* facet (expressed / received) is not distinguishable for Inclusion or Openness. In the previous analysis expressed and received items for both Inclusion and Openness were grouped together whilst in this case expressed and received Inclusion items are loaded onto factor 1 and expressed and received Openness items are loaded onto factor 6. However, unlike the previous analysis, the PCA revealed two distinct and separate factors of Control, factor 2; expressed Control and factor 3; received Control. This indicates that the *Mode* facet is having an effect; participants are distinguishing between received and expressed items, but only in relation to Control items. Factors 4 and 5 pose a challenge to the analysis as they only contain three and four items respectively.

In order to further explore the structure of the FIRO-B, the PCA was repeated, the same procedure as above was used, however, this time only three factors were extracted. The pattern and structure matrix for this analysis can be seen in Appendix 10 (with major loadings for each item in bold).

The results of the repeated analysis confirmed the research conducted by Mahoney and Stasson (2005) and Macrosson (2000) detailed in Chapter 7. In this analysis factor 1 now encompasses all of the received Inclusion items, all of the received Openness items, all of the expressed Inclusion items apart from item 43, and all of the expressed Openness items apart from items 33, 45, and 27. Therefore, factor 1 correlates with Mahoney and Stasson's (2005) Socio-Emotional Affect and Macrosson's (2000) Nurturance. Factor 2 in this analysis again contains all of the expressed Control items, plus items 33 and 45 from expressed Openness. Further analysis of these two items allows some explanation of why they are included in this factor. The items are "There are some things I would not tell anyone" and "There is a part of myself I keep private." These two items could be construed not as a person's desire to be open with others, but rather a desire to control what they inform others about themselves. The third factor contains all of the received Control items plus item 27 from expressed Openness and item 43 from expressed Inclusion. Item 27 "There are some things I do not tell anyone" can be explained in the same terms as items 33 and 45, it represents a desire to control how much personal information other people know. However, the presence of item 43 "I look for people to be with," in factor 3 cannot be so easily explained. In the six factor solution there was a moderate positive correlation between factor 1 and factor 6. Now, as factor 6 has been subsumed into factor 1, we can see from the correlation matrix (Table 10.1.4) that there are very small correlations between all of the factors, indicating that they are indeed measuring distinct aspects of interpersonal behaviour.

Table 10.1.4: Component correlation matrix for the three factor solution

Component	1	2	3
1	1.000	.073	-.063
2	.073	1.000	.120
3	-.063	.120	1.000

The three factor solution therefore corresponds with previous research on the FIRO-B, whereby Inclusion and Openness both combine to form one facet of interpersonal relationships, which can be termed Personal and Social Inclusion, with Control (with a received and expressed mode) constituting the other facet.

In this instance oblique rotation (Direct Oblimin) was used as there are theoretical grounds to suppose that the factors of interpersonal relations should be related to each other. Indeed, it has been argued (Field, 2005) that with psychological constructs and human participants, orthogonal rotation, which assumes the factors to be independent, should not be used. It has further been argued that for exploratory factor analysis oblique rotation should be used before orthogonal rotation as the former provides details as to the extent of the correlation between factors (Pallant, 2007). For the current data-set there was some correlation between the factors, albeit small. Therefore, in the interests of completeness, the PCA was repeated with orthogonal rotation (Varimax). As with oblique rotation, without extracting any factors, nine components reported Eigenvalues above 1. The PCA was repeated with orthogonal rotation extracting both six and three factors. Analysis of the rotated component matrix found that the items loaded onto the same six factors and the same three factors as in the oblique rotation.

In Chapter 7 the facet framework of the FIRO-B was discussed, therefore the FIRO-B data was subjected to Smallest Space Analysis (SSA). The 2-dimensional SSA solution has a Guttman – Lingoes coefficient of alienation of .18615 in 8 iterations, showing a reasonable fit between the Pearson's coefficients of the FIRO-B items and their corresponding geometric distances in the configuration. The two-dimensional solution was adopted as it was found to have a satisfactory coefficient of alienation and was considered to describe the pattern of relationships better that the three-dimensional solution. Figure 10.1.a shows the projection of vector 1 by vector 2 of the two dimensional space. The numbers in the visual representation correspond to the items of the FIRO-B scale which can be seen in Table 10.1.7. Items 6, 10, 22, and 48 belong to the Personal and Social Inclusion facet. They are not identifiable on the plot because their coordinates place them in the cluster of Personal and Social Inclusion items to the left of the plot.

Figure 10.1.a: Visual representation of the SSA for the FIRO-B data

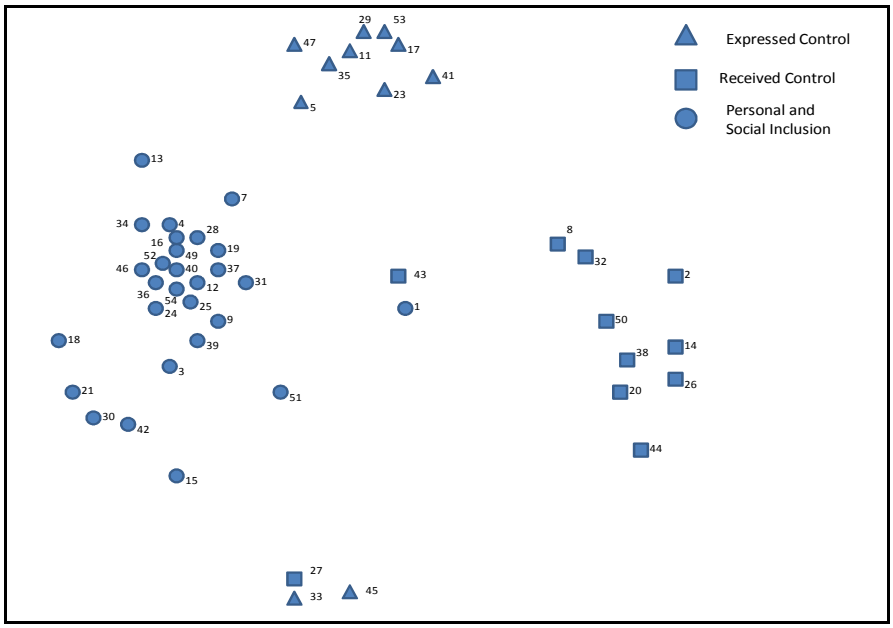
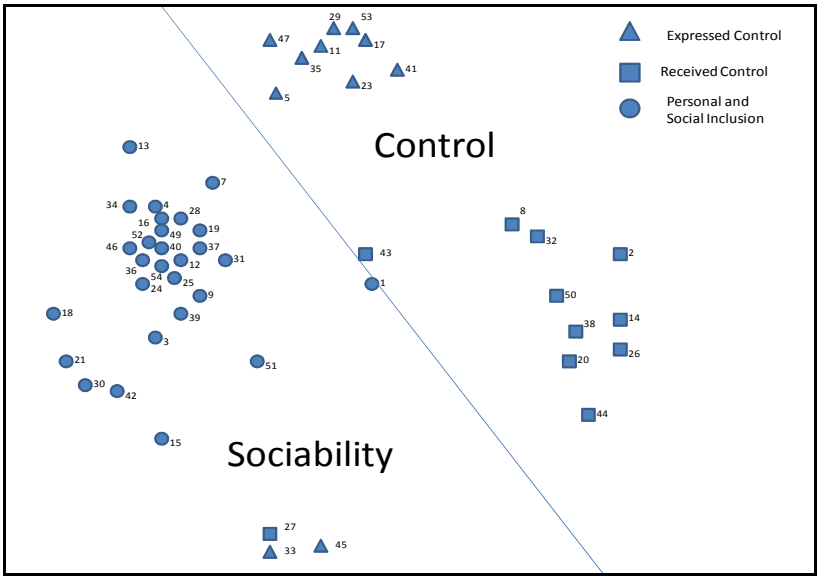


Table 10.1.5: Items of the FIRO-B in each facet

Personal Social Inclusion		Expressed Control	Received Control
1. I seek out people to be with	25. I include other people in my plans	5. I am the dominant person when I am with people	2. People decide what to do when we are together
3. I am totally honest with my close friends	28. People include me in their social affairs	11. I get other people to do things I want done	8. People strongly influence my actions
4. People invite me to do things	30. My closest friends keep secrets from me	17. I strongly influence other people's actions	14. People control my actions
6. My close friends tell me their real feelings	31. I have people around me	23. I take charge when I am with people socially	20. I am easily led by people
7. I join social groups	34. People ask me to participate in their discussions	29. I get people to do things the way I want them done	26. People decide things for me
9. I confide in my close friends	36. My friends confide in me	33. There are some things I would not tell anyone	27. There are some things I do not tell anyone
10. People invite me to join their activities	37. When people are doing things together I join them	35. I take charge when I am with people	32. People strongly influence my ideas
12. My close friends tell me about private matters	39. I have at least one friend to whom I can tell anything	41. I strongly influence other people`s ideas	38. I am strongly influenced by what people say
13. I join social organisations	40. People invite me to parties	45. There is a part of myself I keep private	43. I look for people to be with
15. I am more comfortable when people do not get too close	42. My close friends keep their feelings a secret from me	47. I take charge when I work with people	44. Other people take charge when we work together
16. People include me in their activities	46. People invite me to join them when we have free time	53. I see to it that people do things the way I want them to	50. People often cause me to change my mind
18. My close friends do not tell me about themselves	48. At least two of my friends tell me their true feelings		
19. I am included in informal social activities	49. I participate in group activities		
21. People should keep their private feelings to themselves	51. I have close relationships with a few people		
22. People invite me to participate in their activities	52. People invite me to do things with them		
24. My close friends let me know their real feelings	54. My friends tell me about their private lives		

The SSA plot confirms the results of the factor analysis, and also corroborates the decision to extract three factors. The large cluster to the left of the SSA plot contains all of the items from factor 1, the Personal and Social Inclusion items. The cluster at the top of the plot contains the items from factor 2, the expressed Control items. The cluster to the right of the plot contains the items from factor 3, the received Control items. A line can therefore be drawn through the plot separating the sociability factor (Personal and Social Inclusion) from the Control factors. However, there are three items at the bottom of the plot, in the sociability section that have been labelled expressed Control and received Control. These are ambiguous items. Examination of the loadings of item 33 in the factor analysis show that it loads most highly on to factor 2 (expressed control) but it also loads onto factor 1. Examination of the loadings of items 27 and 45 indicate that they both have very low loadings on all three factors.

Figure 10.1.b: Visual representation of the SSA for the FIRO-B data with regional interpretation



As the structure of the FIRO has been questioned, therefore how the FIRO data in this study is to be analysed is also affected. Due to the widespread use and popularity of the FIRO-B questionnaire, Schutz's (1958) original scales will not be discarded; however, the factor analysis results and the SSA (which are supported by previous research) will also not be ignored. Furthermore, Schutz (1958) proposed that the FIRO-B questionnaire as a whole provided a measure of overall social interactivity, whereby higher scores on the

questionnaire indicate a greater inclination towards social interactivity. Therefore, the FIRO-B will be analysed in terms of the total FIRO-B score, the original scales, and the three new scales.

The reliability of the FIRO-B scale was then assessed. The first assessment considered the reliability of the FIRO-B scale as a whole. The Cronbach alpha coefficient for the scale total was .91, suggesting excellent internal consistency reliability for the scale with this sample. However, analysis of the Corrected Item-Total Correlation values found that the items of the received control scale were reporting low values (less than .3), indicating that the received control scale is measuring a distinct aspect of interpersonal relationships to the rest of the scale. The reliability of the six original scales and the three new scales of the FIRO-B was then assessed. Table 10.1.8 provides the Cronbach alpha coefficients for these scales.

Table 10.1.6: Cronbach alpha coefficients for the original and new scales

	Scale	Cronbach alpha
Original scales		
	Expressed Control	.90
	Expressed Openness	.70
	Expressed Inclusion	.82
	Received Control	.84
	Received Openness	.88
	Received Inclusion	.93
New scales		
	Personal Social Inclusion	.94
	Expressed Control	.82
	Received Control	.77

We can see from Table 10.1.7 that all of the Cronbach alpha coefficients reach and indeed in most cases exceed the ideal figure of .7 (DeVellis, 2003), indicating that the items in each of the individual scales are measuring the same construct.

10.2. The FIRO-B results for the whole sample

Eleven participants did not complete a FIRO questionnaire; they were therefore removed from the analysis of the FIRO-B. Furthermore, one participant from C1 only answered the first eleven questions of the questionnaire; they were also removed, resulting in 514 cases

for analysis. The FIRO data was scrutinised for missing data, the analysis revealed that out of 27,756 items (54 FIRO items multiplied by 514 participants) there were 98 missing items, which equates to 0.35% of the data set. Analysis of those missing items found that the majority of them (19.39%) were missing on question 13. Examination of the questionnaire revealed that the reason for the high rate of answers missing for item 13 is due to the layout of the questionnaire. When the questionnaire is printed, item 13 is a small box at the top of page two, which can be easily missed by a participant. Therefore, it is concluded that participants did not intentionally miss item 13 because of an aversion to the question, but simply because they did not see the question. In order to calculate the total scores for the FIRO as a whole and for the individual scales the missing values required attention. If a value was missing on a particular scale then the mean value of that scale was used to fill in the missing value, then the total FIRO scores were calculated.

Once the issue of missing values had been addressed, the FIRO-B data was screened to assess normality. For both the original scales and the new scales the Kolmogorov-Smirnov statistic was significant, indicating a non-normal distribution of data. Furthermore, as noted in the previous section a number of outliers in the data-set were identified. Statistical transformation of the data was attempted in order to improve the distribution. However, both log transformation and square root transformation failed to improve the distribution of the data, with numerous outliers still remaining. In order to remedy the outlying scores the next most extreme scores in the data file were identified; the outlying scores were changed to one digit above the next extreme score. The above action remedied the outlying scores; however it was felt that this action, for the number of outlying scores present, altered the data-set too greatly, therefore it was decided to leave the outlying scores in the data-set and to employ non-parametric statistics instead to analyse the FIRO-B data.

The total FIRO-B scores were calculated for the whole sample. Schutz (1958) argued that the total FIRO-B score is a measure of a person's social interactivity. The FIRO-B total scores can range from a minimum score of 54 to a maximum score of 324. For the current sample the range of scores was 230, with a minimum score of 72 and a maximum score of 302. The mean score for the total FIRO-B scores for the sample was 218.15 (*SD* = 27.99), which indicates that the sample veers towards greater levels of social interactivity.

The mean scores for the whole sample for each of the scales were calculated. For each of the original scales the maximum possible score is 54, and the minimum possible score is 9. Table 10.2.1 shows the range and mean scores for each of the original scales for the complete sample.

Table 10.2.1: Range and mean scores for each of the original scales for the complete sample

Scale	Minimum score	Maximum score	Mean	Std. deviation
Expressed Control	9	54	30.40	9.30
Expressed Openness	10	54	35.37	6.80
Expressed Inclusion	9	54	39.27	7.69
Received Control	9	53	25.47	7.81
Received Openness	12	54	43.79	7.29
Received Inclusion	9	54	43.84	7.73

The above table shows a wide range of scores for each of the six scales, with almost all of the original scales obtaining the minimum and maximum scores. All of the mean scores are above the median value of the scale (22.5), with participants displaying particularly high levels of received openness and received inclusion, whereas received control has the lowest mean scores.

The mean scores for the three new scales were then calculated, Table 10.2.2 contains these results for the complete sample. For both the expressed control and the received control scales, which both contain 11 items the maximum possible score is 66, and the minimum possible score is 11. For the personal and social inclusion scale, which contains 32 items the maximum possible score is 192 and the minimum possible score is 32.

Table 10.2.2: Range and mean scores for each of the new scales for the complete sample

Scale	Minimum score	Maximum score	Mean	Std. deviation
Personal Social Inclusion	44	192	151.38	22.96
Expressed Control	11	66	34.98	9.32
Received Control	11	63	31.79	8.33

The above table again shows a wide range of scores for each of the six scales, only expressed control obtains the minimum and maximum scores possible. The mean scores for both of the control scales are above the median value of those scales (27.5), whilst the mean scores for personal and social inclusion are well above the median value for that scale (80). Participants are displaying particularly high levels of personal and social inclusion, whereas received control again has the lowest mean scores.

10.2.1. Gender and age differences in the FIRO-B scores for the whole sample

The 514 participants for FIRO-B analysis consist of 235 (45.7%) males and 276 (53.7%) females (3 participants did not indicate their gender). The participants range in age from 16 to 88, the mean age is 31.87 years (Standard Deviation = 14.50), the median age is 25.50 years. Tables 10.2.1.1 and 10.2.1.2 show the range, mean, and median scores for male and female participants for the original and the new FIRO-B scales. From these two tables we can see that the female participants report overall greater levels of social interactivity than the male participants. In-line with the gender norms reported by Schutz (1992), the female participants have scored higher on all of the sociability factors (received openness and inclusion, expressed openness and inclusion, and personal and social inclusion) than the male participants. For the control factors, again in-line with Schutz's norms, males and females report very similar scores; in this case males score slightly higher on expressed control, but virtually the same on received control.

Table 10.2.1.1: Range, median, and mean scores for each of the original scales and the total score for male and female participants

Condition	Scale	Minimum	Maximum	Median	Mean	Std. deviation
Male	Expressed Control	9	53	32.00	31.16	30.07
	Expressed Openness	10	52	34.00	34.02	6.68
	Expressed Inclusion	9	54	38.00	38.15	8.20
	Received Control	9	51	25.00	25.47	7.72
	Received Openness	12	54	43.00	41.49	7.25
	Received Inclusion	9	54	44.00	42.66	8.37
	Total score	72	280	219.00	212.94	30.07
Female	Expressed Control	9	54	30.50	29.77	9.49
	Expressed Openness	17	54	36.00	36.56	6.72
	Expressed Inclusion	11	54	41.00	40.32	7.05
	Received Control	9	53	24.00	25.43	7.90
	Received Openness	18	54	46.00	45.75	6.77
	Received Inclusion	19	54	45.00	44.92	6.97
	Total score	114	302	224.00	222.75	25.38

Table 10.2.1.2: Range, median, and mean scores for each of the new scales and the total score for male and female participants

Condition	Scale	Minimum	Maximum	Median	Mean	Std. deviation
Male	Personal Social Inclusion	44	192	148.00	145.55	24.36
	Expressed Control	11	57	36.00	35.64	9.05
	Received Control	11	63	31.00	31.71	8.38
	Total score	72	280	219.00	212.90	30.12
Female	Personal Social Inclusion	70	192	157.50	156.53	20.41
	Expressed Control	12	66	35.00	34.45	9.55
	Received Control	13	60	31.00	31.80	8.29
	Total score	114	302	224.00	222.78	25.37

Due to the non-normal distribution of the data, parametric statistics were deemed unsuitable, therefore a series of Mann-Whitney U tests were conducted. The first Mann-Whitney U test was conducted to examine the difference in the median scores for the total FIRO-B scores between the male and female participants. The test revealed a significant difference in the total FIRO-B scores of male participants (Median = 219, n = 235) and female participants (Median = 224, n = 276), U = 25643.50, z = -4.08, p < .001, r = .18.

Mann-Whitney U tests were then conducted for each of the six original scales, and for each of the three new scales, comparing male and female participants. Table 10.2.1.3 and Table 10.2.1.4 report the results of these tests.

Table 10.2.1.3: Results of the Mann-Whitney U test with z scores and probability values for the original scales for male and female participants

	Male		Female				
Scale	Median	N	Median	N	*U*	*z*	*p*
Expressed Control	32.00	235	30.50	276	29081.50	-2.01	.04
Expressed Openness	34.00	235	36.00	276	25911.00	-3.93	.000
Expressed Inclusion	38.00	235	41.00	276	27313.00	-3.08	.002
Received Control	25.00	235	24.00	276	31935.50	-.30	.77
Received Openness	43.00	235	46.00	276	20380.50	-7.26	.000
Received Inclusion	44.00	235	45.00	276	27323.00	-3.08	.002

Table 10.2.1.4: Results of the Mann-Whitney U tests with z scores and probability values for the new scales for male and female participants

	Male		Female				
Scale	Median	N	Median	N	*U*	*z*	*p*
Personal Social Inclusion	148.00	235	157.50	276	22896.50	-5.73	.000
Expressed Control	36.00	235	35.00	276	29214.50	-1.93	.05
Received Control	31.00	235	31.00	276	32420.00	-.01	.10

For the original scales male and female participants differed significantly in their scores for all of the scales apart from the received control. For the new scales there was again no statistically significant difference between male and female participants for the received control scale. However, the difference between males and females on the expressed control scale and the personal and social inclusion scale did achieve statistical significance.

Schutz (1992) has reported that scores on the FIRO-B sociability scales and the expressed control scale decrease with age, whilst scores on the received control scale increase with age. The association between age and FIRO-B scores has been assessed for the current sample using Spearman's Rank Order Correlation (rho). There was a small negative correlation between the age of the participant and the total FIRO-B scores, $rho = -.19$, $n = 514$, $p < .001$, with higher FIRO-B scores, therefore higher social interactivity, associated

with younger participants. Tables 10.2.1.5 and 10.2.1.6 detail the results of the correlations for the original scales and the new scales respectively.

