# The Effects of Interrogative Pressure in Simulated Forensic Interviews

Allan McGroarty

A thesis submitted in accordance with the regulations governing the award of Degree of Doctor of Philosophy in Psychology.

Department of Psychology
University of Strathclyde
Glasgow

2007

**Acknowledgements**

Dr James S Baxter for his encouragement, guidance, generosity of spirit, and peerless sense of humour.

Gerry McKenna for his dependability and convincing application of interrogative pressure.

Reuben Millward, Jo Molle, Steve Scales, Jamie Cowan and Callum Wells for acting out my crime scenario with enthusiasm.

Participants from the Departments of Psychology and Engineering, and from the Centre for Lifelong Learning for giving up their time for me.

Judith, Ava and Ari for their support and patience.

# Table of Contents

## List of Tables

## General Abstract

The purpose of this thesis was to investigate the effects of interrogative pressure on interviewee behaviour, particularly the tendency to change previous answers during requestioning. The three main studies used an experimental procedure adapted from the Gudjonsson Suggestibility Scales (GSS; Gudjonsson, 1984, 1987b). Two amendments were made to the standard GSS procedure to increase ecological validity: questions asked of participants were minimally rather than overtly leading and were applied to a video presentation of a simulated crime as opposed to a narrative passage. The repeated questions and negative feedback aspects of the GSS procedure were retained. Study 1 examined the effects of two types of feedback and the presence of a second interviewer. Compared with neutral feedback, negative feedback resulted in more response changes, higher state anxiety and higher ratings of interview difficulty. The results also suggest that the presence of a second interviewer, at least one who has a minimal involvement in questioning, is not likely to influence the recall of interviewees. Study 2 investigated the influence on interviewee behaviour of an intervention made by a supportive "ally". A warning communicated by the ally immediately following feedback was found not to reduce response change. Study 3 examined the effects of two types of interviewer behaviour. An abrupt manner during questioning, compared to a friendly manner, resulted in more response changes. Studies 2 and 3 also assessed the influence of levels of interviewee self-esteem on interviewee responding. Low levels of self-esteem were found to be associated with increased self-reported anxiety following questioning and rated difficulty of the interview, but not with increased response change. These findings

indicate that, if not carefully monitored and controlled, critical forms of verbal feedback and negative demeanour of interviewers can operate as suggestive influences which may compromise the reliability of testimony.

# Chapter 1: An Overview of Police Interrogation and Interviewing: Practices and Empirical Research

The interviewing of witnesses, suspects, complainants and victims is a central aspect of forensic investigation often given priority over other means of gathering evidence (Baldwin, 1994; Pearse & Gudjonsson, 1996). Interviewing of suspects in particular seems to offer a direct route to securing a guilty plea through the attainment of a confession (Irving, 1980). Within the UK, the past three decades have been a period of considerable activity regarding the examination of police interviewing and systematic attempts to improve its practice. Prior to this, the police interrogation room constituted a "black box", closed to public scrutiny.

Recent research has revealed what actually happens in forensic interviews and laboratory-based studies have increased understanding of the effects of factors such as custodial and interrogative pressures, the personality of interviewers, as well as physical and psychological vulnerabilities of interviewees (Gudjonsson, 2006). The result of this research has been to the benefit of both interviewees in terms of protecting their rights, and to police interviewers with regard to highlighting the most effective means of gathering reliable evidence and ensuring that interview-based evidence is admissible in court. Hence, there would seem to be much to commend continuing efforts by a range of professionals to seek to understand and improve police interviewing. McKenzie (2001) states this point firmly: "A clear responsibility rests firmly upon criminal justice

educators, forensic psychologists, enlightened senior police officers, and lawyers, both defence and prosecution, to take every possible step to ensure that the dynamics of interviewing and interrogation are understood" (p. 447). The present thesis, based on four experimental studies, is directed towards this objective.

This chapter begins with a discussion of early police interrogation techniques and considers the progression of interrogation as a deliberately coercive, and often physically brutal practice, towards one characterised by psychological manipulation. Much of the initial discussion centres on American police interrogation practice, not only because the bulk of early literature came from the US but, despite differences between the American legal system and legal systems elsewhere, including the UK, American interrogation tactics have been widely influential (Gudjonsson, 2003). Following consideration of some of the psychological factors operative during interrogation which may influence the behaviour of suspects, attention is given to criticisms of American-style interrogation techniques. The chapter then focuses on the movement away from interrogation, particularly in the UK, towards investigative interviewing methods which have sought to apply ethical procedure and psychological knowledge to the information gathering process. Finally, a review of recent published research on real police interviews with suspects and witnesses is presented.