Table 10.2.1.5: Spearman's Rank Order Correlation between participant age and the original FIRO-B scales

Measure	Age
Expressed Control	.10*
Expressed Openness	-.21**
Expressed Inclusion	-.17**
Received Control	-.06
Received Openness	-.16**
Received Inclusion	-.21**

* $p < .05$ (2-tailed)
** $p < .001$ (2-tailed)

Table 10.2.1.6: Spearman's Rank Order Correlation between participant age and the new FIRO-B scales

Measure	Age
Personal Social Inclusion	-.23**
Expressed Control	.06
Received Control	-.08

** $p < .001$ (2-tailed)

For the sociability scales (expressed and received openness and inclusion) and the received control scale there is a small negative correlation. As the participants get older their scores on these scales decrease, although for the received control scale this association is not significant. For the expressed control scale there is a small significant positive correlation with the age of the participant, as participants get older they score higher on the expressed control scale. For the new scales, there is a small significant negative correlation between the personal and social inclusion scale and the age of the participant. The control scales show correlations in the same direction as the original scales, although both are very small correlations and not significant. The correlations for the sociability scales conform to the results reported by Schutz (1992), younger people report higher scores on these scales. However, for the control scales, correlations opposite to Schutz's results have been found,

received control scores decreased as age increased, and expressed control scores increased as age increased, although both correlations were very small.

10.3. Analysis of the differences between the informed groups (A1 and B1), the uninformed groups (A2 and B2), and the control group (C1)

Firstly the mean and median scores for the total FIRO-B scores for the informed groups (A1 and B1) were compared to the mean and median scores for the uninformed groups (A2 and B2), and the control group (C1). The results can be seen in Table 10.3.1.

Table 10.3.1: Range, median, and mean scores for the total FIRO-B for the informed, the uninformed, and the control groups

Condition	Minimum	Maximum	Median	Mean	Std. deviation
Informed	72	302	221.00	217.47	33.45
Uninformed and Control	140	280	222.00	218.62	23.46

We can see from Table 10.3.1 that participants in the informed condition report a greater range of total FIRO-B scores than those in the uninformed and control conditions. The uninformed and control conditions report a higher mean and median score than the informed condition.

The mean and median scores were then compared for the informed, uninformed and control conditions for both the original scales and the new scales. These results can be seen in Table 10.3.2 and Table 10.3.3.

Table 10.3.2: Range, median, and mean scores for each of the original scales for the informed, uninformed, and control groups

Condition	Scale	Minimum	Maximum	Median	Mean	Std. deviation
Informed	Expressed Control	9	54	31.00	31.12	9.64
	Expressed Openness	10	53	35.00	35.43	6.72
	Expressed Inclusion	9	54	40.00	39.12	8.90
	Received Control	10	52	25.00	25.74	8.15
	Received Openness	12	54	44.00	43.40	8.06
	Received Inclusion	9	54	44.00	42.67	8.79
Uninformed and Control	Expressed Control	9	54	32.00	29.89	9.04
	Expressed Openness	14	54	35.00	35.33	6.86
	Expressed Inclusion	18	54	39.00	39.38	6.72
	Received Control	9	53	24.00	25.29	7.57
	Received Openness	20	54	45.00	44.07	6.69
	Received Inclusion	16	54	45.00	44.66	6.78

Table 10.3.3: Range, median, and mean scores for each of the new scales for the informed, uninformed and control groups

Condition	Scale	Minimum	Maximum	Median	Mean	Std. deviation
Informed	Personal Social Inclusion	44	192	156.00	149.44	26.32
	Expressed Control	11	62	36.00	35.97	9.88
	Received Control	15	63	31.50	32.06	8.73
Uninformed and Control	Personal Social Inclusion	70	192	155.00	152.74	20.20
	Expressed Control	12	66	35.88	34.28	8.84
	Received Control	11	60	31.00	31.60	8.04

10.3.1. Mann-Whitney U tests to compare the informed, uninformed, and control groups

The first Mann-Whitney U test was conducted to examine the difference in the median scores for the total FIRO-B scores between the informed and the uninformed and control groups. The test revealed no significant difference in the total FIRO-B scores of informed participants (Median = 221, n = 212) and uninformed and control participants (Median = 222, n = 302), U = 31422.00, z = -.36, p = .72, r = .02.

Mann-Whitney U tests were then conducted for each of the six original scales, and for each of the three new scales, comparing participants in both the informed and uninformed and control groups. Table 10.3.1.1 and Table 10.3.1.2 report the results of these tests.

Table 10.3.1.1: Results of the Mann-Whitney U tests with z scores and probability values for the original scales for the informed, uninformed, and control conditions

	Informed		Uninformed and Control				
Scale	Median	N	Median	N	*U*	*z*	*p*
Expressed Control	31.00	212	32.00	302	30095.50	-1.16	0.25
Expressed Openness	35.00	212	35.00	302	30846.50	-0.70	0.48
Expressed Inclusion	40.00	212	39.00	302	31493.50	-0.31	0.75
Received Control	25.00	212	24.00	302	30975.50	-0.63	0.53
Received Openness	44.00	212	45.00	302	31282.00	-0.44	0.66
Received Inclusion	44.00	212	45.00	302	28281.00	-2.26	0.02

Table 10.3.1.2: Results of the Mann-Whitney U test with z scores and probability values for the new scales for the informed, uninformed and control conditions

	Informed		Uninformed and Control				
Scale	Median	N	Median	N	*U*	*z*	*p*
Personal Social Inclusion	156.00	212	155.00	302	31192.50	-0.49	0.62
Expressed Control	36.00	212	35.88	302	28750.50	-1.97	0.05
Received Control	31.50	212	31.00	302	31143.00	-0.53	0.60

For the original scales there were no statistically significant differences between the informed and uninformed and control groups for all of the expressed scales and the received control and received openness scales. The only statistically significant result was for the received inclusion scale, with the uninformed and control group scoring significantly higher than the informed group (r = .01). For the new scales there were no statistically significant differences between the informed and uninformed and control groups for the personal social inclusion scale and the received control scale. However, the difference between the two groups on the expressed control scale did achieve statistical significance, with the informed group scoring significantly higher than the uninformed and control group (r = .09).

10.4. Analysis of the participants in the informed condition who identified the target and a matched comparison group

In order to determine whether there is any difference in the FIRO-B scores for those who identified the target in the informed condition, a matched comparison group was formed. The 36 participants who identified the target in the informed condition were matched on age, gender, occupation, and educational achievement to 36 participants also from the informed condition that had not identified the target. Before further analysis was conducted to examine the differences between the target group and the matched comparison group the normality of the distribution of the data was tested. This sub-set of data for both the original scales and the new scales was considerably more normally distributed than the whole data-set. Of the six original scales only two (expressed openness and received inclusion) reported a statistically significant Kolmogorov-Smirnov result, and there were only seven outlying scores. These extreme scores were changed to the next highest score for that scale plus one. This action remedied the extreme scores and improved the normality of the distribution. For the three new scales the Kolmogorov-Smirnov result was non-significant for the two control scales, the personal social inclusion scale was significant, and there were two outlying scores, which were dealt with as before. It was therefore decided to use parametric statistics to compare the scores for both the original scales and the new scales.

The mean scores for the total FIRO-B, the original scales, and the three new scales were calculated comparing the group of participants in the informed condition who identified the target and the matched comparison group. For the original scale total FIRO-B scores the informed group who identified the target reported higher scores of social interactivity (M = 228.26, SD = 33.87) than the matched comparison group (M = 216.93, SD = 22.91). The total FIRO-B scores for the new scales, which differ slightly due to the missing values inputted and the outlying values that were changed, show the same relationship between the groups. The informed group who identified the target reported higher scores of social interactivity (*M* = 227.98, *SD* = 34.35) than the matched comparison group (*M* = 217.13, *SD* = 23.60). Tables 10.4.1 to 10.4.2 show the mean scores for the original scales and the new scales respectively.

Table 10.4.1: Mean scores for the original scales for the informed group who identified the target and the matched comparison group

	Target group		Comparison group	
Scale	Mean	Std. Deviation	Mean	Std. Deviation
Expressed Control	33.39	11.02	29.93	8.22
Expressed Openness	34.94	5.37	36.13	5.78
Expressed Inclusion	41.68	9.52	39.47	6.62
Received Control	27.97	8.53	23.77	6.49
Received Openness	45.41	7.41	44.34	7.21
Received Inclusion	44.86	9.15	43.28	8.28

Table 10.4.2: Mean scores for the new scales for the informed group who identified the target and the matched comparison group

	Target group		Comparison group	
Scale	Mean	Std. Deviation	Mean	Std. Deviation
Personal Social Inclusion	156.04	25.48	152.05	21.81
Expressed Control	37.83	11.16	35.10	9.38
Received Control	34.11	9.05	29.98	7.31

For the original scales we can see that the informed group who identified the target scored higher than the matched comparison group on all of the scales apart from the expressed openness scale. For the new scales similar results are reported. The informed group scored higher than the matched control group on all of the scales.

10.4.1. T-tests to compare the informed group who identified the target and the matched comparison group

A series of Independent-samples t-tests were conducted to compare the mean scores of the informed group who identified the target and the matched comparison group. For the original scale FIRO-B total scores there was no significant difference in scores for the informed group (*M* = 228.26, *SD* = 33.87) and the matched comparison group (M = 216.93, *SD* = 22.91); *t* (61.49) = 1.66, p = .10 (two-tailed). The magnitude of the differences in the means (mean difference = 11.33, 95% CI: -2.30 to 24.95) was small to moderate (eta squared = .04). As the mean scores only differ slightly for the new scale total FIRO-B scores

the t-test result was the same. Table 10.4.1.1 and Table 10.4.1.2 report the results of the independent-samples t-tests for the original scales and the new scales respectively.

Table 10.4.1.1: Results of the Independent-Samples t-tests comparing the original scale mean scores for the target group and the matched comparison group

Scale	*t*	*df*	*p*	*Mean difference*	*eta squared*
Expressed Control	1.51	64.76	.14	3.45	.03
Expressed Openness	0.90	70.00	.37	-1.19	.01
Expressed Inclusion	1.14	62.44	.26	2.21	.02
Received Control	2.35	70.00	.02	4.20	.08
Received Openness	0.62	70.00	.54	1.07	.005
Received Inclusion	0.77	70.00	.45	1.58	.008

Table 10.4.1.2: Results of the Independent-Samples t-tests comparing the new scale mean scores for the target group and the matched comparison group

Scale	*t*	*df*	*p*	*Mean difference*	*eta squared*
Personal Social Inclusion	.71	70.00	.48	3.99	.007
Expressed Control	1.13	70.00	.26	2.74	.02
Received Control	2.13	70.00	.04	4.13	.06

For the original scales only one of the comparisons of the mean scores between the informed group who identified the target and the matched comparison group achieved statistical significance, the received control scale (p = .02). For the new scales the received control scale again was the only scale to achieve statistical significance (p = .04), with the informed group who identified the target scoring significantly higher than the matched comparison group.

10.4.2. Logistic regression analysis

For the original scales direct logistic regression was performed to assess the impact of the six scales on the likelihood that participants would identify the target. The model contained the six original scales as independent variables (expressed control, openness, inclusion, and received control, openness, inclusion). The full model containing all predictors was not statistically significant, χ^2 (6, *N* = 72) = 9.50, *p* = .15, however, the Hosmer-Lemeshow Goodness of Fit Test was non-significant (*p* = .47), indicating that the model was able to

distinguish between participants who identified the target and those that did not. The model as a whole explained between 12.4% (Cox and Snell R square) and 16.5% (Nagelkerke R squared) of the variance in identifications, and correctly classified 61.1% of cases. As shown in Table 10.4.2.1, only one of the independent variables made a unique statistically significant contribution to the model (received control, p = .04).

Table 10.4.2.1: Logistic regression predicting the likelihood of correctly identifying the target for the original scales

							95% C.I. for Odds Ratio	
	B	S.E.	Wald	df	Sig.	Odds Ratio	Lower	Upper
Expressed Control	.04	.03	1.94	1	.16	1.04	.98	1.10
Expressed Openness	-.05	.05	.77	1	.38	.96	.87	1.06
Expressed Inclusion	.003	.06	.003	1	.96	1.00	.90	1.12
Received Control	.08	.04	4.26	1	.04	1.08	1.00	1.17
Received Openness	.07	.05	1.68	1	.20	1.07	.97	1.18
Received Inclusion	-.02	.06	.12	1	.72	.98	.87	1.10
Constant	-3.89	2.52	2.38		.12	.02		

For the new scales direct logistic regression was also performed to assess the impact of the personal social inclusion, expressed control, and received control scores on the likelihood that participants would identify the target. The model contained three independent variables (personal social inclusion, expressed control, and received control). The full model containing all predictors was not statistically significant, χ^2 (3, *N* = 72) = 5.52, *p* = .14, however, the Hosmer-Lemeshow Goodness of Fit Test was non-significant (*p* = .44), indicating that the model was able to distinguish between participants who identified the target and those that did not. The model as a whole explained between 7.4% (Cox and Snell R square) and 9.8% (Nagelkerke R squared) of the variance in identifications, and correctly classified 52.8% of cases. As shown in Table 10.4.2.2, only one of the independent variables made a unique statistically significant contribution to the model (received control, p = .05).

Table 10.4.2.2: Logistic regression predicting the likelihood of correctly identifying the suspect for the new scales

							95% C.I. for Odds Ratio	
	B	S.E.	Wald	df	*p*	Odds Ratio	Lower	Upper
Personal Social Inclusion	.005	.01	.23	1	.63	1.01	.98	1.03
Expressed Control	.02	.03	.59	1	.44	1.02	.97	1.07
Received Control	.06	.03	3.73	1	.05	1.06	.10	1.13
Constant	-3.39	1.97	2.96		.09	.03		

10.4.3. Partial Order Scalogram Analysis

In order to establish the quantitative and qualitative variation in the FIRO-B scores between the informed group who identified the target and the matched comparison group, Partial Order Scalogram Analysis was conducted. The 72 individuals in this analysis were given a profile or structuple according to their score on the three scales of the FIRO-B. To construct these structuples the maximum and minimum scores for each of the scales was calculated, three equal groups were then formed encompassing the range of scores. Each individual was then given a score from 1 to 3 for each of the three scales. Scores of 1 indicates a low score, scores of 2 are classed as medium scores, whilst scores of 3 are classed as high scores. Table 10.4.3.1 shows the grouping of scores for the low, medium, and high groups for the three scales.

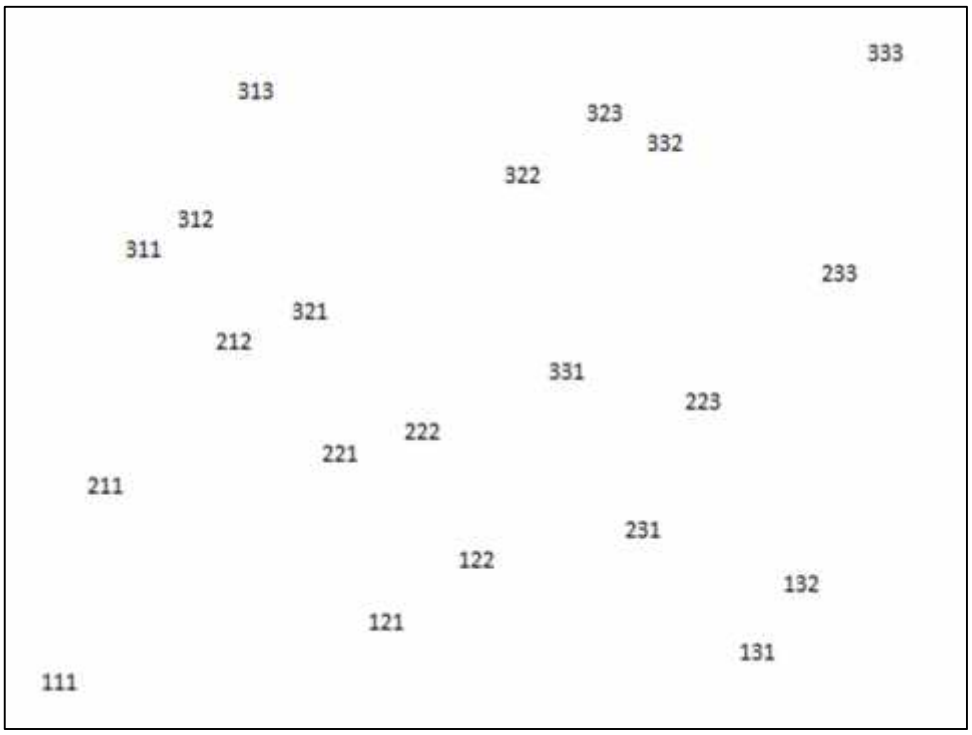
Table 10.4.3.1: Range of scale scores to constitute each structuple group

	Structuple group		
Scale scores	Low (1)	Medium (2)	High (3)
Personal & Social Inclusion	98 - 128	129 - 159	160 - 191
Expressed Control	15 - 30	31 - 46	47 - 62
Received Control	17 - 28	29 - 40	41 - 51

Therefore a person with a high Personal and Social Inclusion score, a medium Expressed Control score, and a low Received Control score, would have the structuple 321. Structuples can range from 111, indicating low scores on each of the three scales, to 333, indicating high scores on each of the three scales, and can be any combination of the three numbers in between. The structuples can be quantitatively ordered, a structuple of 333 is greater than a structuple of 222, which is then greater than a structuple of 111. Structuples can also be quantitatively the same; structuples of 121 and 211 have the same total score of 4 however they are qualitatively different. The first example, 121, consists of low Personal and Social Inclusion, medium Expressed Control, and low Received Control, which is qualitatively different to 211, medium Personal and Social Inclusion, low Expressed Control, and low Received Control. The POSA therefore partially orders the structuples according to their quantitative similarities and qualitative differences.

The structuples are represented in the POSA as points in a geometric space, placed according to their partial order. Figure 10.4.3.a shows the geometric space of the POSA which consists of 21 structuples from the 72 individuals from the informed and matched comparison groups. The analysis uses the structuples not the raw data, and each structuple is only represented in the space once. Therefore, each structuple may represent more than one individual; that is how 72 individuals are reduced to 21 structuples. The geometric space shows the increase in the scores on the three scales running along the joint axis from low scores on all three scales (111) in the lower left of the space to high scores on all three scales (333) in the upper right. We can see that structuples with the same total score occupy the same location on the joint axis, for example structuples 332, 323, and 233 in the upper right of the space, all total 8. However, they do not occupy the exact same position because they differ in how that total is constructed.

Figure 10.4.3.a: A partial order scalogram analysis of 21 profiles derived from the 72 individuals in the target and matched comparison groups

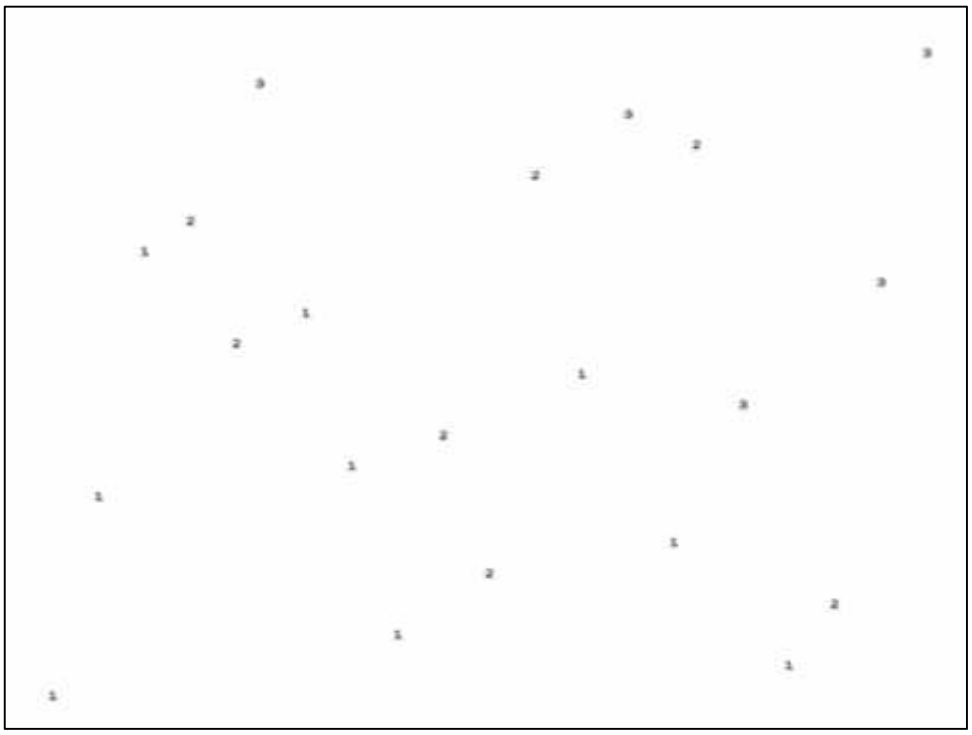


The two-dimensional solution of the POSA, shown above has a coefficient of correct representation of 0.86, indicating that there is a correct representation of 86% of structuple pairs (Borg & Shye, 1995). There is an even distribution of structuples throughout the space, indicating that the individuals in the informed and matched comparison groups show quantitative and qualitative variation in their FIRO-B scores.

In order to examine how each of the three scales contributes to the composition of the structuples in the geometric space, it is necessary to examine the item diagrams of each of the three scales. Examination of these item diagrams can determine whether the scale produces a clear partition of the space. If distinct regions of the space can be identified, then this framework may be used to distinguish FIRO-B scores between participants in the informed group and participants in the matched comparison group.

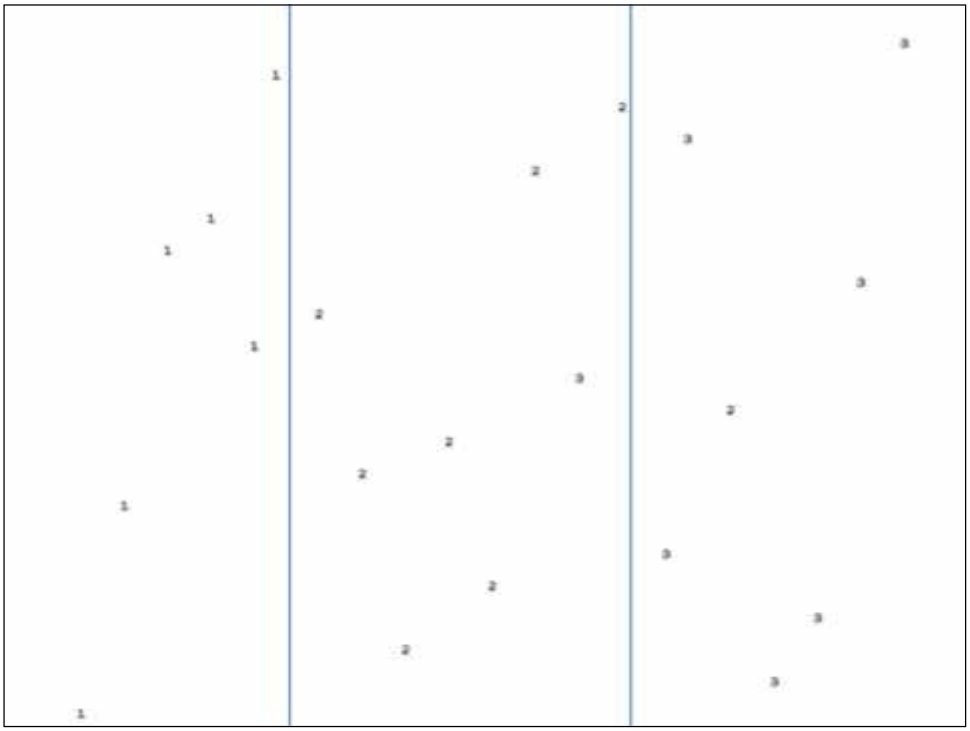
We can see from Figure 10.4.3.b that scores of 1, 2, or 3 on the received control scale do not occupy distinct regions of the space. The space cannot be partitioned due to the scores, therefore the received control scale cannot be used to distinguish between the two groups.

Figure 10.4.3.b: Item diagram for the received control scale



For the expressed control scale we can see from Figure 10.4.3.c that low, medium, and high scores on the scale can be distinguished along the X axis. The left-hand section contains only individuals who scored 1 on the expressed control scale. The majority of individuals in the middle section scored 2 on the scale, although there is a score of 3 in this section. The right-hand section consists of individuals who scored 3 on the scale, although there is a score of 2 in this section also.

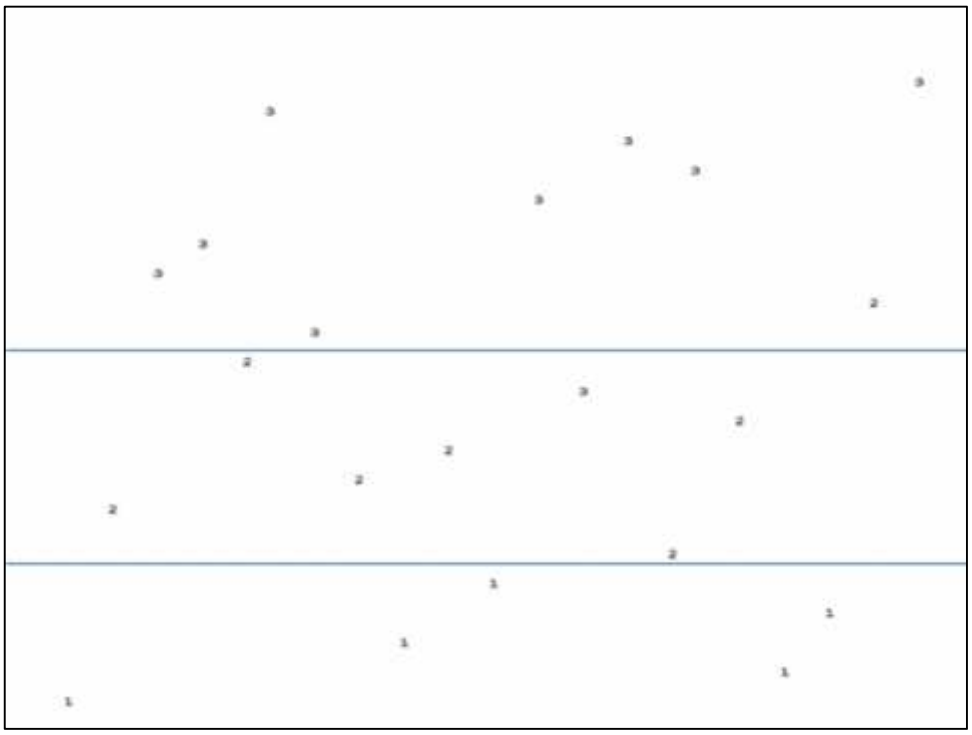
Figure 10.4.3.c: Partitioned item diagram for the expressed control scale



Note: X partitioning loading coefficient = 0.99

For the personal social inclusion scale we can see from Figure 10.4.3.d that low, medium, and high scores on the scale can be distinguished along the Y axis. The lower section contains all of the individuals who scored 1 on the scale. The majority of individuals in the middle section scored 2 on the scale, although there is a score of 3 in this section. The upper section consists of individuals who scored 3 on the scale, although there is a score of 2 in this section also.

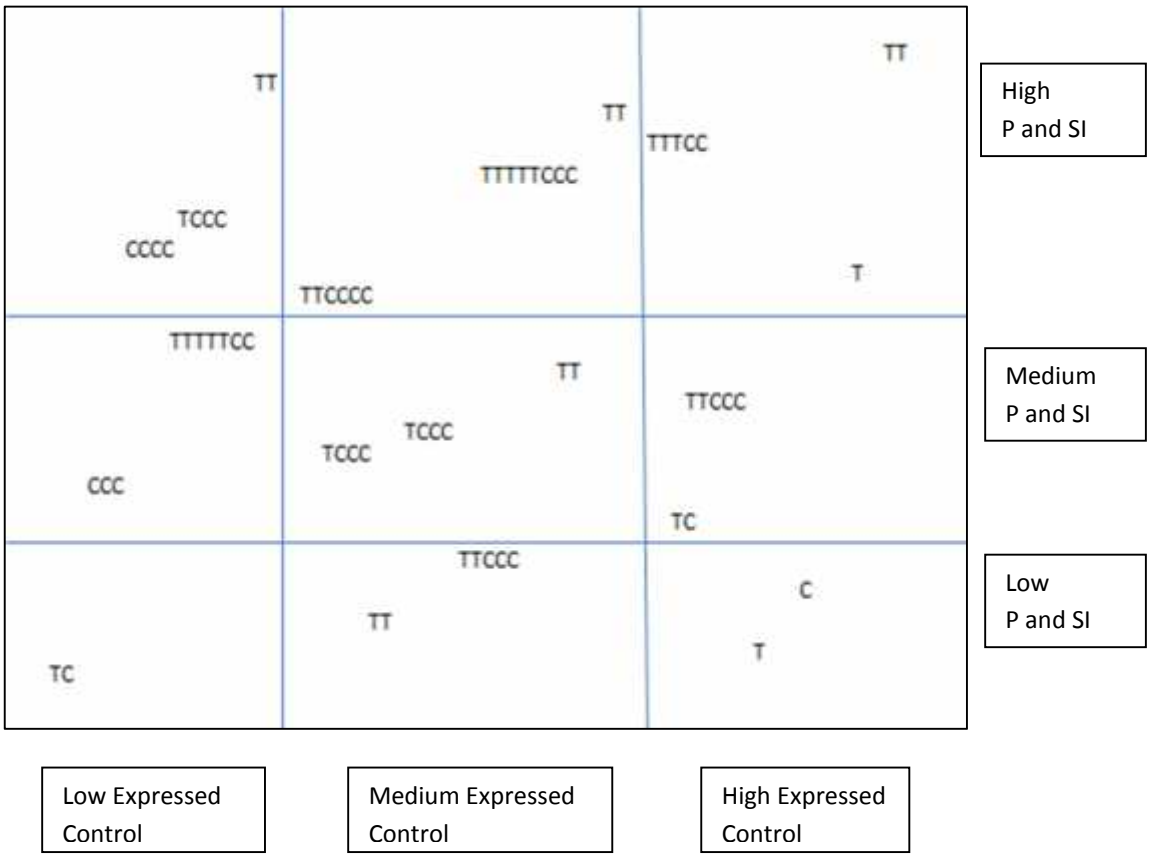
Figure 10.4.3.d: Partitioned item diagram for the personal and social inclusion scale



Note: Y partitioning loading coefficient = 1.00

In Figure 10.4.3.e each one of the 72 individuals is placed in the space according to their score on the three scales. A 'T' corresponds to an individual from the informed group who identified the target; a 'C' is an individual from the matched comparison group. Therefore in the upper right of the space there are two individuals from the informed group who scored 333 on the three scales. The partitioning of the space along the X and Y axis results in nine sections of the space; comprised of high, medium, and low personal and social inclusion, and high, medium, and low expressed control.

Figure 10.4.3.e: A partial order scalogram analysis of 25 profiles showing the location of the informed individuals who identified the target (T) and the matched comparison individuals (C) with the superposition of the two partition lines



With the partitions of the space in place it is clear that individuals in the informed and matched comparison groups display qualitative and quantitative differences in their FIRO-B scores, as T's and C's are spread throughout the space. However, there are some areas of the plot that distinguish between individuals in the informed and matched comparison groups. The high expressed control and high personal and social inclusion sections mainly contain individuals from the informed condition who identified the target. On the opposite side of the space though, the section with high personal and social inclusion but low expressed control, this area is mainly constituted with individuals from the matched comparison group.

10.5. Analysis of those participants who identified the target in the informed, uninformed, and control conditions

In the informed condition (A1 and B1) 36 participants identified either number eight or number four, and were therefore deemed to have made a target identification. In the

uninformed condition (A2 and B2) 27 participants made a target identification, whilst 25 participants made a target identification (either number four or number eight) in the control condition (C1). This section will compare the FIRO-B scores (total, original scales and three new scales) for the participants in each of these groups.

Firstly the normality of the distribution of the data for the three groups was assessed. For the original FIRO-B scales the data was not particularly normally distributed. Five of the seven scales reported a significant Kolmogorov-Smirnov statistic, only the expressed control scale and the total FIRO-B score reported a non-significant statistic. Sixteen outlying scores were changed to less extreme scores, which marginally improved the normality of the distribution. However, three of the scales (expressed openness, received openness, and received inclusion) still reported significant Kolmogorov-Smirnov statistics, therefore for the original scales non-parametric statistics were used to further analyse the data.

The FIRO-B data for the three new scales was substantially more normally distributed. Only the received control scale reported a significant Kolmogorov-Smirnov statistic. Five outlying scores were changed to less extreme scores which improved the normality of the data. As the outlying scores had been sufficiently dealt with parametric statistics were employed to analyse the new FIRO-B scales.

10.5.1. Comparison of the mean scores of the total FIRO-B scale, the original scales and the new scales for the informed, uninformed, and control groups

Firstly, the mean scores for the total scores for the three groups were compared. We can see from Tables 10.5.1.1 and 10.5.1.2 that there is a substantial difference in the total scale scores between the three groups, with the informed group reporting higher mean scores of social interactivity than the uninformed group and the control group reporting the lowest social interactivity scores. For the original scales the informed group scored higher than the uninformed group and the control group on the expressed control, expressed inclusion, received openness, and received control scales, but lower on the expressed openness scale. For the received inclusion scale the informed group scored lower than the control group but higher than the uninformed group.

Table 10.5.1.1: Mean scores of the original scales for the informed, uninformed and control groups who identified the target

	Informed group		Uninformed group		Control group	
Scale	Mean	Std. deviation	Mean	Std. deviation	Mean	Std. deviation
Expressed Control	33.39	11.02	31.22	6.74	23.08	10.75
Expressed Openness	35.03	5.58	35.48	5.62	36.72	7.83
Expressed Inclusion	41.41	10.23	39.35	6.86	37.56	6.08
Received Control	28.44	9.55	24.96	6.40	25.48	8.49
Received Openness	45.41	7.41	43.70	7.62	42.28	7.59
Received Inclusion	44.86	9.15	44.26	7.61	45.18	5.88
Total score	228.54	34.78	218.98	24.05	210.30	26.25

For the new scales the informed group scored higher than the uninformed group and the control group on all three scales. The uninformed group in turn scored higher than the control group on all three scales.

Table 10.5.1.2: Mean scores of the new scales for the informed, uninformed and control groups who identified the target

	Informed group		Uninformed group		Control group	
Scale	Mean	Std. deviation	Mean	Std. deviation	Mean	Std. deviation
Personal Social Inclusion	156.04	25.48	154.20	17.90	150.02	19.33
Expressed Control	37.83	11.16	34.81	7.15	28.80	10.94
Received Control	33.97	8.78	31.07	6.71	31.11	9.04
Total score	227.84	34.21	220.09	20.71	209.93	26.25

10.5.2. Further analysis comparing the informed, uninformed and control groups who identified the target

For the original scales of the FIRO-B a series of Kruskal-Wallis tests were conducted to determine if there were any significant differences in the FIRO-B mean scores between the informed, uninformed, and control groups. The differences in the total FIRO-B score approached significance (p = .06), however, only the expressed control scale reported statistically significant differences in their scores between the three groups χ^2 (2, n = 88) = 11.29, p < .01. The control group recorded lower median scores (Md = 26) than both the informed group (Md = 30.50) and the uninformed group (Md = 33). Three post hoc Mann-

Whitney U tests were conducted. To control for a Type I error a Bonferroni correction was applied resulting in a revised alpha level of .017. The Mann-Whitney U tests revealed a significant difference in the expressed control scores between the informed group and the control group (U = 240.50, z = -3.08, p = .002, r = .39), and the uninformed group and the control group (U = 190.00, z = -2.71, p = .007, r = .38). There was no significant difference between the scores of the informed group and the uninformed group.

A series of one-way between-groups ANOVA's with post-hoc tests were conducted to determine if there were any significant differences in the FIRO-B mean scores between the informed, uninformed, and control groups for the three new scales. Only the total FIRO-B score and the expressed control scale reported statistically significant differences in their scores between the three groups at the $p < .05$ level. For the total FIRO-B score, according to Levene's test, the variance in the scores was not homogenous for each of the three groups, therefore the Brown-Forsythe F-ratio is reported; F (2, 80.76) = 3.20, p = .05. Despite only just reaching statistical significance the actual difference in mean scores is quite large, which is evidenced in the medium effect size obtained (eta squared = .06), which may be due to the small sample in this analysis. Because of the heterogeneity of variance Games-Howell post-hoc tests were conducted. This test is particularly recommended when there are unequal sample sizes and heterogeneity of variance (Field, 2005). The Games-Howell post-hoc tests indicated a difference between the informed group (M = 227.84, SD = 34.21) and the control group (M = 209.93, SD = 26.25) suggesting a trend towards statistical significance (p = .06). The uninformed group (M = 220.09, SD = 20.71) did not differ significantly from either the informed group or control group.

The expressed control scale of the new scales also violated Levene's test for homogeneity of variance, therefore the Brown-Forsythe F-ratio is again reported; F (2, 74.48) = 6.23, p = .003, eta squared = .12. As before Games-Howell post-hoc tests were conducted. The post-hoc comparisons indicated that the expressed control mean score for the informed group (M = 37.83, SD = 11.16) was significantly different to the control group (M = 28.80, SD = 10.94, $p < .01$). The uninformed group (M = 34.81, SD = 7.15) did not differ significantly from either the informed group or the control group.

10.6. Further analysis of the participants in the informed condition

In this section the participants in the informed condition (groups A1 and B1) are further analysed. In particular the FIRO-B scores for those who identified the target in the informed group will be compared to those who did not identify the target in the informed group. 36 participants identified the target, compared to 176 participants who did not identify the target. This results in a sample for these analyses of 212 participants. Firstly the normality of the distribution of the data was analysed to determine the suitability of parametric testing. For the original scales only the expressed control and the received control scales achieved a non-significant Kolmogorov-Smirnov statistic. Once the outlying scores had been dealt with the Kolmogorov-Smirnov statistics did not change, for the expressed openness, expressed inclusion, received openness, and received inclusion scales the statistic was still significant. For the three new scales the personal social inclusion scale achieved a significant Kolmogorov-Smirnov statistic, even when the outlying scores had been changed. Due to the non-normal distribution of data non-parametric statistics were employed for both the original scales and the new scales.

10.6.1. Comparison of the FIRO-B mean scores for the participants in the informed condition who made target and non-target identifications

From Tables 10.6.1.1 and 10.6.1.2 we can see that participants who made a target identification in the informed condition scored higher on all of the FIRO-B new scales and all of the original scales apart from expressed openness.

Table 10.6.1.1: Mean scores of the original FIRO-B scales for the target and non-target identifications in the informed condition

	Target identification		Non-target identification	
Scale	Mean	Std. deviation	Mean	Std. deviation
Expressed Control	33.39	11.02	30.66	9.30
Expressed Openness	35.03	5.58	35.51	6.94
Expressed Inclusion	41.41	10.23	38.65	8.56
Received Control	28.44	9.55	25.18	7.75
Received Openness	45.41	7.41	42.99	8.15
Received Inclusion	44.86	9.15	42.22	8.67
Total score	228.54	34.78	215.21	32.81

Table 10.6.1.2: Mean scores of the new FIRO-B scales for the target and non-target identifications in the informed condition

	Target identification		Non-target identification	
Scale	Mean	Std. deviation	Mean	Std. deviation
Personal Social Inclusion	156.04	25.48	148.09	26.36
Expressed Control	37.83	11.16	35.59	9.59
Received Control	34.66	10.34	31.52	8.29
Total score	228.54	34.85	215.20	32.80

10.6.2. Mann-Whitney U tests to compare the participants in the informed condition who identified the target to those in the informed condition who did not identify the target

A series of Mann Whitney U tests were conducted to compare the original scale FIRO-B scores of those in the informed condition who identified the target to those who did not identify the target. There was a significant difference in the total FIRO-B scores for those who identified the target (*Md* = 238.00, *n* = 36) and those who did not identify the target (*Md* = 219.38, *n* = 176), *U* = 2456, *z* = -2.12, p < .05, *r* = .15. Target identifiers were significantly more socially interactive than non-identifiers. Table 10.6.2.1 shows the results of the Mann-Whitney U tests for the original scales.

Table 10.6.2.1: Results of the Mann-Whitney U tests comparing the original scale scores for participants making target and non-target identifications in the informed condition

	Target		Non-Target				
Scale	Median	N	Median	N	*U*	*z*	*p*
Expressed Control	30.50	36	31.00	176	2744.50	-1.26	.21
Expressed Openness	34.00	36	35.00	176	2953.00	-0.64	.52
Expressed Inclusion	41.25	36	40.00	176	2549.00	-1.85	.07
Received Control	28.00	36	25.00	176	2531.50	-1.90	.06
Received Openness	46.00	36	44.00	176	2634.50	-1.59	.11
Received Inclusion	47.00	36	44.00	176	2515.50	-1.95	.05

The Mann-Whitney U tests conducted on the original scales found that only the scores for the received inclusion scale were significantly different for the two groups (*r* = .13), with target identifiers reporting significantly higher received inclusion scores than non-

identifiers. Target identifiers also reported higher received control scores than non-identifiers, although this was just above significance.

As expected the Mann-Whitney U test for the total FIRO-B scores for the new scales also reported a significant difference between those who identified the target (*Md* = 238.00, *n* = 36) and those who did not identify the target (*Md* = 219.30, *n* = 176) in the informed condition; *U* = 2452.50, *z* = -2.13, *p* < .05, *r* = .15. The Mann-Whitney U tests for the new scales revealed no significant differences in the scores for those who did identify the target and those who did not identify the target in the informed condition. As Table 10.6.2.2 shows the personal and social inclusion scale was the only scale close to achieving a statistically significant difference between the two groups.

Table 10.6.2.2: Results of the Mann-Whitney U tests comparing the new scale scores for participants making target and non-target identifications in the informed condition

	Target		Non-Target				
Scale	Median	N	Median	N	*U*	*z*	*p*
Personal Social Inclusion	160.00	36	154.00	176	2574.50	-1.77	.08
Expressed Control	35.00	36	36.00	176	2803.50	-1.09	.28
Received Control	33.00	36	31.00	176	2653.00	-1.54	.12

10.7. Analysis of the FIRO-B scores of the administrators

Nine of the sixteen administrators returned a completed FIRO-B questionnaire. In the first stage of the analysis the mean scores for the administrator sample were compared to the participant sample on the six original scales, the three new scales and the total score.

Table 10.7.1 shows that overall for the total FIRO-B scores the administrators and the participants barely differ at all in their reported levels of social interactivity. Indeed on many of the scales the two groups report comparable scores. However, the scores on the received control scale for both the original scales and the new scales are noticeably different for the two groups, with the administrator sample reporting lower scores than the participant sample.