## 1.1 Early history of police interrogation

An historical perspective on police interrogation techniques shows that the methods used over the course of the twentieth century underwent radical change. In the first three

decades of the previous century, interrogation practice was characterised by the infliction of physical pain and mental suffering on criminal suspects towards the objective of extracting information and obtaining confessions. According to Leo (2004), the characteristic features of such methods, known commonly as the "third degree", are that "(a) it typically occurs during custodial detention; (b) it involves the use of physical force and/or psychological duress; and (c) its fundamental purpose is to extort admissions and confessions of guilt from criminal suspects" (p. 52).

To cite an early example, Munsterberg (1908) describes an instance of the third degree following which a young man who had previously denied the murder of a girl, eventually confessed:

"The detectives had taken the shabby young man to the undertaking rooms, led him to the side of the coffin, suddenly whipped back the sheet, exposing the white bruised face, and abruptly demanded, "When did you see her?" He sank to his knees and put his hands over his face; but they dragged him to his feet and ordered him to place his right hand on the forehead of the body. Shuddering, he obeyed, and the next moment again collapsed. The detectives pulled him again to his feet, and fired at him question after question, forcing him to stroke the girl's hair and cheeks; and, evidently without control of his mind, he affirmed all that his torturers asked, and, in his half-demented state, even added details to his untrue story" (pp. 76-77).

Clearly, the treatment of the suspect in this case constitutes a form of inflicted terror. He was eventually acquitted by the jury who had presumably been made aware of the methods used to obtain the confession. The use of corpses, or parts of corpses, to obtain confessions was by no means unusual. As Hopkins (1931) noted:

"A visit to the morgue is a fairly standard piece of police work in murder cases … it is like flipping a coin with two heads. If the sight of a mutilated cadaver causes the suspected person to show any human signs of shock, then the police naturally interpret it as a betrayal of guilt, and feel more warranted in "putting on the screws" after that. On the other hand, if no shock is revealed, that may demonstrate that the suspect is a hardened monster" (p. 257).

Leo (1992, 2004) has delineated the various coercive interrogation strategies that comprise the third degree. The first of these is "direct physical abuse" in the form of beating, clubbing, whipping and punching detainees on various parts of their body. Keedy (1937) cites a case in which a suspect was hanged with a rope from the limb of a tree then whipped until he confessed. As a result of such abuse, suspects would often be rendered unconscious, require medical assistance and be permanently scarred.

A distinction can be made between "direct" and "deniable" physical abuse. In the latter,

some care is taken to prevent lasting physical markings in order that accusations of physical coercion might be later refuted in court. Among other treatments, suspects have been struck with rubber hose (the "gold fish"), pieces of tyre, poked in the ribs with a blackjack, tied to an armchair and struck on the side of the head, winded by blows to the solar plexus (a "fist to the wind"), administered tear gas, and forced to stand for hours on end (Haller, 1976; Leo, 2004).

Leo (2004) makes a further a distinction under the general category of physical force in describing examples of "orchestrated physical abuse". This involved a series of more premeditated, and no less coercive, tactics designed to inflict pain. Examples of this approach include locking suspects in a small cell heated to an extreme temperature (a "Sweat Box"), the "Water Cure" in which water was poured into a suspect's mouth or nostrils to give the sensation of drowning, and the administration of electric shocks, often in a real electric chair. Suspects were also threatened with the "cannon ball". A heavy cannon ball was suspended above the suspect's head and a rope attached to the release mechanism at one end and the suspect's outstretched leg at the other. A false move from the suspect would release the ball which would crush his or her head. Under such a circumstance, suspects would most often feel compelled to confess. Another technique, involving two police officers, was the "Mutt and Jeff' routine, also known as "Good Cop/Bad Cop" in which one interrogator took an accusatory, negative and often violent stance towards the suspect and the other displayed sympathy and support.

Physical abuse of the kinds noted above was also accompanied by a number of psychological pressures. Often, criminal suspects would "disappear" for days on end while police officers questioned them for prolonged periods of time, a practice known as "incommunicado interrogation". An illustration of this method is given by Hopkins (1931). He details the case of Ziang Sung Wan, a Chinese resident of New York accused of murdering three inmates of the Chinese Educational Mission in Washington. Wan was seriously ill with influenza and stomach trouble when two Washington detectives arrived at his home in New York and, without official warrants of any kind, took him back to a Washington hotel room, i.e. not the police station, for interrogation. He was subjected to persistent accusations of guilt and lengthy cross-examination by the Superintendent of Police and one or more detectives. Wan was questioned morning, afternoon, evening and after midnight, and this gruelling schedule of interrogation continued despite his clear ill health. He was held incommunicado for a total of eight days and only formally arrested on the ninth day when, weak and exhausted, he finally confessed to the murder of the three men. Ziang Sung Wan was tried and convicted largely on the basis of his confession and he was sentenced to hang. Sometime later, an appeal taken to the Supreme Court of the United States was upheld and Wan was eventually released. Later still, the Washington police were indicted for unlawful arrest and for using the third degree.