Table 10.7.1: Comparison of the mean scores of the administrators and the participants for the total scores, the original scales, and the new scales

	Administrator sample		Participant sample	
	Mean	Std. deviation	Mean	Std. deviation
Expressed Control	30.67	9.39	30.40	9.30
Expressed Openness	36.78	7.48	35.37	6.80
Expressed Inclusion	41.89	10.12	39.27	7.69
Received Control	16.78	3.96	25.47	7.81
Received Openness	43.22	9.65	43.79	7.29
Received Inclusion	47.78	5.29	43.84	7.73
Personal Social Inclusion	158.33	24.30	151.38	22.96
Expressed Control	35.56	9.14	34.98	9.32
Received Control	23.22	5.26	31.79	8.33
Total	217.11	26.55	218.15	27.99

A series of Mann-Whitney U tests were conducted to compare the scores of the administrators and the participants. For the FIRO-B total scores there was no significant difference between the two groups. For the original scales only the received control scale reported a significant difference between the administrators (*Md* = 16, *n* = 9) and the participants (*Md* = 25, *n* = 514), *U* = 711.50, *z* = -3.57, *p* < .001, *r* = .16. For the new scales only the difference in the received control scores was significantly different for the administrators (*Md* = 22, *n* = 9), and the participants (*Md* = 31, *n* = 514), *U* = 892.50, *z* = -3.16, *p* < .01, *r* = .14.

In the second stage of the analysis the FIRO-B scores of those administrators in the informed condition are more closely scrutinized. Of the six administrators in the informed condition who were informed of the location of the target in the line-up, four returned a completed FIRO-B questionnaire. Table 10.7.2 shows the percentage of target identifications that each administrator who returned a FIRO-B is responsible for.

Table 10.7.2: Percentage of target identifications for the Administrators in the informed condition who returned a FIRO-B questionnaire

Administrator	Target identifications %
2	19.44
3	19.44
8	30.56
10	5.56

From the above table it is clear that Administrator 8 is responsible for a large share of the target responses in the informed condition. Administrator 10 however, is responsible for a small share of those responses. Is there a difference then in the FIRO-B scores of these two individuals which can account for the difference in the number of target responses they achieved? As we can see from Table 10.7.3 the FIRO-B scores for these two individuals indicate that there is a large difference in the FIRO-B total scores, with Administrator 8 reporting a lower score of social interactivity than Administrator 10. For the original scales Administrator 8 reported a higher expressed control score, but scored lower on the five remaining scales. In particular Administrator 8 scored substantially lower on the received control and the received openness scales than Administrator 10.

Table 10.7.3: Comparison of the scores for the four Administrators in the informed condition for the total scores, the original scales, and the new scales

	Administrator 2	Administrator 3	Administrator 8	Administrator 10
Expressed Control	39	34	33	18
Expressed Openness	42	44	27	34
Expressed Inclusion	44	54	30	45
Received Control	14	13	11	23
Received Openness	45	51	22	51
Received Inclusion	48	53	45	54
Personal Social Inclusion	162	187	119	176
Expressed Control	48	40	35	20
Received Control	22	22	14	29
Total	232	249	168	225

For the new scales Administrator 8 scored substantially lower on the personal social inclusion scale and the received control scale, but higher on the expressed control scale than Administrator 10. In sum the Administrator who achieved the most target identifications from their line-ups was less socially interactive than the Administrator who achieved the least target identifications. They reported lower scores on all of the scales apart from the expressed control scales, indicating that the Administrator who achieved the most target identifications relates to being the dominant person in social interactions. As this is only a comparison between two individuals, definitive conclusions cannot be drawn between higher levels of expressed control and the ability to influence the outcome of a line-up.

When the analysis of the remaining two Administrators (who were both responsible for 19.44% of target identifications) from the informed condition is considered, see Table 10.7.3, further conclusions can be intimated. As before those Administrators responsible for the greater number of target identifications reported higher levels of expressed control, but lower levels of received control. Therefore those Administrators who exert greater levels of influence on their participants relate more to the dominant role in an interpersonal relationship, whereby they influence the actions of others and take charge of the social situation. Analysis of four individuals again does not provide a conclusive link between tendencies towards dominance in social situations and the ability to influence; however, it does provide an interesting avenue for further research. Therefore in the next chapter the transcripts of the line-ups will be analysed for any evidence that dominant interpersonal tendencies has had an impact on the line-up identification.

10.8. Summary of the main findings of this chapter

There are three findings of particular note from the analysis of the FIRO-B data. Firstly, when the participants in the informed condition who identified the target were matched on age, gender, educational attainment, and occupation to a group of participants from the informed condition that did not identify the target, the target identifiers reported significantly higher received control scores than the matched comparison group. The differences between the two groups on the other scales were not significant. Secondly, when comparing the participants in the informed condition who identified the target to

participants in the informed condition who did not identify the target, the target identifiers were significantly higher in their overall social interactivity and their received inclusion scores. Furthermore, although not quite reaching statistical significance, target identifiers reported higher received control scores than non-identifiers. Finally, the third finding of note suggests that Administrators who are informed of the location of the target and who are responsible for more target identifications report higher expressed control scores and lower received control scores, than informed Administrators who are responsible for more non-target identifications.

Analysis Section 4 – Analysis of the Transcripts from the Line-ups: Examining Verbal Cues

Chapter 11:

As part of the line-up procedure, administrators were asked to audio-record their line-ups. Eleven of the sixteen administrators returned the transcripts of their recordings, resulting in 381 cases for analysis. These transcripts were analysed for the presence of a variety of verbal cues that may influence the participants in their choice of line-up member.

11.1. Coding of the transcripts

The transcripts were analysed by the author, and a coding dictionary of verbal behaviours was created. Table 11.1.1 provides a full list of the verbal cues identified and an example of each from the transcripts.

Table 11.1.1: Verbal behaviours of the Administrators with examples from the transcripts

Verbal behaviours identified	Example
Tell witness to look carefully	"Examine them all carefully"
Tell witness to take time	"It's ok take your time"
Ask if sure after identification	"Ok, number eight, are you certain?"
Ask if sure after non-identification	"Are you sure number one?"
Ask to look again after identification	"Do you want to look again?"
Ask to look again after non-identification	"You can have another look if you like"
Repeat choice with questioning tone	"Number twelve?"
Repeat choice with confirmatory tone	"Number five, thank you"
Show or offer to show line-up photographs again	"Would you like to see the photos again"
Press for an identification	"I need you to give me an answer"
Say do not know who suspect is	"Well I don't know who he is"

The 381 transcripts were coded by the author for the presence or absence of the above eleven behaviours. A postgraduate student blind to the hypotheses of the study also coded the transcripts using the coding dictionary. Kappa Measure of Agreement tests were used to assess inter-rater agreement. According to Peat (2001) a Kappa Measurement of Agreement value above .8 represents very good agreement. From Table 11.1.2 we can see that the coding of the transcripts was very consistent, with ten of the eleven behaviours reporting Kappa values above .9.

Table 11.1.2: Kappa Measure of Agreement values for the eleven behaviours identified

Administrator Behaviours	Kappa Measure of Agreement
Tell witness to look carefully	.93
Tell witness to take time	.98
Ask if sure after identification	1.00
Ask if sure after non-identification	.96
Ask to look again after identification	1.00
Ask to look again after non-identification	1.00
Repeat choice with questioning tone	.93
Repeat choice with confirmatory tone	.96
Show or offer to show line-up photographs again	1.00
Press for an identification	.89
Say do not know who suspect is	1.00

Note: All values are significant at the $p < .001$ level

A second postgraduate student, also blind to the hypotheses and conditions of the study was employed to resolve the disagreements between the two raters. The proportion of the administrator behaviours were then compared for the informed condition, the uninformed condition and the control condition. Table 11.1.3 shows the percentage of the administrator behaviours for the three conditions.

Table 11.1.3: Percentage of each of the Administrator's behaviours in their line-ups in the informed, uninformed, and control conditions (frequencies in parentheses)

	Condition		
Verbal behaviours identified	Informed	Uninformed	Control
Tell witness to look carefully	4.20% (6)	1.30% (2)	0.00% (0)
Tell witness to take time	12.00% (17)	6.90% (11)	3.80% (3)
Ask if sure after identification	2.80% (4)	1.30% (2)	0.00% (0)
Ask if sure after non-identification	3.50% (5)	5.00% (8)	0.00% (0)
Ask to look again after identification	0.00% (0)	0.60% (1)	0.00% (0)
Ask to look again after non-identification	1.40% (2)	0.60% (1)	0.00% (0)
Repeat choice with questioning tone	12.00% (17)	8.20% (13)	1.30% (1)
Repeat choice with confirmatory tone	23.20% (33)	11.90% (19)	1.30% (1)
Show or offer to show photographs again	9.90% (14)	5.70% (9)	3.80% (3)
Press for an identification	14.80% (21)	8.20% (13)	20.00% (16)
Say do not know who suspect is	0.00% (0)	0.00% (0)	3.80% (3)

From the above table we can see that the informed condition has the highest percentage of administrator behaviours for seven of the eleven identified behaviours. The most prevalent behaviours identified were the 'repeat choice with a confirmatory tone', 'press for an identification', 'repeat choice with a questioning tone', and 'tell the witness to take their time'. The least frequent administrator behaviours were 'asking the witness to look at the photos again after an identification', and 'asking the witness to look again after a non-identification'. The control condition has the lowest percentage of administrator behaviours for all of the identified behaviours, apart from the 'press for an identification' behaviour. Furthermore, the behaviour of 'saying they do not know who the suspect is' is only perpetrated by administrators in the control condition.

In order to compare the administrator behaviours between the three groups, the frequencies of each of the administrator behaviours was calculated. A score of one was given for each incidence of an administrator behaviour, therefore, an administrator who told the witness to take their time, pressed for an identification, and repeated their choice with questioning tone would obtain a score of three. The total scores for each of the three groups were calculated and can be seen in Table 11.1.4.

Table 11.1.4: Number and percentage of administrator behaviours identified and average number of line-ups per behaviour in each condition

Condition	Number of line-ups	Number of behaviours identified	Average number of line-ups per behaviour
Informed	142	119 (52.89%)	1.19
Uninformed	159	79 (35.11%)	2.01
Control	80	27 (12.00%)	3.96
Total	381	225 (100%)	1.69

We can see from this table that the administrators in the informed condition contributed more of the administrator behaviours than those in the uninformed condition. Administrators in the uninformed condition in turn contributed more of the administrator behaviours than those administrators in the control condition. We can also see that overall administrator behaviours occurred on average in nearly every other line-up. For the informed condition they occurred in almost four out of every five line-ups, whilst for the

uninformed condition this dropped to every other line-up, and for the control condition they occurred on average in one in three line-ups.

11.2. The length of the transcripts

The transcripts were then analysed for their length and the number of words spoken between the administrator and the participant. It would be expected that those administrators who exhibited the greatest number of the behaviours identified above, and were the most influential with their participants would have the most discourse with their participants. In order to determine the length of the interaction, the number of words spoken by both the administrator and participant were calculated. Some administrators transcribed the whole of their interaction, whilst others only transcribed the line-up section. Counting the whole of the interaction would skew the results towards those who transcribed the whole of the interaction. Therefore, only the words spoken by the administrator and the participant during the line-up are calculated. Administrators spoke during the line-up an average of 40.66 (SD = 17.59) words, whilst the participant spoke an average of 15.43 (SD = 12.31) words.

Table 11.2.1 shows the mean number of words spoken by the administrator and participant in the informed, uninformed and control groups. We can see from this table that administrators in the informed group did indeed speak to their participants more than administrators in the uninformed and control groups. However, surprisingly administrators in the control condition spoke to their participants more than administrators in the uninformed condition.

Table 11.2.1: Mean number of words spoken by Administrators and participants in the three conditions

	Administrator	Participant
Informed	48.83 (21.86)	18.95 (12.76)
Uninformed	30.20 (9.03)	9.94 (9.11)
Control	46.48 (7.99)	20.10 (12.92)

We can also see from this table that administrators were responsible for more discourse than the participants in all three conditions. Participants in the informed condition spoke on average double the amount of the participants in the uninformed condition. However, the

participants in the control condition spoke marginally more than the participants in the informed condition. Kruskal-Wallis tests revealed statistically significant differences in the number of words spoken by both the administrator; χ^2 (2, n = 381) = 94.58, p < .001, and the participant; χ^2 (2, n = 381) = 71.53, p < .001, in the three conditions.

A series of Mann-Whitney U tests with Bonferronni adjustments were conducted (adjusted alpha level .017). For the number of words spoken by the administrator there was a significant difference between the informed condition (*Md* = 58.00, n = 142) and the uninformed condition (*Md* = 28.00, n = 159), U = 6199.00, z = -6.76, p < .001, r = .39, and between the uninformed condition and the control condition (*Md* = 47.50, n = 80), U = 1231.50, z = -10.18, p < .001, r = .67. However the difference between the informed condition and the control condition was non-significant, U = 5199.00, z = -1.05, p > .05, r = .07. For the number of words spoken by the participant there was a significant difference between the informed condition (*Md* = 16.50, n = 142) and the uninformed condition (*Md* = 7.00, n = 159), U = 6037.50, z = -6.98, p < .001, r = .40, and between the uninformed condition and the control condition (*Md* = 16.00, n = 80), U = 2737.00, z = -7.21, p < .001, r = .47. However the difference between the informed condition and the control condition for the number of words spoken by the participant was non-significant, U = 5237.50, z = -.97, p > .05, r = .07.

11.3. Comparison of the transcripts obtaining target identifications with those obtaining non-target identifications

Of the 381 transcripts returned, 301 belonged to the informed and uninformed conditions. Of those 301 transcripts, 42 were transcripts where the participant had made a target identification, and 259 were transcripts where the participant had made a non-target identification. For the 42 target transcripts 39 administrator behaviours were observed, showing that on average behaviours were occurring in almost every line-up. For the 259 non-target transcripts 159 administrator behaviours were observed, for this group behaviours occurred on average in every 1.62 line-ups. The target transcripts were compared to the non-target transcripts to see if any differences in the number of words spoken by the administrator and the participant existed between them. Table 11.3.1 shows the mean number of words spoken by the administrator, and mean number of words spoken by the participant for target and non-target transcripts.

Table 11.3.1: Comparison of the mean number of words spoken by the Administrator and the participant, for target and non-target transcripts

	Words spoken administrator	Words spoken participant
Target	43.64 (21.09)	13.95 (9.45)
Non-target	38.23 (18.37)	14.23 (12.21)

This table indicates that administrators who obtained target identifications from their participants spoke to them more during the line-up procedure, than administrators who obtained non-target identifications. However, for the number of words spoken by the participant the reverse is found. Participants who made target identifications spoke marginally less than participants who made non-target identifications. When Mann-Whitney U tests were conducted, the differences between the target and non-target transcripts for the number of words spoken by the administrator and participant were non-significant.

11.4. The relationship between the transcripts and the Administrator's FIRO-B scores

Six of the Administrators who returned transcripts of their line-ups also returned a completed FIRO-B questionnaire. A series of Spearman's Rank Order Correlations (rho) were conducted to explore the relationship between the FIRO-B scale scores of the Administrators, the mean number of words spoken by the Administrators, and the mean number of words spoken by the participants.

For the mean number of words spoken by an Administrator there was a strong positive correlation with expressed control scale scores (*rho* = .77, *n* = 6, *p* =.07), and a strong negative correlation with received control (*rho* = -.77, *n* = 6, *p* =.07), both of these correlations were just above the alpha level of significance. These correlations indicate a trend suggesting that an Administrator who reports higher levels of expressed control and lower levels of received control speaks more during the line-up

For the mean number of words spoken by the participants there was a strong negative correlation with the Administrator's reported expressed control (*rho* = -.89, *n* = 6, $p < .05$), indicating that participants use fewer words with Administrators who seek to control the interaction. However, there was a strong positive correlation between the mean number of

words spoken by participants and the Administrators received openness score (*rho* = .84, *n* = 6, *p* < .05), and received inclusion score (*rho* = .79, *n* = 6, *p* = .06), the received inclusion correlation was just above the alpha level of significance. Participants therefore, use more words in the line-ups when the Administrators report higher scores of received openness and inclusion.

11.5. Summary of the main findings of this chapter

When the transcripts of Administrators in the informed, uninformed, and control conditions were compared, Administrators in the informed condition were found to interact for longer with their participants, and exhibit more of the identified verbal cues. However, participants in the control condition were found to speak marginally more than participants in the informed condition. In the informed condition, when the transcripts of those identifying the target were compared to those that did not identify the target, the Administrators of the target identifiers used more words, and exhibited more of the verbal cues.

Finally, from the correlation analysis it appears that Administrators who report higher levels of expressed control, those who wish to control the social interaction, spoke more words during the line-up. Whereas, Administrators who report higher levels of received control, those who do not wish to control the social interaction, spoke fewer words. For the mean number of words spoken by the participant, participants spoke more words when the Administrator was more receptive to openness and inclusion from others. However, the participants used fewer words when the Administrator reported higher levels of expressed control.

Chapter 12: Discussion

12.1. Discussion

The aim of this research was to examine whether an experimenter's knowledge could influence a participant to pick a particular person from a photographic line-up. Facilitators (A and B) were informed of the location of the target in the line-up. They informed half of their administrators of the location of the target (administrators A1 and B1), but did not inform the other half (administrators A2 and B2). There was also a control group where both the Facilitator (C) and the administrators (C1) were not informed of the location of the target. In condition 1 administrators recruited participants and asked them to pick the person that they thought was most likely to be the man convicted of the Lockerbie bombing. Participants then completed the FIRO-B questionnaire. For participants in condition 2, once they had completed the FIRO-B, they were again asked to make a choice, but this time from a target-absent line-up. The line-up procedures were audio-recorded.

12.1.1. Evidence for the experimenter expectancy effect

The research conducted by Greathouse and Kovera (2009) and others (Haw & Fisher, 2004; Phillips et al., 1999) has found evidence that the administrator of a line-up, when they know the identity of the suspect, can influence the response of the witness. However the research to date has proved contradictory, with the effect being found with both simultaneous and sequential line-ups, and biased and unbiased instructions. Furthermore, the eyewitness's memory may limit the influence of the administrator. Therefore, Greathouse and Kovera (2009) called for research to examine the effect of the quality of the memory on investigator bias. In response to this, this research has removed the memory component, thereby leaving a view of the administrator's influence not confounded by the witness's memory.

With regard to the presence of experimenter expectancy effects, the research made two hypotheses. Firstly, it stated that those participants in the informed condition would identify the target more often than those participants in the uninformed condition and the control condition. Secondly, those participants in the uninformed condition would identify the target more often than those in the control condition. The author would therefore tentatively state that there is evidence of experimenter expectancy effects in this research. In condition A1 the target suspect was identified the most frequently. In condition B1 the

target was the joint most frequently identified. In conditions A2 and B2 the target was the second most frequent identification. A chi-square goodness of fit test indicates that the identifications of the target in condition A1 was significantly greater than the identifications of the target in the control condition.

Furthermore, similar results were found in the target absent line-up. Identifications of the target were the most frequent in conditions A1, B1, and B2, and the second most frequent in condition A2. A non-significant chi-square test indicates that there was no association between correct or incorrect identifications in the first line-up and correct or incorrect identifications in the second line-up. This finding suggests that a conclusion that the administrator's knowledge of the location of the target impacted the participant's choice can be drawn.

The frequencies of the identifications also show a high frequency of identifications of number 7 in conditions A1 and A2. Indeed in condition A1 identifications of number 7 were the second most frequent and in A2 were the most frequent identifications. In condition B1 identifications of number 3 are the joint most frequent identifications, whilst in B2 identifications of number 3 are the third most frequent. When the line-up is viewed (see Appendix 4), it can be seen that line-up members 3 and 7 come immediately before the targets 4 and 8. Therefore some of the identifications may be explained by some leakage of an experimenter expectancy effect, but where the participant misinterprets the cues from the administrator. Indeed an explanation for this may be provided by the processes highlighted in the original experimenter expectancy research conducted by Rosenthal (1976). In those original person perception experiments, experimenters were led to expect a range of scores, either positive or negative, from their participants. Therefore, their influence of their subjects was on a continuum, and was not for a specific number. The same process may have been in evidence in this research. The experimenter may have been exhibiting cues in the general area of numbers 4 and 8, but the participants could not distinguish the cues and therefore were influenced to pick number 3 or 7.

It should also be considered that in the criminal justice system personnel are particularly motivated to obtain a particular result. Whether that is a police officer motivated to obtain an identification of a suspect from a line-up, a finger-print analyst motivated to match a

latent print to a suspects print, or a lawyer motivated to successfully prosecute a case (Koppl & Sacks, 2013). In this research however, administrators were not particularly motivated to obtain an identification. They were all misinformed as to the true rationale of the experiment, and were led to believe that they were examining different types of line-up procedure. The instructions given to the administrators in the informed and uninformed conditions provided a very subtle motivating instruction that each participant should make a selection, and would hopefully make a correct selection. However, the extent to which the administrators construed that statement as an invitation to influence the participant is unknown.

Indeed, the participants were most probably oblivious to any influencing behaviours exhibited by the administrator. Clark et al. (2009) found that even when experimenters actively tried to influence their participants, the participants rated the influence as very low, a mean score of 1.66 from a scale of 1 to 9. Clark et al. (2009) suggest that this may be because they either did not notice or remember the interaction or they did not feel that the statements made by the experimenter were particularly manipulative. The motivation of the administrator to obtain a particular selection may explain why the expectancy effect was not as pronounced in condition B1. For condition A the Facilitator was informed that the suspect was number 8 in the line-up, and was told to make this known to the administrators in A1. For condition B though, the Facilitator was told that the administrator in B1 would be told that the suspect was number 4, even though that was not true. The Facilitator may have consciously or unconsciously communicated the information that number 4 was not the actual target. The administrator may not have been as motivated to influence the participant to make a selection that they may have known was wrong.

Furthermore, if the participants were attending to the demand characteristics of the experiment (Orne, 1962), it is not completely clear why. There is some evidence for Orne's (1962) good subject effect. When asked to pick someone from a line-up without having seen the person before, the majority of participants did not query the request. Some participants expressed doubt that they would be correct but they picked someone anyway. Indeed, the vast majority of the participants selected a person from the line-up, only a very small minority refused to comply. Analysis of the transcripts found that some participants asked the Administrator if they "had got it right" or had selected "the right person". This indicates

that for some participants they were apprehensive of giving the wrong answer, and were apprehensive of being negatively evaluated by the Administrator (Rosenberg, 1965). It is difficult to determine whether the participant was obedient to the authority of the Administrator. As stated before, the vast majority were obedient to the experimental procedure because they complied with it. Therefore, whilst the Administrator in their role as the experimenter commands some authority, perhaps participants would have been more obedient to the procedure if an authority figure such as a police officer had conducted the line-up.

12.1.2. The impact of interpersonal relations

The FIRO-B questionnaire has provided some interesting insights into the relationship between experimenter expectancy effects and interpersonal behavioural style. As the FIRO-B has not been utilised in this type of research before, hypotheses about the direction of the relationship were not proposed. Therefore, exploratory comparisons of FIRO-B scores were conducted. Comparisons between those identifying the target in the informed group and a matched comparison group found significantly higher received control scores for the target identifiers. This group were also found to have higher received control scores when compared to the participants in the informed condition who did not make a correct identification. Although this result just missed statistical significance (p = .06) it does suggest a trend for this group of target identifiers to have high received control scores and therefore to be controlled by others in social situations. Furthermore, those correctly identifying in the informed group were significantly higher in both their total FIRO-B score and their received inclusion score than those incorrect in the informed group. It can be tentatively concluded then that individuals who are more socially interactive, but also score higher in received control, which is characterised as a tendency to let others dominate social situations, and higher in received inclusion, which is characterised as a need to be included in social situations, may be more susceptible to cues from the experimenter communicating their expectancy.

When considering the FIRO-B scores of the administrators, further interesting results emerged. However, caution must be taken when drawing conclusions from a small sample. Nevertheless, subtle differences were found when comparing the FIRO-B scores of the

administrators achieving a high number of correct identifications to the administrators achieving a low number of correct identifications. In particular higher scores of expressed control and lower scores of received control were reported by administrators achieving a high number of correct identifications compared to administrators achieving a low number of correct identifications. It would therefore appear that the communication of expectancies is particularly conducive if the administrator relates to the dominant controlling role in a social interaction, and the participant relates to the submissive and compliant role.

12.1.3. The impact of verbal cues

Analysis of the transcripts of the line-up procedures identified the presence of 11 verbal statements spoken by the administrator that could be construed as cues to the participant. Overall it was found that administrators in the informed condition exhibited more of these verbal behaviours than administrators in both the uninformed and control conditions. Certain behaviours were identified that informed administrators exhibited at greater rates than uninformed and control administrators, For example, informed administrators were more likely to repeat the participant's choice in a confirmatory tone, repeat the choice in a questioning tone, and tell the participant to take their time. Uninformed administrators exhibited some verbal behaviours; most notably repeating the choice with both a confirmatory and questioning tone. However, the control administrators were the most likely to press for an identification out of all the administrators, and were the only group to tell the participant that they did not know the identity of the suspect.

Analysis of the length of the interaction between the administrator and the participant found, perhaps unsurprisingly that the informed administrators, who exhibited the most verbal behaviours also spoke to their participants the most. One possible explanation for this is the higher expressed control scores reported by administrators in the informed condition. Their wish to control the social situation may manifest itself by controlling the dialogue of the interaction. More surprising perhaps is the finding that control administrators spoke to their participants more than the uninformed administrators, and did not differ much from the informed administrators. Furthermore, when the number of words spoken by the participant is considered, the participants in the control condition spoke more than in the other two conditions. This may be because the participants in the

control condition were not getting any cues from the administrator as to who the suspect was, so they felt they had to draw the conversation out to help them make an identification.

Further interesting results were found when Administrator's FIRO-B scores were compared to their transcripts. Perhaps not surprisingly, Administrators who spoke more reported higher scores of expressed control; they dominate the social interaction through their dialogue. Whereas Administrators who spoke fewer words reported higher received control scores; they let others dominate the social interaction with their dialogue. Of more interest though is the effect of the Administrators interpersonal style on the number of words spoken by the participant. When the Administrator reported higher expressed control scores, where they dominate the social interaction, the participant spoke fewer words. However, participants spoke more words when the Administrator scored higher on the received openness and received inclusion scales; Administrators who exhibit an open and inclusive style therefore encourage participants to interact with them.

When transcripts for correct identifications were compared to transcripts for incorrect identifications, it was found that administrators who obtained correct identifications spoke to their participants more and exhibited more verbal behaviours. However, these differences were not significant. Nevertheless, this still indicates that whilst some expectancies are communicated verbally, the kinesic communication highlighted by Rosenthal (1976) is also an important channel of communication.

12.2. Implications of the research

We have seen in this research that participants are susceptible to the cues exhibited by administrators who know the location of the suspect in the line-up. Moreover, through the high incidences of identifications of number 3 and 7 from the line-up, we have seen that either, administrators are not always successful at correctly communicating their knowledge to the participant, or participants are not always successful at accurately interpreting the cues. Furthermore, we have seen that subtle, non-directive verbal statements made by administrators can influence the choice of the participant. The real-life implications of this research are clear then; eyewitnesses who are susceptible to the demand characteristics of the line-up, but have a poor memory of the event may be at greater risk of making a false identification. This is particularly true if the administrator makes seemingly innocuous

statements to the eyewitness. The risk of a non-identification of a culprit or an identification of a foil is also increased if the witness misinterprets the cues from the administrator, and makes their identification on the basis of this.

Moreover, as Clark et al, (2009) argue a line-up administrator can still exert influence even when blind to the location of the suspect, and unbiased instructions are given. This research therefore highlights the importance of considering the social interaction between the experimenter and the participant and by extension the interaction between the real-life line-up administrator and the eyewitness. If an administrator can provide influence with their mere presence during a line-up, perhaps research should experiment with line-ups where the administrator leaves the room after giving their instructions. However, considering the resistance to double-blind procedures it is unlikely that the criminal justice system would submit to such a procedure. Therefore, it is imperative that a better understanding of the factors that moderate the communication of experimenter expectancy effects is gained.

This research therefore highlights the importance of conducting double-blind line-up procedures. However, law enforcement agencies have continued to resist this recommendation, citing procedural hurdles and financial costs as reasons for not adopting this procedure, regardless of the growing body of research. The results of this research have further suggested that a Facilitator informed of the location of the suspect may inadvertently communicate this knowledge to a line-up administrator. The implication then is that a non-blind investigator asking a blind investigator to conduct a line-up may not result in a double-blind line-up. It is difficult to see how the procedural hurdles, particularly in small police stations with limited man power, could be overcome to ensure a true double-blind line-up, where all those concerned with the line-up were unaware of the identity of the suspect.

Law enforcement agencies are not alone though in resisting the double-blind recommendation. In a review of leading journals in experimental science, examining the extent of the utilisation of blind procedures, Sheldrake (1998) has identified “the widespread neglect of possible experimenter effects” (p.76). He reports that out of 237 reviewed papers from the physical sciences, none employed a blind methodology. In the

biological sciences, 0.8% employed a blind methodology, this figure rose to 5.9% in the medical sciences, and 4.9% for psychology and animal behaviour. However, for parapsychology 85.2% of the papers reviewed used a blind methodology. Furthermore, a survey of 11 British university science departments found that blind methodologies are rarely taught or used. However as this body of research continues to expand, to include the potential for bias in the wider criminal justice system, the laboratories, the crime scenes, and the court rooms, it must become ever harder for the criminal justice system to ignore the implications of this body of research.

12.3. Limitations of the research

As this research utilised the photograph of a suspect from a high profile case, there is the possibility that the participants recognised Mr Megrahi. This would then invalidate the suggestion that the Administrator's knowledge of the identity and location of the suspect influenced the participant to pick that photograph. However, there are a number of results from the study that suggest that the participants did not have a memory for Mr Megrahi. Firstly, if participants did indeed recognise Mr Megrahi we would expect participants to choose his picture in all of the conditions. In conditions B1, B2, and C1, the percentage of participants choosing Mr Megrahi's photograph were 5.40%, 5.00%, and 8.20% respectively. These figures show that in these groups the majority of participants did not recognise Mr Megrahi from the line-up. The target absent condition also provides evidence that the participants did not have a memory for Mr Megrahi. Participants in this condition did not refuse to pick a photograph because the photograph of Mr Megrahi was not present. Furthermore, in the informed condition (A1) of the target absent line-up participants picked photograph number 8, even though it was not Mr Megrahi, more often than the other photographs. Both of these points suggest that the participants were not choosing based on their memories of Mr Megrahi. Nevertheless, in order to eradicate the possibility that the participants had a memory of Mr Megrahi, this research should be repeated with a line-up of foils.

Although administrators were given guidelines as to how to conduct the line-ups, they were also given flexibility as to the setting and location of the line-ups. As a result there are a number of factors, which were discussed in Chapters 1 and 4, which may have influenced

the results of the study. For example, the study did not control for the lighting conditions for the line-ups, or the length of time that participants were allowed to view the photographs. Whilst the returned transcripts do indicate that administrators did follow the instructions for conducting the line-ups, as some transcripts were not returned, we cannot be completely confidant that all of the line-ups followed the instructed procedure.

Furthermore, the instructions given to Facilitators in conditions A and B differed. Group A facilitators were told that the target was number 8 in the line-up, however group B facilitators were instructed to tell their administrators that the target was number 4 even though they knew this not to be true. Therefore, variations in instructions, location, and how the line-ups were conducted, mean that participants were not all subjected to the exact same experimental procedure. These factors may then confound the results suggesting that it was the administrator's knowledge of the location of the target which led to the higher rates of identification of that target.

In the real world, identification procedures are conducted in a variety of settings and locations, by administrators using their own interpretation of the guidelines for how to conduct those procedures. This study is therefore akin to real-life investigative practices, and has shown that administrators of line-ups, who know the location of the target, may transfer this knowledge to the witness and influence them to pick the target. Nevertheless, to further the understanding of how this knowledge is communicated this research should be repeated with standardised line-up procedures, where each participant undergoes the exact same procedure.

There are many factors which were explored in Chapter 1, that were not examined in this study. This research chose to concentrate on the impact of interpersonal behaviour. Whilst the FIRO-B has revealed interesting relationships between aspects of interpersonal behaviour and susceptibility to influence, there are many more potentially important factors to consider, which may have had a bearing on the results of this study. Indeed, Rosenthal examined many factors when exploring experimenter expectancy effects, including the experience and motivation of the experimenter. The complex interplay of biosocial factors, including gender, age, and ethnicity, and psychosocial attributes between the experimenter and participant have also been considered. However, no concrete rules concerning the

relationship between these factors and experimenter expectancy effects have been advanced. Therefore, in order to fully explore the impact of the administrators knowledge of the location of the target, in this new type of experimental procedure where the memory component is removed, further research is needed to unpick the tangled web of these biosocial and psychosocial factors.

The sample for this research was relatively large, however, we cannot generalise that this sample is representative of the wider population. We only have the responses of the participants who agreed to take part in the experiment, and therefore do not know if the current sample is particularly susceptible to or immune from influence. Furthermore, as only nine of the administrators returned a completed FIRO-B questionnaire it is difficult to draw definitive conclusions from such a small sample. More research is therefore needed with a greater number of administrators to obtain a better understanding of the differences between them, and the impact of these differences on the experimenter expectancy effect.

The FIRO-B is a questionnaire and therefore employs a self-report methodology. As with all self-report methodologies there is a danger that participants may answer in a socially desirable manner. In Chapter 2 the research of Orne (1962) and Rosenberg (1965) suggested that participants may behave or state that they behave in a manner which portrays them as a 'good' person. Although this effect is particularly seen with questionnaires asking about illicit or illegal behaviours, this effect may also occur with a personality questionnaire such as the FIRO-B. In particular participants may wish to make themselves appear more popular by exaggerating on the social inclusion questions. Or participants worried about appearing too controlling or too submissive may underestimate their levels of expressed and received control. In order to alleviate some of these concerns this research employed an instrument that has been in use for many years, and has proved a valid measure of interpersonal behaviour. Furthermore, it was stressed to the participants that their responses to the FIRO-B would be completely anonymous.

As the line-up procedures were audio-recorded but not video-recorded, we are limited to analysing the verbal behaviour of the administrator and the participant. Therefore the kinesic communication, highlighted by Rosenthal (1976) and Greathouse and Kovera (2009), including the administrator's expression, or how long they left the photo in front of the

participant, was not available for analysis. Furthermore, not all of the administrators returned their transcripts of the line-up procedure. Those that were not returned may have indicated administrator behaviours not identified in the transcripts that were returned. They may also have highlighted particularly influencing verbal behaviours that would have corresponded with or contradicted the results of this research.

12.4. Future directions for research

This research has suggested that experimenters can influence participants to pick a particular person from a photo-array, in a line-up type procedure, without the confounding effect of a memory of the person. This therefore supports the research that has suggested that line-up administrators with knowledge of the location of the suspect can also influence an identification by an eyewitness.

Research has in recent years started to tackle the problem of bias in the criminal justice system. This research has moved from examining the effect of the administrator's knowledge of the location of the suspect to more recently examining the potential for bias in forensic laboratories, crime scenes, and court rooms. Whilst experimenter expectancy effects have been well established, the recommendations to reduce these effects, namely double-blind procedures are still resisted by the criminal justice system. Therefore, more research is needed to provide an irrefutable argument for the need for double-blind procedures. In particular this research needs to concentrate on the interaction between the experimenter and participant.

Despite the magnitude of the body of research demonstrating experimenter expectancy effects, there is a paucity of understanding of how these expectancies are communicated, and what factors moderate their communication. Rosenthal (1976) suggested that both kinesic and paralinguistic channels were important for the communication of expectancies, a finding that has been validated by subsequent research (Garrioch & Brimacombe, 2001; Greathouse & Kovera, 2009). However, this research has failed to date to determine why an innocuous verbal statement such as 'examine the line-up carefully,' or a smile, or a body movement would influence a person to make a particular choice from a line-up. Furthermore, research to date has largely ignored the moderating factors of the

experimenter and the participant, their gender, age, personality, and the effect these have on the communication and interpretation of expectancies.

This research therefore attempted to measure just one small aspect of personality, interpersonal behaviour, using just one instrument of measurement. The results of this research have indicated that certain interpersonal styles may moderate the communication of expectancies, whilst other interpersonal styles may aid in the interpretation of those expectancies. More research then is needed in general to examine the factors that moderate the communication of expectancies. But specifically there is a need to focus on the social interaction between the experimenter and the participant and the role of interpersonal relations in this interaction. In particular, research could pit high expressed control administrators against low received control participants, or examine the effects of different combinations of interpersonal style of the administrator and the participant on the communication and interpretation of expectancies. Interestingly though, research has suggested that some people are not able to read and interpret other peoples cues and therefore may be immune to experimenter expectancy effects (Rosnow et al., 1994; Rosnow, Skleder, & Bind, 1995). Is there then an interpersonal style that renders the expectancies of the experimenter obsolete?

References

Ainsworth, P.B. (1998). *Psychology, Law and Eyewitness Testimony,* Chichester: Wiley.

Ainsworth, P.B. (2000). *Psychology and Crime: Myths and Reality,* Essex: Pearson Education Limited.

Ainsworth, P.B., & King, E. (1988). Witnesses' perceptions of identification parades. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical Aspects of Memory: Current Research and Issues, Vol 1* (pp. 66-70). Chichester, UK: Wiley.

Anthony, T., Cooper, C., & Mullen, B. (1992). Cross-racial identification: A social cognitive integration. *Personality and Social Psychology Bulletin, 18,* 296-301.

Bartlett, M.S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society, 16* (Series B), 296-298.

Bartlett, J.C., & Memon, A. (2006). Eyewitness memory of young and older eyewitnesses. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Handbook of eyewitness psychology: Memory for people* (pp. 309-338). Mahwah, NJ: Lawrence Erlbaum Associates.

Behrman, B.W., & Davey, S.L. (2001). Eyewitness Identification in Actual Criminal Cases: An Archival Analysis. *Law and Human Behavior, 25,* 475-491.

Berkson, J., Magath, T.B., & Hurn, M. (1940). The error of estimate of the blood cell count as made with the haemocytometer. *American Journal of Physiology, 128*, 309-323.

Binder, A., McConnell, D., & Sjoholm, N.A. (1957). Verbal conditioning as a function of experimenter characteristics. *Journal of Abnormal and Social Psychology, 55,* 309-314.

Borg, I., & Lingoes, J.C. (1987). *Multidimensional Similarity Structure Analysis*, New York: Springer.

Borg, I., & Shye, S. (1995). *Facet theory: Form and content*. Newbury Park, CA: Sage.

Bradfield, A.L., Wells, G.L., & Olson, E.A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology, 87,* 112-120.

Brewer, N., & Wells, G.L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity and target-absent base rates. *Journal of Experimental Psychology: Applied, 12,* 11-30.

Brigham, J.C., Van Verst, M., & Bothwell, R.K. (1986). Accuracy of children's eyewitness identifications in a field setting. *Basic and Applied Social Psychology, 7,* 295-306.

Brigham, J.C., Wasserman, A.W., Meissner, C.A. (1999). Disputed Eyewitness Identification Evidence: Important Legal and Scientific Issues. *Court Review, 36,* 12-25.

Brimacombe, C.A.E., Quinton, N., Nance, N., & Garrioch, L. (1997). Is age irrelevant? Perceptions of young and old adult's eyewitnesses. *Law and Human Behavior, 21,* 619-634.

Brown, E., Deffenbacher, K., & Sturgill, W. (1977). Memory for faces and the circumstances of the encounter. *Journal of Applied Psychology, 62,* 311-318.

Buckhout, R. (1974). Eyewitness testimony. *Scientific American, 231,* 23-31.

Burt, D.B., Zember, M.J., & Niederehe, G. (1995). Depression and memory impairment: A meta-analysis of the association, its pattern, and specificity. *Psychological Bulletin, 17,* 285-305.

Canter, D.V. (2009). Relating to the Appeal of Abdelbaset Ali Mohmed al-Megrahi: With particular attention to the evidence of Mr Anthony Gauci. Retrieved November 28, 2012, from http://www.iaip.org/uploads/Examination%20of%20Eyewtiness%20Testimony%20used%20to%20Convict%20the%20Lockerbie%20Bomber.pdf

Canter, D., & Fritzon, K. (1998). Differentiating arsonists: A model of firesetting actions and characteristics. *Legal and Criminological Psychology, 3*, 73-96.

Canter, D., & Youngs, D. (2009). *Investigative Psychology: Offender Profiling and the Analysis of Criminal Action*, West Sussex, UK: Wiley.

Canter, D.V., Youngs, D. & Hammond, L. (2009). *Experimenter Effects on Respondent Choice: The 'Fundamental Administrator Effect' Example from Photo Line-up Identifications.* Unpublished report on file with the author.

Canter, D.V., Youngs, D., & Hammond, L. (2013). Cognitive bias in line-up identifications: The impact of administrator knowledge. *Science and Justice, 53,* 83-88.

Carlsmith, J.M., Ellsworth, P.C., & Aronson, E. (1976). *Methods of Research in Social Psychology*. Philippines: Addison-Wesley Publishing Company Inc.

Catell, R.B. (1966). The scree test for number of factors. *Multivariate Behavioral Research, 1,* 245-276.

Chance, J.E., & Goldstein, A.G. (1984). Face-recognition memory: Implications for children's eyewitness testimony. *Journal of Social Issues, 40,* 69-85.

Clark, S.E. (2005). A Re-examination of the Effects of Biased Lineup Instructions in Eyewitness Identification. *Law and Human Behavior, 29,* 575-604.

Clark, S.E. (2008, December 18). Report on Identification Procedures: Abdelbaset Ali Mohmed al-Megrahi v. H.M. Advocate. Retrieved November 28, 2012, from http://www.megrahimystory.net/downloads/Professor%20Steve%20Clark's%20report%2018%2012%2008.pdf

Clark, S.E. & Godfrey, R.D. (2009). Eyewitness identification evidence and innocence risk. *Psychonomic Bulletin and Review, 16*, 22-42.

Clark, S.E., Howell, R.T., & Davey, S.L. (2008). Regularities in eyewitness identification. *Law and Human Behaviour, 32,* 187-218.

Clark, S.E. & Tunnicliff, J.L. (2001). Selecting lineup foils in eyewitness identification experiments: Experimental control and real-world simulation. *Law and Human Behavior, 25,* 199-216.

Clark, S.E., Marshall, T.E., & Rosenthal, R. (2009). Lineup Administrator Influences on Eyewitness Identification Decisions. *Journal of Experimental Psychology: Applied, 15,* 63-75.

Cohen, J.W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cordaro, L., & Ison, J.R. (1963). Observer bias in classical conditioning of the planarian. *Psychological Reports, 13,* 787-789.

Cutler, B.L., & Penrod, S.D. (1988). Improving the reliability of eyewitness identification: Lineup construction and presentation. *Journal of Applied Psychology, 73,* 281-290.

Cutler, B.L., & Penrod, S.D. (1995). *Mistaken Identity.* Cambridge: Cambridge University Press.

Cutler, B.L., Penrod, S.D., & Martens, T.K. (1987a). The reliability of eyewitness identifications: The role of system and estimator variables. *Law and Human Behavior, 11,* 223-258.

Cutler, B.L., Penrod, S.D., & Martens, T.K. (1987b). Improving the reliability of eyewitness identifications: Putting context into context. *Journal of Applied Psychology, 72,* 629-637.

Cutler, B.L., Penrod, S.D., O'Rourke, T.E., Martens, T.K. (1986). Unconfounding the effects of contextual cues on eyewitness identification accuracy. *Social Behaviour, 1,* 113-134.

Deffenbacher, K.A. (1983). The influence of arousal on reliability of testimony. In B.R. Clifford, & S. Lloyd-Bostock (Eds.), *Evaluating witness evidence: Recent psychological research and new perspectives* (pp. 235-251). Chichester: John Wiley & Sons.

Deffenbacher, K.A., Carr, T., & Leu, J. (1981). Memory for words, pictures and faces: Retroactive inference. *Journal of Experimental Psychology, Human Learning and Memory, 7,* 299-305.

Dekle, D.J. (1997). Testing delays resulting in increased identification accuracy in line-ups and show-ups. *Journal of Offender Rehabilitation, 25,* 35-49.

DeVellis, R.F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, California: Sage.

Devlin, Lord P. (1976). *Report to the Secretary of State for the Home Department on the Departmental Committee on Evidence of Identification in Criminal Cases.* London: HMSO.

Dodson, C.S., & Krueger, L.E. (2006). I misremember it well: Why older adults are unreliable eyewitnesses. *Psychonomic Bulletin and Review, 13,* 770-775.

Doherty, W., & Colangelo, N. (1984). The Family FIRO Model: a modest proposal for organizing family treatment. *Journal of Marital and Family Therapy, 11,* 299-303.

Doob, A.N., & Kirshenbaum, H.M. (1973). Bias in police line-ups – partial remembering. *Journal of Police Science and Administration, 1,* 287-293.

Douglass, A.B., & Steblay, N.M. (2006). Memory Distortion in Eyewitnesses: A Meta-Analysis of the Post-identification Feedback Effect. *Applied Cognitive Psychology, 20,* 859-869.

Dror, I. (2012a). Cognitive forensics and experimental research about bias in forensic casework. *Science and Justice, 52,* 128-130.

Dror, I. (2012b). Expectations, contextual information, and other cognitive influences in forensic laboratories. *Science and Justice, 52,* 132.

Dror, I. (2013). The ambition to be scientific: Human expert performance and objectivity. *Science and Justice, 53,* 81-82.

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

Dror, I., & Hampikian, G. (2011). Subjectivity and bias in forensic DNA mixture interpretation. *Science and Justice, 51,* 204-208.

Dror, I., & Rosenthal, R. (2008). Meta-analytically Quantifying the Reliability and Biasability of Forensic Experts. *Journal of Forensic Science, 53,* 900-903.

Duncan, B.L. (1976). Different social perceptions and attribution of intergroup violence: testing the lower limits of stereotyping of blacks. *Journal of Personality and Social Psychology, 34,* 590-598.

Duncan, S., Rosenberg, M.J., & Finkelstein, J. (1969). The paralanguage of experimenter bias. *Sociometry, 32,* 207-219.

Dunning, D., & Perretta, S. (2002). Automaticity and eyewitness accuracy: a 10-12 second rule for distinguishing accurate from inaccurate positive identifications. *Journal of Applied Psychology, 87,* 951-962.

Dvir, T., Eden, D., & Banjo, M.L. (1995). Self-fulfilling Prophecy and Gender: Can Women be Pygmalion and Galatea? *Journal of Applied Psychology, 80,* 253-270.

Ebbesen, E., & Flowe, S. (2002). *Simultaneous versus sequential lineups: what do we really know?* Retrieved April 11, 2013, from http://www2.le.ac.uk/departments/psychology/ppl /hf49/ SimSeq%20Submit.pdf

Ellis, H.D., Shepherd, J.W., Flin, R., & Davies, G. (1989). Identification from a computer-driven retrieval system compared with a traditional mug-shot album search: A new tool for police investigations. *Ergonomics, 32,* 167-177.

Ellison, K.W., & Buckhout, R. (1981). *Psychology and Criminal Justice.* New York: Harper & Row.

Ferguson, D.C., & Buss, A.H. (1960). Operant conditioning of hostile verbs in relation to experimenter and subject characteristics. *Journal of Consulting and Clinical Psychology, 24,* 324-327.

Field, A. (2005). Discovering Statistics Using SPSS (2nd Ed). London: Sage Publications Ltd.

Fruzzetti, A.E., Toland, K., Teller, S.A., & Loftus, E.F. (1992). Memory and Eyewitness Testimony. In M. Gruneberg, & P. Morris (Eds.), *Aspects of Memory, Vol 1: The Practical Aspects* (pp. 18-50). London: Routledge.

Garrioch, L., & Brimacombe, C.A.E. (2001). Lineup Administrators' Expectations: Their Impact on Eyewitness Confidence. *Law and Human Behavior, 25,* 299-315.

Geddes, L. (2012). Miscarriages of justice will occur. *New Scientist, 7 February,* 10-12.

Giannelli, P.C., Weatherhead, A.J. III, Weatherhead, R.W. (2010). Independent Crime Laboratories: The Problem of Motivational and Cognitive Bias. *Utah Law Review, 2,* 247-266.

Gonzalez, R., Ellsworth, P.C., & Pembroke, M. (1993). Response biases in lineups and showups. *Journal of Personality and Social Psychology, 64,* 525-537.

Greathouse, S.M., & Kovera, M.B. (2009). Instruction Bias and Lineup Presentation Moderate the Effects of Administrator Knowledge on Eyewitness Identification. *Law and Human Behavior, 33,* 70-82.

Green, G. (2001). Vulnerability of witnesses with learning disabilities: Preparing to give evidence against a perpetrator of sexual abuse. *British Journal of Learning Disabilities, 29,* 103-109.

Groves, R.M., Salfati, C.G., & Elliot, D. (2004). The Influence of Prior Offender/Victim Relationship on Offender Stalking Behavior. *Journal of Investigative Psychology and Offender Profiling, 1*, 153-167.

Guttman, L. (1959). A Structural Theory for Intergroup Beliefs and Action. *American Sociological Review, 24*, 318-328.

Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika, 33*, 469-506.

Häkkänen, H., Lindlöf, P., & Santilla, P. (2004). Crime Scene Actions and Offender Characteristics in a Sample of Finnish Stranger Rapes. *Journal of Investigative Psychology and Offender Profiling, 1*, 17-32.

Halverson, A.M., Hallahan, M., Hart, A.J., & Rosenthal, R. (1997). Reducing the biasing effects of judges' nonverbal behavior with simplified jury instruction. *Journal of Applied Psychology, 82,* 590-598.

Hare, R.D. (2003). *The Hare Psychopathy Checklist Revised* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Harley, E.M., Carlsen, K.A., & Loftus, G.R. (2004). The "saw-it-all-along" effect: Demonstrations of visual hindsight bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 960-968.

Harris, M.J., & Rosenthal, R. (1985). Mediation of Interpersonal Expectancy Effects: 31 Meta-Analyses. *Psychological Bulletin, 97,* 363-386.

Harris, M.J., & Rosenthal, R. (1986). Four factors in the mediation of teacher expectancy effects. In R.S. Feldman (Ed.), *The social psychology of education* (pp. 91-114). New York: Cambridge University Press.

Hart, A.J. (1995). Naturally occurring expectation effects. *Journal of Personality and Social Psychology, 68,* 109-115.

Haw, R.M., & Fisher, R.P. (2004). Effects of Administrator-Witness Contact on Eyewitness Identification Accuracy. *Journal of Applied Psychology, 89,* 1106-1112.

Home Office. (2010). Police and Criminal Evidence Act 1984: Code D: Code of practice for the identification of persons by police officers. London: The Stationery Office. Retrieved August 19, 2013, from http://www.officialdocuments.gov.uk/document/other/9789999106344/9789999106344.pdf

Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179-185.

Hosch, H.M., & Cooper, D.S. (1982). Victimization as a determinant of eyewitness accuracy. *Journal of Applied Cognitive Psychology, 67,* 649-652.

Howells, T.H. (1938). A study of ability to recognize faces. *Journal of Abnormal Psychology, 33,* 124-127.

Howitt, D. (2006). *Introduction to Forensic and Criminal Psychology* (2nd ed.). Essex: Pearson Education Limited.

Hurley, J.R. (1990). Does FIRO-B relate better to interpersonal or intrapersonal behavior? *Journal of Clinical Psychology, 46,* 454-466.

Hurley, J.R. (1991). FIRO-B's dissociation from two central dimensions of interpersonal behavior. *Psychological Reports, 68,* 243-254.

Hurley, J.R. (1992). Further evidence against the construct validity of the FIRO-B scales. *Psychological Reports, 70,* 639-649.

Innocence Project (2013). Downloaded 04 March 2013 from http://www.innocence project. org/understand/Eyewitness-Misidentification.php

International Research Centre of Investigative Psychology (IRCIP). (2010). *The Facet Structure of the FIRO-B Model of Interpersonal Personality.* Unpublished study on file with the author, IRCIP, Huddersfield.

Jamieson, A. (2004). A rational approach to the principles and practice of crime scene investigation. *Science and Justice, 44,* 3-7.

Jones, E.E. (1991). *Interpersonal Perception.* New York: W.H. Freeman and Company.

Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika, 39,* 31-36.

Karras, M., McCarron, E., Gray, A., & Ardasinski, S. (2006). *On the edge of justice: The legal needs of people with a mental illness in NSW.* Sydney: Law and Justice Foundation of NSW.

Kassin, S.M., Ellsworth, P.C., & Smith, V.L. (1989). The "general acceptance" of psychological research on eyewitness testimony: A survey of experts. *American Psychologist, 44*, 1089-1098.

Kebbell, M.R., & Hatton, C. (1999). People with mental retardation as witnesses in court: A review. *Mental Retardation, 37,* 179-187.

Kebbell, M.R., & Milne, R. (1998). Police officers' perceptions of eyewitness performance in forensic investigations. *Journal of Social Psychology, 138,* 323-330.

Kebbell, M.R., & Wagstaff, G.F. (1999). *Face value: Evaluating the accuracy of eyewitness information.* London: Home Office.

Kohnken, G., & Maass, A. (1988). Eyewitness testimony: False alarms on biased instructions? *Journal of Applied Psychology, 73,* 363-370.

Koppl, R., & Sacks, M. (2013). The Criminal Justice System Creates Incentives for False Convictions. *Criminal Justice Ethics, 32,* 126-162.

Krane, D.E., Ford, S., Gilder, J.R., Inman, K., Jamieson, A., Koppl, R., ... Thompson, W.C. (2008). Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation. *Journal of Forensic Science, 53,* 1006-1007.

Lab mouse test results depend on scientist's gender. (2014, April 29). *CBC News.* Retrieved April 30, 2014, from http://www.cbc.ca/news/technology/lab-mouse-test-results-depend-on-scientist-s-gender-1.2625550

Lindsay, R.C.L., & Wells, G.L. (1980). What price justice? Exploring the relationship of lineup fairness to identification accuracy. *Law and Human Behavior, 4,* 303-314.

Lindsay, R.C.L., & Wells, G.L. (1985). Improving eyewitness identifications form lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70,* 556-564.

Lindsay, R.C.L., Lea, J.A., & Fulford, J.A. (1991). Sequential lineup presentation: Technique matters. *Journal of Applied Psychology, 76,* 741-745.

Lindsay, R.C.L., Wallbridge, H., & Drennan, D. (1987). Do the clothes make the man? An exploration of the effect of lineup attire on eyewitness identification accuracy. *Canadian Journal of Behavioural Science, 19,* 463-478.

Lindsay, R.C.L., Lea, J.A., Nosworthy, G.J., Fulford, J.A., Hector, J., Le Van, V., & Seabrook, C. (1991). Biased lineups: Sequential presentation reduces the problem. *Journal of Applied Psychology, 76,* 796-802.

Lingoes, J.C. (1973). *The Guttman-Lingoes Non-Metric Program Series*. Ann Arbor, Michigan: Mathesis Press.

Loftus, E.F. (1976). Unconscious transference. *Law and Psychology Review, 2,* 93-98.

Loftus, E.F. (1996). *Eyewitness Testimony*. Cambridge, Mass: Harvard University Press.

Loftus, E.F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning and Memory, 12,* 361-366.

Loftus, E.F., & Ketcham, K. (1991). *Witness for the defense: The accused, the eyewitness, and the expert who puts memory on trial.* New York: St. Martin's Press.

Loftus, E.F., & Palmer, J.C. (1974). Reconstructions of automobile destruction: an example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 585-589.

Loftus, E.F., Loftus, G.R., & Messo, J. (1987). Some facts about "weapon focus." *Law and Human Behaviour, 11,* 55-62.

Loftus, E.F., Miller, D.G., & Burns, H.J. (1978). Semantic integration of verbal information into visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 19-31.

Loftus, G.R., & Harley, E.M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin and Review, 12, 43-65.*

Luus, C.A.E., & Wells, G.L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behaviour, 15,* 43-57.

Luus, C.A.E., & Wells, G.L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology, 79,* 714-724.

McNemar, Q. (1946). Opinion-Attitude Methodology. *Psychological Bulletin, 43,* 289-374.

Macrosson, W.D.K. (2000). FIRO-B: Factors and Facets. *Psychological Reports, 86,* 311-320.

Mahoney, J.M., & Stasson, M.F. (2005). Interpersonal and Personality Dimensions of Behavior: FIRO-B and the Big Five. *North American Journal of Psychology, 7,* 205-216.

Malpass, R.S., & Devine, P.G. (1981). Eyewitness Identification: Lineup Instructions and the Absence of the Offender. *Journal of Applied Psychology, 66,* 482-489.

Malpass, R.S., Tredoux, C.G., & McQuiston-Surrett, D. (2006). Lineup Construction and Lineup Fairness. In R. Lindsay, D. Ross, J.D. Read, & M.P. Toglia (Eds.), *Handbook of Eyewitness Psychology (Vol. 2): Memory for People.* London: Lawrence Erlbaum Associates.

Mecklenburg, S.H. (2006, March 17). Report to the Illinois Legislature: Illinois Pilot Program on Sequential Double-Blind Lineup Procedures. Retrieved April 3, 2013, from http://www.chicagopolice.org/IL%20Pilot%20on%20Eyewitness%20ID.pdf

Mecklenburg, S.H., Bailey, P.J., & Larson, M.R. (2008). The Illinois Field Study: A Significant Contribution to Understanding Real World Eyewitness Identification Issues. *Law and Human Behavior, 32,* 22-27.

Meissner, C., & Brigham, J.C. (2001). Twenty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology, Public Policy and Law, 7,* 3-35.

Memon, A., Gabbert, F., & Hope, L. (2004). The ageing eyewitness. In J.R. Adler (Ed.), *Forensic Psychology: Concepts, Debates and Practice* (pp. 96-112). Devon: Willan Publishing.

Merton, R.K. (1948). The self-fulfilling prophecy. *The Antioch Review, 8,* 193-210.

Mickes, L., Flowe, H.D., & Wixted, J.T. (2012). Receiver Operating Characteristic Analysis of Eyewitness Memory: Comparing the Diagnostic Accuracy of Simultaneous vs. Sequential Lineups. *Journal of Experimental Psychology, 18,* 361-376.

Milgram, S. (1974). *Obedience to Authority: An Experimental View.* New York: Harper & Row.

Mitchell, K.J., Johnson, M.K., & Mather, M. (2003). Source monitoring and suggestibility to misinformation: Adult age-related differences. *Applied Cognitive Psychology, 17,* 107-119.

Morgan, C.A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baranoski, M., & Southwick, S. (2004). Accuracy of eyewitness memory for persons encountered during exposure to high intense stress. *International Journal of Law and Psychiatry, 27,* 265-279.

Murrie, D.C., Boccaccini, M.T., Guarnera, L.A., Rufino, K.A. (2013). Are Forensic Experts Biased by the Side That Retained Them? *Psychological Science.* Retrieved August 25, 2013, from http://pss.sagepub.com/content/early/2013/08/21/0956797613481812

Murrie, D.C., Boccaccini, M.T., Johnson, J.T., Janke, J. (2008). Does interrater (dis)agreement on Psychopathy Checklist scores in sexually violent predator trials suggest partisan allegiance in forensic evaluations? *Law and Human Behavior, 32,* 352-362.

Murrie, D.C., Boccaccini, M.T., Turner, D., Meeks, M., Woods, C., & Tussey, C. (2009). Rater (dis)agreement on risk assessment measures in sexually violent predator proceedings: Evidence of adversarial allegiance in forensic evaluations? *Psychology, Public Policy, and Law, 15,* 19-53.

National Academy of Sciences. (2009). Strengthening forensic science in the United States: a path forward. Washington D.C: National Academies Press.

Odinot, G., Wolters, G. & Lavender, T. (2009). Repeated partial eyewitness questioning causes confidence inflation but not retrieval-induced forgetting. *Applied Cognitive Psychology, 23,* 90-97.

Orne, M.T. (1959). The nature of hypnosis: Artifact and essence. *Journal of Abnormal and Social Psychology, 58,* 277-299.

Orne, M.T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17,* 776-783.

Orne, M.T. (1969). Demand characteristics and the concept of quasi-controls. In R. Rosenthal & R. Rosnow (Eds.), *Artifact in behavioral research* (pp. 143-179). New York: Academic Press.

Orne, M. T., &, Evans, F. J. (1965). Social control in the psychological experiment: Antisocial behavior and hypnosis. *Journal of Personality and Social Psychology, 1,* 189-200.

O'Rourke, T.E., Penrod, S.D., Cutler, B.L., Stuve, T.E. (1989). The external validity of eyewitness identification research: Generalizing across subject populations. *Law and Human Behavior, 13,* 385-397.

Pallant, J. (2007). *SPSS Survival Manual* (3rd Ed.). Maidenhead, UK: Open University Press.

Patterson, K.E., & Baddeley, A.D. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 406-407.

Peat, J. (2001). *Health science research: A handbook of quantitative methods.* Sydney: Allen & Unwin.

Pfungst, O. (1911). *Clever Hans (The Horse of Mr. Von Osten).* New York: Henry Holt.

Phillips, M.R., McAuliff, B.D., Kovera, M.B., & Cutler, B.L. (1999). Double-Blind Photoarray Administration as a Safeguard Against Investigator Bias. *Journal of Applied Psychology, 84,* 940-951.

Pichert, J.W., & Anderson, R.C. (1977). Taking Different Perspectives on a Story. *Journal of Educational Psychology, 69,* 309-310.

Pozzulo, J.D. (2006). Person description and identification by child eyewitnesses. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Handbook of eyewitness psychology: Memory for people* (pp. 283-308). Mahwah, NJ: Lawrence Erlbaum Associates.

Pozzulo, J.D., & Lindsay, R.C.L. (1998). Identification accuracy of children versus adults: A meta-analysis. *Law and Human Behaviour, 22,* 549-570.

R. v. Turnball, 98 Cr. App. R. 313 (1977).

Raudenbush, S.W. (1984). Magnitude of Teacher Expectancy effects on Pupil IQ as a Function of the Credibility of Expectancy Induction: A Synthesis of Findings from 18 Experiments. *Journal of Educational Psychology, 76,* 85-97.

Risinger, D.M., Saks, M.J., Thompson, W.C., & Rosenthal, R. (2002). The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion. *California Law Review, 90,* 1-56.

Roberts, A. (2009). Eyewitness Identification Evidence: Procedural Developments and the Ends of Adjudicative Accuracy. *International Commentary on Evidence, 6.* Retrieved March 20, 2013, from http://wrap.warwick.ac.uk/887/1/WRAP_Roberts_eyewitness.pdf

Roethlisberger, F.J., & Dickson, W.J. (1939). *Management and the Worker.* Cambridge, Mass: Harvard University Press.

Rosenberg, M.J. (1965). When dissonance fails: On eliminating evaluation apprehension from attitude measurement. *Journal of Personality and Social Psychology, 1,* 28-42.

Rosenberg, M.J. (1969). The Conditions and Consequences of Evaluation Apprehension. In R. Rosenthal, & R.L. Rosnow (Eds.), *Artifact in Behavioral Research* (pp. 279-349). New York: Academic Press.

Rosenberg, M.J. (1980). Experimenter expectancy, evaluation apprehension, and the diffusion of methodological angst. *Behavioral and Brain Sciences, 3,* 472-474.

Rosenthal, R. (1966). *Experimenter Effects in Behavioral Research.* New York: Appleton-Century-Crofts.

Rosenthal, R. (1967). Covert communication in the psychology experiment. *Psychological Bulletin, 67,* 356-367.

Rosenthal, R. (1976). *Experimenter Effects in Behavioral Research* (2nd Ed.). USA: Irvington Publishers Inc.

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science, 3,* 176-179.

Rosenthal, R. (2003). Covert Communication in Laboratories, Classroom, and the Truly Real World. *Current Directions in Psychological Science, 12,* 151-154.

Rosenthal, R., & Fode, K.L. (1961). The problem of experimenter outcome-bias. In D.P. Ray (Ed.), *Series research in social psychology.* Symposia series No. 8, Washington, D.C.: National Institute of Social and Behavioral Science.

Rosenthal, R., & Fode, K.L. (1963a). The effect of experimenter bias on the performance of the albino rat. *Behavioral Science, 8,* 183-189.

Rosenthal, R., & Fode, K.L. (1963b). Three experiments in experimenter bias. *Psychological Reports, 12,* 491-511.

Rosenthal, R., & Jacobson, L. (1992). *Pygmalion in the Classroom: Teacher Expectation and Pupils' Intellectual Development* (2nd ed.). New York: Irvington Publishers Inc.

Rosenthal, R., & Lawson, R. (1964). A longitudinal study of the effects of experimenter bias on the operant learning of laboratory rats. *Journal of Psychiatric Research, 2,* 61-72.

Rosenthal, R., & Rosnow, R.L. (1975). *The Volunteer Subject.* New York: Wiley.

Rosenthal, R., & Rosnow, R.L. (1991). *Essentials of Behavioral Research: Methods and Data Analysis* (2nd ed.). New York: McGraw-Hill Inc.

Rosenthal, R., & Rubin, D.B. (1978). Interpersonal expectancy effects: the first 345 studies. *The Behavioral and Brain Sciences, 3,* 377-415.

Rosenzweig, S. (1933). The experimental situation as a psychological problem. *Psychological Review, 40,* 337-354.

Rosnow, R.L. & Rosenthal, R. (1997). *People Studying People: Artifacts and ethics in behavioral research.* New York: W.H. Freeman and Company.

Rosnow, R.L., Skleder, A.A., & Rind, B. (1995). Reading Other People: A Hidden Cognitive Structure? *General Psychologist, 31,* 1-10.

Rosnow, R.L., Skleder, A.A., Jaeger, M.E., & Rind, B. (1994). Intelligence and the Epistemics of Interpersonal Acumen: Testing Some Implications of Gardner's Theory. *Intelligence, 19,* 93-116.

Russano, M.B., Dickinson, J.J., Greathouse, S.M., & Kovera, M.B. (2006). "Why don't you take another look at number three?": Investigator knowledge and its effects on eyewitness confidence and identification decisions. *Cardozo Public Law, Policy and Ethics Journal, 4,* 354-380.

Sacks, E.L. (1952). Intelligence scores as a function of experimentally established social relationships between child and examiner. *Journal of Abnormal and Social Psychology, 47,* 354-358.

Saks, M.J., Risinger, D.M., Rosenthal, R., & Thompson, W.C. (2003). Context effects in forensic science: A review and application of the science of science to crime laboratory practice in the United States. *Science and Justice, 43,* 77-90.

Salfati, C.G., & Canter, D. (1999). Differentiating Stranger Murders: Profiling Offender Characteristics from Behavioral Styles. *Behavioral Sciences and the Law, 17*, 391-406.

Sarason, I.G., & Harmatz, M.G. (1965). Test anxiety and experimental conditions. *Journal of Personality and Social Psychology, 1,* 499-505.

Sauer, J., Brewer, N., Zweck, T., & Weber, N. (2010). The Effect of Retention Interval on the Confidence-Accuracy Relationship for Eyewitness Identification. *Law and Human Behavior, 34,* 337-347.

Schacter, D., Dawes, R., Jacoby, L., Kahneman, D., Lempert, R., Roediger, H., & Rosenthal, R. (2008). Policy Forum: Studying eyewitness identifications in the field. *Law and Human Behavior, 32,* 3-5.

Schuler, R.C. (2009). Science, Bias and the Crime Scene Investigation. Retrieved August 17, 2013, from https://onlinecourses.science.psu.edu/frnsc200/sites/onlinecourses.science.psu.edu.frnsc200/files/u3/Chapter_2_Science_Bias_and_Crime_Scene.pdf

Schutz, W. (1958). *FIRO: A three dimensional theory of interpersonal behaviour.* New York: Holt, Rinehart, and Winston.

Schutz, W. (1976). *Leaders of Schools.* San Diego, CA: University Associates.

Schutz, W. (1992). Beyond FIRO-B – Three new theory-derived measures – Element-B: Behaviour, Element-F: Feelings, Element-S: Self. *Psychological Reports, 70,* 915-937.

Schutz, W. (1994). *The Human Element.* San Francisco: Jossey-Bass.

Searcy, J.H., Bartlett, J.C., & Memon, A. (1997). Age differences in accuracy and choosing in eyewitness identification and face recognition. *Memory and Cognition, 27,* 538-552.

Shapiro, P., & Penrod, S. (1986). Meta-Analysis of facial identification studies. *Psychological Bulletin, 100,* 139-156.

Shaw, J.I., & Skolnick, P**.** (1994). Sex differences, weapon focus, and eyewitness reliability. *The Journal of Social Psychology, 134,* 413-420.

Shaw, J.I., & Skolnick, P. (1999). Weapon focus and gender differences in eyewitness accuracy: Arousal versus salience. *Journal of Applied Social Psychology, 29,* 2328-2341.

Shaw, J.S. III. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied, 2,* 126-146.

Shaw, J.S. III, McClure, K.A. (1996). Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law and Human Behavior, 20,* 629-654.

Sheldrake, R. (1998). Experimenter Effects in Scientific Research: How Widely Are They Neglected? *Journal of Scientific Exploration, 12,* 73-78.

Shepherd, J.W., Ellis, H.D., & Davies, G.M. (1982). *Identification Evidence: A Psychological Evaluation.* Aberdeen: Aberdeen University Press.

Shiffman, H.R., & Bobko, D.J. (1975). Effects of stimulus complexity on the perception of brief temporal intervals. *Journal of Experimental Psychology, 103,* 156-159.

Smith, S.M., Lindsay, R.C.L., & Pryke, S. (2000). Postdictors of eyewitness errors: Can false identifications be diagnosed? *Journal of Applied Psychology, 85,* 542-550.

Sporer, S.L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology, 78,* 22-33.

Sporer, S.L., Penrod, D., Read, S., & Cutler, B.L. (1995). Choosing, confidence and accuracy: A meta-analysis of the confidence relation in eyewitness identification studies. *Psychological Bulletin, 118,* 315-327.

Stark, C.E.L., Okado, Y., & Loftus, E.F. (2010). Imaging the reconstruction of true and false memories using sensory reactivation and the misinformation paradigms. *Learning and Memory, 17,* 485-488.

Steblay, N.M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior, 16,* 413-424.

Steblay, N.M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior, 21,* 283-297.

Steblay, N.M., Dysart, J., Fulero, S., & Lindsay, R.C.L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior, 25,* 459-473.

Steblay, N.M., Dysart, J., Fulero, S., & Lindsay, R.C.L. (2003). Eyewitness accuracy rates in police showup and lineup presentations: A meta-analytic comparison. *Law and Human Behavior, 27,* 523-540.

Stevenson, H.W., Keen, R., & Knights, R.M. (1963). Parents and strangers as reinforcing agents for children's performance. *Journal of Abnormal and Social Psychology, 67,* 183-186.

Tabachnick, B.G., & Fidell, L.S. (2007). *Using Multivariate Statistics* (5th ed.). Boston: Pearson Education.

Tollestrup, P., Turtle, J., & Yuille, J. (1994). Actual victims and witnesses to robbery and fraud: An archival analysis. In D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 144-160). New York: Cambridge University Press.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327-352.

United States v. Wade, 388 U.S. 218 (1967).

Valentine, E.R. (1992). *Conceptual Issues in Psychology* (2nd ed.). London: Routledge.

Valentine, T. (2008, December 19). Report on the eyewitness evidence in the case of al Megrahi v HMA. Retrieved November 28, 2012, from http://www.megrahimystory.net /downloads/VMA%20report%20al%20Megrahi%20vs%20%20HMA%2019-12-2008.pdf

Valentine, T., & Heaton, P. (1999). An evaluation of the fairness of police line-ups and video identifications. *Applied Cognition Psychology, 13,* S59-S72.

Valentine, T., Darling, S., & Memon, A. (2007). How can psychological science enhance the effectiveness of identification procedures? An international comparison. *Public Interest Law Reporter, 11,* 21-39.

Van Avermaet, E. (2001). Social influence in small groups. In M. Hewstone & W. Stroebe (Eds.), *Introduction to Social Psychology* (pp. 403-443). Oxford: Blackwell Publishing Ltd.

Watkins, M.W. (2000). *Monte Carlo PCA for parallel analysis* (computer software). State College, PA: Ed & Psych Associates.

Weber, N., Brewer, N., & Wells, G.L., Semmler, C., & Keast, A. (2004). Eyewitness identification accuracy and response latency: the unruly 10-12 second rule. *Journal of Experimental Psychology: Applied, 10,* 139-147.

Weingardt, K.R., Toland, H.K., & Loftus, E.F. (1994). Reports of Suggested Memories: Do People Truly Believe Them? In D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Adult Eyewitness Testimony: Current Trends and Developments.* Cambridge: Cambridge University Press.

Weinstein, D. (1979). Fraud in science. *Social Science Quarterly, 59,* 639-652.

Wells, G.L. (1978). Applied eyewitness testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36,* 1546-1557.

Wells, G.L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology, 14,* 89-103.

Wells, G.L. (1988). *Eyewitness identification: A system handbook.* Toronto: Carswell Legal Publications.

Wells, G.L. (1993). What do we know about eyewitness identification? *American Psychologist, 48,* 553-571.

Wells, G.L., & Bradfield, A.L. (1998). "Good you identified the suspect": Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology, 83,* 360-376.

Wells, G.L., & Bradfield, A.L. (1999). Distortions in eyewitness' recollections: Can the postidentification-feedback effect be moderated? *Psychological Science, 10,* 138-144.

Wells, G.L., & Hryciw, B. (1984). Memory for faces: Encoding and retrieval operations. *Memory and Cognition, 12,* 338-344.

Wells, G.L., & Lindsay, R.C.L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin, 88,* 776-784.

Wells, G.L., & Luus, C.A.E. (1990). Police Lineups as Experiments: Social Methodology as a Framework for Properly Conducted Lineups. *Personality and Social Psychology Bulletin, 16,* 106-117.

Wells, G.L., & Murray, D.M. (1983). What can psychology say about the Neil vs. Biggers criteria for judging eyewitness identification accuracy? *Journal of Applied Psychology, 68,* 347-362.

Wells, G.L., & Olson, E.A. (2003). Eyewitness Testimony. *Annual Review of Psychology, 54,* 277-295.

Wells, G.L., & Quinlivan, D.S. (2009). Suggestive Eyewitness Identification Procedures and the Supreme Court's Reliability Test in Light of Eyewitness Science: 30 Years Later. *Law and Human Behaviour, 33,* 1-24.

Wells, G.L., Rydell, S., Seelau, E. (1993). On the selection of distracters for eyewitness lineups. *Journal of Applied Psychology, 78,* 835-844.

Wells, G.L., Seelau, E., Rydell, S., & Luus, C.A.E. (1994). Recommendations for conducting lineups. In D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments,* pp. 223-244. New York: Cambridge University Press.

Wells, G.L., Malpass, R.S., Lindsay, R.C.L., Fisher, R.P., Turtle, J.W., & Fulero, S.M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist, 55,* 581-598.

Wells, G.L., Small, M., Penrod, S., Malpass, R.S., Fulero, S.M., & Brimacombe, C.A.E. (1998). Eyewitness Identification Procedures: Recommendations for Lineups and Photospreads. *Law and Human Behaviour, 22,* 1-39.

Wixted, J.T. & Ebbesen, E.B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory and Cognition, 25*, 731-739.

Wogalter, M.S., Malpass, R.S., & McQuiston, D.E. (2004). A national survey of U.S. police on preparation and conduct of identification lineups. *Psychology, Crime, & Law, 10,* 69-82.

Wogalter, M.S., Marwitz, D.B., & Leonard, D.C. (1992). Suggestiveness in photospread line-ups: Similarity induces distinctiveness. *Applied Cognitive Psychology, 6,* 443-453.

Wright, D.B., & McDaid, A.T. (1996). Comparing system and estimator variables using data from real line-ups. *Applied Cognitive Psychology, 10,* 75-84.

Wright, D.B., & Skagerberg, E.M. (2007). Post-identification feedback affects real eyewitnesses. *Psychological Science, 18,* 172-178.

Wright, D.B., Carlucci, M.E., Evans, J.R., & Compo. N.S. (2010). Turning a Blind Eye to Double Blind Line-Ups. *Applied Cognitive Psychology, 24,* 849-867.

Yarmey, A.D. (1996). The Elderly Witness. In S.L. Sporer, R.S. Malpass, & G. Koehnken (Eds.), *Psychological Issues in Eyewitness Identification* (pp. 259-278). New Jersey: Lawrence Erlbaum Associates.

Yarmey, A.D., Jones, H.T., & Rashid, S. (1984). Eyewitness memory of elderly and young adults. In D.J. Muller, D.E. Blackman, & A.J. Chapman (Eds.), *Psychology and Law* (pp. 215-228). Chichester: John Wiley and Sons.

Yarmey, A.D., Yarmey, M.J., & Yarmey, A.L. (1996). Accuracy identifications in showups and lineups. *Law and Human Behavior, 20,* 459-477.

Yuille, J.C. (1993). We must study forensic eyewitnesses to know about them. *American Psychologist, 48,* 572-573.

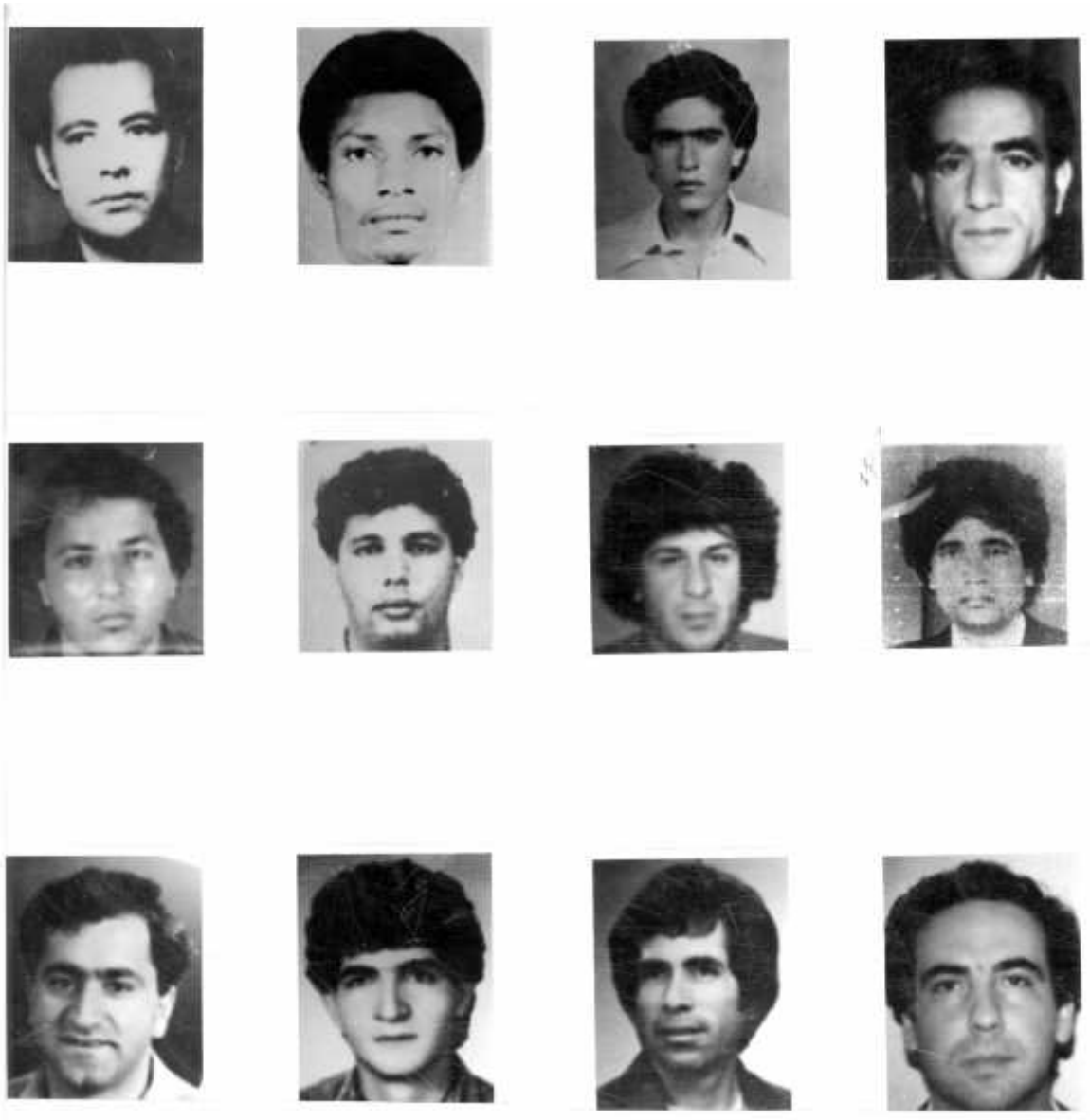
Zhu, B., Chen, C., Loftus, E.F., Lin, C., He, Q., Chen, C., ... Dong, Q. (2010a). Individual differences in false memory from misinformation: Cognitive factors. *Memory, 18,* 543-555.

Zhu, B., Chen, C. Loftus, E.F., Lin, C., He, Q., Chen, C., ... Dong, Q. (2010b). Individual differences in false memory from misinformation: Personality characteristics and their interactions with cognitive abilities. *Personality and individual Differences, 48,* 889-894.

Zwick, W.R., & Velicer, W.F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432-442.

Appendices

Appendix 1: Photo line-up containing Mr Megrahi's picture shown to Mr Gauci on the 15th February 1991



Appendix 2: SREP approval certificate

University of
HUDDERSFIELD
Inspiring tomorrow's professionals

10 July 2012

Ms Jemma Hollinshead
Research Student
School of Human and Health Sciences
University of Huddersfield

Dear Jemma

School Research Ethics Panel (SREP) Submission
Title of Study: 'Uncovering the interpersonal nature of photographic identification line-ups'

I confirm that your project as titled above has received ethical approval from the School of Human and Health Sciences Research Ethics Panel, University of Huddersfield.

I also confirm that indemnity for this project will be covered by the insurance policy held by the University of Huddersfield, as it falls within the normal range of research activity.

With best wishes for the success of your research.

Yours sincerely

K. Ousey

Dr Karen Ousey
Deputy Chair, School Research Ethics Panel (SREP)
School of Human and Health Sciences

Direct Tel: +44 (0)1484 473462
Email: K.J.Ousey@hud.ac.uk

Queensgate Huddersfield HD1 3DH UK Telephone +44 (0) 1484 422288 Fax +44 (0) 1484 516151
Vice-Chancellor: Professor Bob Cryan BSc MBA PhD DSc
An exempt charity and a centre of excellence for vocational education and research

Appendix 3: Consent form

Informed Consent for Participation in Research
International Centre for Investigative Psychology
University of Huddersfield
Project Title: Eyewitness Testimony

Researcher:

As part of my postgraduate studies at the University of Huddersfield I am conducting a study on the effects of different types of line-ups on eyewitness testimony. I am speaking to a random selection of people for this. I would therefore be grateful it you could take part in an experiment for me.

You will just be asked to look at different line-ups then complete a personality questionnaire.

As I am also interested in how people choose other people from line-ups I will be recording the line-up.

It is completely anonymous. Neither your name nor any other identifying details will be recorded in connection with your responses. It should not take any more than 10 minutes.

By taking part in the experiment you have consented to be in the study. Participation is voluntary. You may decide not to participate in this study and if you begin participation you may still decide to stop and withdraw at any time. Your decision will be respected, but please be aware that if you choose to withdraw or omit information, we cannot use any of your answers for analysis.

Please tick the boxes if you are happy to take part in the study

I agree to take part in the experiment ☐

I consent for my responses to be used in further analysis and research ☐

I agree for the experiment to be recorded ☐

A copy of this form will be given to you to retain for future reference.

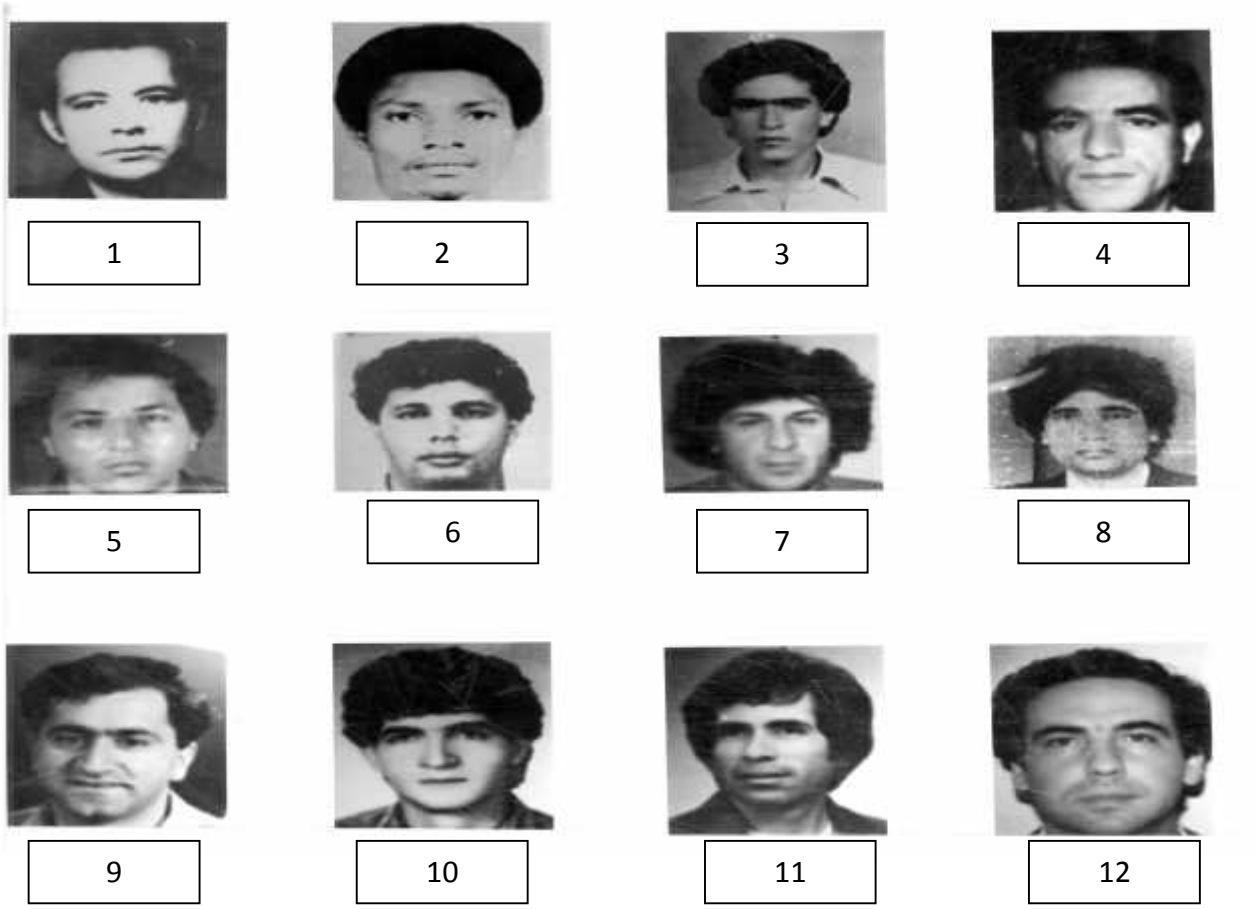
If you have been affected by any of the issues raised in the research and would like some free confidential advice or someone to listen to, the following organisations and support services are available to you:

For confidential advice and support:

Tel: 01484 472227
E-mail: internalcounsel@hud.ac.uk

This project is being carried out under the supervision of Jemma Hollinshead. If you have any comments or questions about the study please contact her at u0972880@hud.ac.uk

Appendix 4: Photo-array X1 and X2



Appendix 5: Photo-array Y1 and Y2

1

2

3

4

5

6

7

8

9

10

11

12

Appendix 6: Respondent answer sheet

Participant code - _____

Experimenter name - _____

Administrator name - _____

Respondent condition –

Age – _____

Gender – M /F (Please circle)

Occupation – _____

Highest level of education – _____

Have you ever been asked to identify a suspect from a police line-up?

Yes / No (please circle)

How confident are you that you have identified the correct person? (Please circle)

Very unconfident 1 2 3 4 5 Very confident

<u>First Condition</u>

Identification - _____

<u>Second Condition</u>

First identification - _____

Second identification - _____

Appendix 7: The FIRO questionnaire

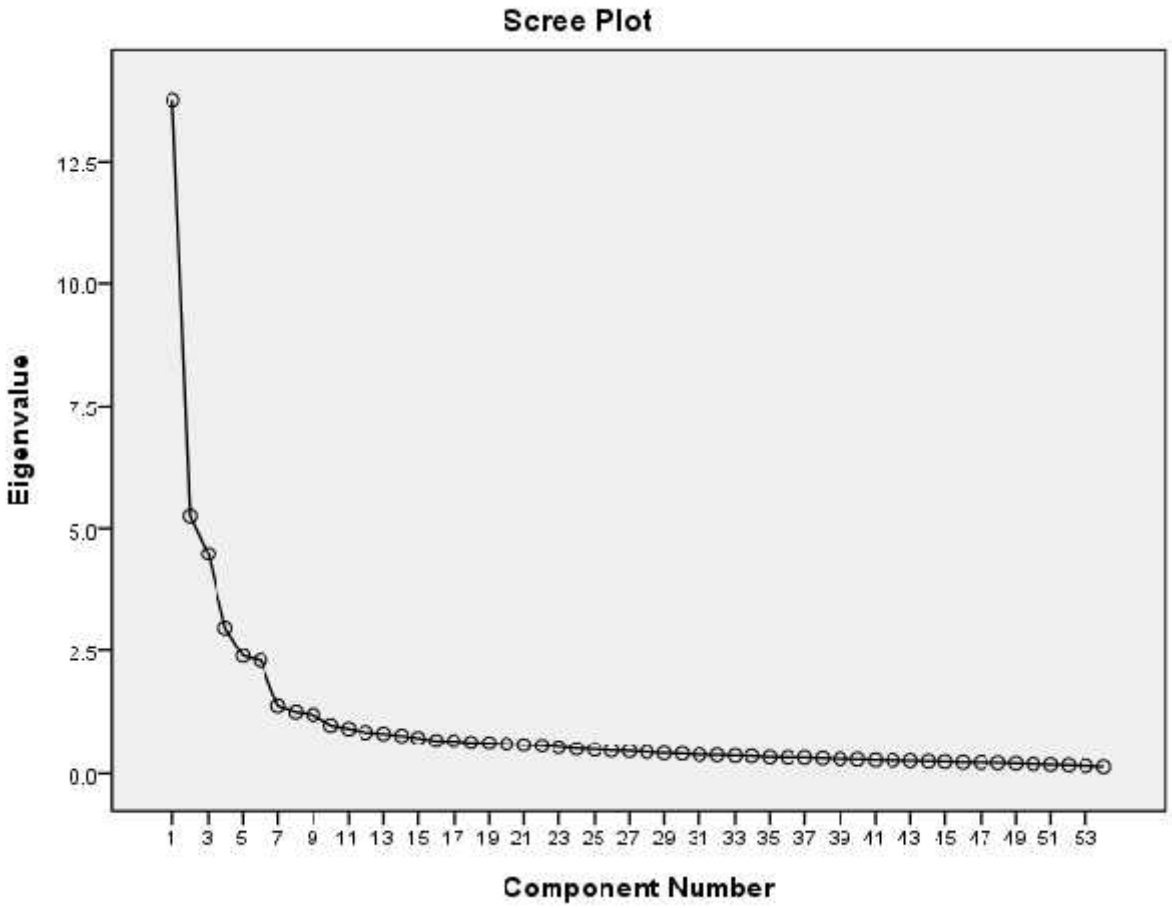
Below is a list of some different ways of behaving towards others. Read each statement and put an X in one of the 6 boxes to show how much you agree that the statement is *true.* The more you agree it is true, the nearer your X should be to the AGREE side.

1. I seek out people to be with.	DISAGREE							AGREE
2. People decide what to do when we are together.	DISAGREE							AGREE
3. I am totally honest with my close friends.	DISAGREE							AGREE
4. People invite me to do things.	DISAGREE							AGREE
5. I am the dominant person when I am with people.	DISAGREE							AGREE
6. My close friends tell me their real feelings.	DISAGREE							AGREE
7. I join social groups.	DISAGREE							AGREE
8. People strongly influence my actions.	DISAGREE							AGREE
9. I confide in my close friends.	DISAGREE							AGREE
10. People invite me to join their activities.	DISAGREE							AGREE
11. I get other people to do things I want done.	DISAGREE							AGREE
12. My close friends tell me about private matters.	DISAGREE							AGREE
13. I join social organisations.	DISAGREE							AGREE
14. People control my actions.	DISAGREE							AGREE
15. I am more comfortable when people do not get too close.	DISAGREE							AGREE
16. People include me in their activities.	DISAGREE							AGREE
17. I strongly influence other people's actions.	DISAGREE							AGREE

18. My close friends do not tell me about themselves.	DISAGREE							AGREE
19. I am included in informal social activities.	DISAGREE							AGREE
20. I am easily led by people.	DISAGREE							AGREE
21. People should keep their private feelings to themselves.	DISAGREE							AGREE
22. People invite me to participate in their activities.	DISAGREE							AGREE
23. I take charge when I am with people socially.	DISAGREE							AGREE
24. My close friends let me know their real feelings.	DISAGREE							AGREE
25. I include other people in my plans.	DISAGREE							AGREE
26. People decide things for me.	DISAGREE							AGREE
27. There are some things I do not tell anyone.	DISAGREE							AGREE
28. People include me in their social affairs.	DISAGREE							AGREE
29. I get people to do things the way I want them done.	DISAGREE							AGREE
30. My closest friends keep secrets from me.	DISAGREE							AGREE
31. I have people around me.	DISAGREE							AGREE
32. People strongly influence my ideas.	DISAGREE							AGREE
33. There are some things I would not tell anyone.	DISAGREE							AGREE
34. People ask me to participate in their discussions.	DISAGREE							AGREE
35. I take charge when I am with people.	DISAGREE							AGREE
36. My friends confide in me.	DISAGREE							AGREE

37. When people are doing things together I join them.	DISAGREE							AGREE
38. I am strongly influenced by what people say.	DISAGREE							AGREE
39. I have at least one friend to whom I can tell anything.	DISAGREE							AGREE
40. People invite me to parties.	DISAGREE							AGREE
41. I strongly influence other people`s ideas.	DISAGREE							AGREE
42. My close friends keep their feelings a secret from me.	DISAGREE							AGREE
43. I look for people to be with.	DISAGREE							AGREE
44. Other people take charge when we work together.	DISAGREE							AGREE
45. There is a part of myself I keep private.	DISAGREE							AGREE
46. People invite me to join them when we have free time.	DISAGREE							AGREE
47. I take charge when I work with people.	DISAGREE							AGREE
48. At least two of my friends tell me their true feelings.	DISAGREE							AGREE
49. I participate in group activities.	DISAGREE							AGREE
50. People often cause me to change my mind.	DISAGREE							AGREE
51. I have close relationships with a few people.	DISAGREE							AGREE
52. People invite me to do things with them.	DISAGREE							AGREE
53. I see to it that people do things the way I want them to.	DISAGREE							AGREE
54. My friends tell me about their private lives.	DISAGREE							AGREE

Appendix 8: The Screeplot of the Principal Components Analysis for the FIRO-B data



Appendix 9: Pattern and structure matrix for PCA with oblimin rotation of six factor solution of FIRO items

Item	Pattern Coefficients						Structure Coefficients					
	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6
49	**.819**	-.040	.009	-.020	.070	.019	**.810**	.103	-.002	-.062	.140	-.358
22	**.756**	.013	-.054	-.131	-.023	-.133	**.824**	.162	-.056	-.174	.059	-.483
16	**.744**	.028	-.099	-.080	-.027	-.090	**.793**	.171	-.100	-.120	.052	-.434
19	**.739**	-.084	.046	-.242	.010	.063	**.708**	.052	.050	-.277	.072	-.279
10	**.735**	.069	.000	-.057	-.087	-.177	**.824**	.221	.002	-.103	-.014	-.508
7	**.735**	.121	.080	.115	.109	.185	**.676**	.236	.069	.080	.131	-.156
13	**.707**	.058	-.001	.171	.079	.227	**.613**	.166	-.010	.143	.109	-.095
28	**.692**	.063	-.032	-.054	-.105	-.166	**.773**	.208	-.028	-.096	-.034	-.477
52	**.660**	-.038	-.043	-.136	-.041	-.236	**.763**	.099	-.047	-.179	.046	-.538
37	**.632**	.025	.081	.028	.052	-.190	**.726**	.146	.064	-.017	.111	-.479
40	**.625**	-.037	-.008	-.104	-.079	-.309	**.757**	.101	-.011	-.149	.008	-.588
4	**.616**	.062	-.062	.135	-.098	-.220	**.712**	.191	-.066	.094	-.027	-.488
25	**.613**	-.089	.102	.000	.125	-.179	**.688**	.021	.077	-.044	.190	-.461
46	**.589**	-.063	-.074	.120	-.145	-.362	**.723**	.072	-.080	.078	-.050	-.609
34	**.551**	.087	-.117	.213	-.076	-.212	**.646**	.199	-.126	.178	-.008	-.452
31	**.437**	.026	.114	-.074	-.070	-.317	**.583**	.130	.111	-.115	-.017	-.509
35	.019	**.842**	-.078	-.103	.049	-.010	.188	**.842**	-.067	-.118	-.017	-.071
53	.030	**.817**	-.040	.153	-.049	.071	.136	**.820**	-.025	.142	-.127	.030
47	.013	**.790**	-.166	.068	.018	-.061	.184	**.790**	-.160	.055	-.032	-.108
29	.001	**.778**	-.024	.194	-.079	-.026	.139	**.783**	-.011	.180	-.150	-.047
17	-.035	**.756**	.007	-.048	.010	-.010	.111	**.750**	.019	-.061	-.063	-.032
11	.060	**.700**	-.017	.228	-.069	-.015	.177	**.714**	-.008	.213	-.130	-.058
41	-.034	**.675**	.140	-.242	-.112	-.011	.097	**.687**	.170	-.257	-.190	-.023
23	-.007	**.641**	.052	-.555	.003	-.021	.149	**.652**	.079	-.569	-.055	-.073
5	.056	**.610**	-.087	-.480	.122	-.051	.226	**.619**	-.077	-.495	.090	-.143
20	.044	-.120	**.797**	-.093	.103	.082	-.006	-.111	**.790**	-.117	.025	.082
38	-.071	-.042	**.773**	.051	-.023	-.082	-.050	-.038	**.771**	.025	-.103	-.015

Appendix 9: Pattern and structure matrix for PCA with oblimin rotation of six factor solution of FIRO items

Item	Pattern Coefficients						Structure Coefficients					
	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6
50	-.053	.035	**.736**	-.042	-.092	-.097	-.012	.050	**.744**	-.068	-.170	-.042
26	.015	-.040	**.729**	-.104	.092	.201	-.074	-.042	**.729**	-.120	.000	.208
14	-.003	-.007	**.710**	-.114	.063	.149	-.064	-.008	**.712**	-.132	-.027	.165
32	.066	.156	**.603**	.558	-.041	-.066	.089	.174	**.588**	.528	-.117	-.050
44	.066	-.254	**.567**	-.041	-.129	.057	-.019	-.224	**.581**	-.054	-.166	.070
8	.114	.226	**.550**	.524	.079	-.024	.143	.239	**.526**	.493	.002	-.047
2	-.267	-.018	**.359**	.233	-.151	-.280	-.170	-.039	**.359**	.223	-.188	-.120
51	.114	-.152	.103	**-.597**	-.023	-.323	.261	-.101	.111	**-.619**	.028	-.392
1	.057	.135	.276	**-.555**	.091	-.190	.203	.160	.280	**-.581**	.078	-.248
43	.187	.162	.350	**-.398**	.097	-.096	.287	.204	.351	**-.429**	.074	-.204
45	-.020	-.022	.091	-.162	**.814**	.070	.020	-.101	.010	-.171	**.800**	-.001
33	-.080	-.092	.046	-.164	**.811**	-.060	.006	-.176	-.040	-.173	**.815**	-.102
27	-.001	.094	-.020	.456	**.721**	-.012	.058	.018	-.114	.445	**.709**	-.062
15	.117	-.091	-.044	.071	**.455**	-.224	.237	-.104	-.107	.051	**.498**	-.314
24	.091	.034	-.013	-.070	-.044	**-.769**	.446	.091	-.034	-.112	.035	**-.811**
48	.153	.103	-.014	-.105	-.049	**-.710**	.495	.171	-.031	-.149	.024	**-.785**
12	.191	.092	.045	-.064	-.076	**-.705**	.524	.169	.030	-.111	-.007	**-.790**
36	.171	.082	-.023	-.066	-.076	**-.672**	.488	.153	-.036	-.108	-.003	**-.750**
6	.187	.074	.019	-.086	-.090	**-.670**	.502	.150	.008	-.130	-.019	**-.754**
54	.257	.050	.000	-.087	-.091	**-.656**	.560	.138	-.012	-.132	-.011	**-.771**
42	-.137	-.100	-.108	-.199	.197	**-.646**	.165	-.111	-.147	-.221	.270	**-.612**
9	.183	.035	.134	.040	.117	**-.589**	.464	.086	.097	-.005	.169	**-.678**
3	.109	-.011	.045	.201	.106	**-.545**	.352	.021	.006	.166	.160	**-.592**
39	.099	.038	.030	-.111	.053	**-.527**	.355	.078	.009	-.145	.106	**-.584**
30	-.084	-.001	-.194	.062	.253	**-.524**	.172	-.019	-.243	.043	.316	**-.514**
18	.177	-.090	-.208	.128	.052	**-.389**	.336	-.050	-.235	.108	.133	**-.471**
21	.135	-.089	-.138	.302	.279	**-.371**	.295	-.081	-.194	.279	.344	**-.444**

Appendix 10: Pattern and structure matrix for PCA with oblimin rotation of three factor solution of FIRO items

Item	Pattern Coefficients			Structure Coefficients		
	Component 1	Component 2	Component 3	Component 1	Component 2	Component 3
40	**.800**	.015	.030	**.800**	.077	-.018
10	**.789**	.145	.040	**.797**	.207	.008
22	**.784**	.089	-.009	**.791**	.145	-.048
52	**.779**	.020	-.002	**.781**	.077	-.048
46	**.769**	-.009	-.054	**.771**	.040	-.103
54	**.766**	.023	-.006	**.768**	.078	-.051
12	**.752**	.043	.025	**.754**	.101	-.017
28	**.737**	.144	.009	**.747**	.199	-.020
48	**.730**	.048	-.037	**.736**	.097	-.077
16	**.730**	.110	-.058	**.741**	.156	-.090
37	**.719**	.046	.083	**.717**	.108	.043
6	**.719**	.035	.007	**.721**	.089	-.033
24	**.716**	-.036	-.043	**.716**	.011	-.092
49	**.712**	.028	.039	**.712**	.085	-.002
36	**.704**	.040	-.040	**.709**	.086	-.080
25	**.701**	-.086	.102	**.688**	-.023	.048
4	**.690**	.120	-.050	**.702**	.164	-.079
9	**.664**	-.060	.077	**.655**	-.003	.028
31	**.650**	.045	.133	**.645**	.108	.098
19	**.625**	.006	.116	**.618**	.065	.077
34	**.618**	.133	-.123	**.636**	.163	-.146
39	**.546**	-.028	.003	**.544**	.012	-.035
3	**.533**	-.106	-.026	**.527**	-.070	-.073
7	**.504**	.177	.082	**.512**	.224	.072
42	**.453**	-.225	-.158	**.447**	-.211	-.214
18	**.450**	-.123	-.242	**.456**	-.120	-.285
51	**.435**	-.150	.173	**.413**	-.097	.128
13	**.426**	.128	.006	**.435**	.160	-.005
21	**.411**	-.195	-.234	**.411**	-.194	-.283
30	**.376**	-.128	-.282	**.385**	-.135	-.321
15	**.334**	-.222	-.136	**.327**	-.214	-.184
1	**.314**	.097	.307	**.302**	.157	.299
35	.100	**.822**	-.131	.168	**.813**	-.038
53	-.005	**.819**	-.099	.060	**.807**	.000
29	.038	**.772**	-.087	.100	**.764**	.003
47	.101	**.770**	-.233	.172	**.749**	-.147
17	.037	**.734**	-.043	.093	**.732**	.043
11	.072	**.698**	-.078	.128	**.694**	.001
41	.047	**.687**	.140	.088	**.707**	.219

Cont. Appendix 10: Pattern and structure matrix for PCA with oblimin rotation of three factor solution of FIRO items

Item	Pattern Coefficients			Structure Coefficients		
	Component 1	Component 2	Component 3	Component 1	Component 2	Component 3
23	.130	**.645**	.072	.173	**.664**	.141
5	.209	**.593**	-.094	.258	**.597**	-.035
33	.113	**-.311**	-.08	.095	**-.312**	-.125
45	.065	**-.224**	-.030	.051	**-.223**	-.061
20	.026	-.179	**.802**	-.037	-.080	**.778**
38	.028	-.107	**.762**	-.028	-.013	**.748**
50	.062	-.007	**.743**	.015	.087	**.738**
26	-.094	-.081	**.737**	-.146	.001	**.733**
14	-.069	-.048	**.719**	-.118	.034	**.717**
44	.006	-.238	**.616**	-.050	-.164	**.587**
32	.071	.096	**.529**	.045	.165	**.536**
8	.102	.148	**.460**	.084	.211	**.471**
43	.335	.135	**.370**	.322	.204	**.365**
2	-.041	-.075	**.329**	-.067	-.038	**.322**
27	.058	-.115	**-.211**	.063	-.136	**-.228**