The final subcategories of the third degree described by Leo (1992, 2004) relate to various deprivations and explicit threats of harm. While being held for long periods of

questioning, suspects were often deliberately deprived of basic needs such as food, water, sleep and toilet facilities to further weaken their resistance and induce admissions of guilt. In addition, they were psychologically intimidated by threats of severe injury or death if they did not confess.

## 1.2 The demise of the third degree

In general, although the police did not publicly admit to the use of extreme physical force or psychological duress on suspects (see Fiaschetti (1930), however, for a candid account of the use of third degree methods), it was widely accepted as a normal aspect of investigative work. As such, third degree methods were used on both adult and juvenile suspects, on women as well as men (Hopkins, 1931), and these practices were sanctioned by senior officials. As Haller (1976) points out, representatives of the State's Attorney and police chiefs themselves were sometimes present while the third degree was being administered.

Not surprisingly, third degree practices came to the attention of the popular media and newspaper reports often referred to the use of coercive measures by police officers as these extracts from 1906 illustrate:

"John L. Voss, accused by the police of the murder of his wife and the burning of his home to destroy the evidence of his crime, yesterday admitted to Assistant Chief of Police Schuettler and Inspector George M. Shippy that he had purchased a revolver and

a box of cartridges some time prior to the crime on Sunday morning. The admission, wrung from the prisoner after three days of cross examination, is regarded as important" (quoted in Haller, 1976, p. 319).

"Every police department in this country has its 'sweaters' or inquisitors, and long practice has made them adepts at the art, if so it may be called" (quoted in Haller, 1976, p. 319).

Media reports of police brutality would generally be denied by police officials and it was very difficult for aggrieved citizens to bring charges against policemen. The subculture in police departments was that of an intense group loyalty and, in the event of accusations of malpractice, officers would close ranks and sometimes would harass witnesses for the prosecution, arrest them, or drive them out of town to prevent them testifying against accused police officers (Haller, 1976). Nevertheless, a steady stream of media reports and accusations by former detainees generated a public outcry and the movement opposing third degree practice was to prove irresistible (Leo, 2004).

By present day standards of course, it is ironic that the police, an organised civil force for maintaining order, preventing and detecting crime, and law enforcement did not perceive their responsibilities from a legal perspective. Haller (1976) notes that, up until the mid 1920s in America, the police were not oriented towards the norms of the legal

system. Officers were untrained in the law and they operated in a criminal justice system that placed little emphasis on legal procedure. The use of violence was not officially recorded and, at the time, there were no civil liberties organisations to protect citizens from police violations of individual rights. It was in this loosely regulated climate that the extreme coercion of suspects persisted. As will be noted, however, changes in investigative practice and the eventual demise of the third degree would happen largely in response to legal pressures.

The National Commission of Law Observance and Law Enforcement was introduced by President Herbert Hoover as a means of examining the current state of American policing. In 1931, the commission published the Wickersham Report which was highly critical of policing methods and served as a record of public distaste as can be seen from the statement below. The Wickersham Report was to have far-reaching and irreversible implications for police practice generally and this included methods of interrogation.

"The general failure of the police to detect and arrest criminals guilty of the many murders, spectacular bank, pay-roll, and other hold-ups, and sensational robberies with guns, frequently resulting in the death of the robbed victim, has caused a loss of public confidence in the police of our country."

Another important development came in a landmark case brought to the United States

Supreme Court, Brown v. Mississippi (1936). In a one-day trial, three black defendants were convicted of murder and sentenced to death. It transpired that the confessions were not made voluntarily, but had been extracted by violent means. Specifically, one of the defendants was twice hung from a tree by a deputy sheriff then tied to the tree and whipped. He was later arrested and whipped further until he agreed to confess. The other two defendants (one of them an Ed Brown) were arrested, taken to jail, made to strip then whipped with leather straps until they also confessed. The Supreme Court ruled that there had been "compulsion by torture to extort a confession" and the treatment of the defendants amounted to a "clear denial of due process". A number of subsequent Supreme Court decisions would castigate the police and eventually third degree methods were declared unlawful in the United States (McKenzie, 2001).

## 1.3 The move to psychological interrogation

The public revelation of police misconduct and abuse created a legitimacy crisis in policing which prompted a redefinition and professionalisation of practices. Police leaders openly denounced third degree methods not only on moral and legal grounds, but on the basis that they were ineffective towards the goal of eliciting confessions. Concerned with restoring public attitudes towards policing and enhancing effectiveness, a movement took place away from the random brutality of the third degree and towards "scientific" forms of criminal investigation (Leo, 2004).

As part of this movement, from 1940 in America, there began a steady stream of police interrogation training manuals – among them, Inbau (1942), Mulbar (1951), O'Hara

(1956), and Arther and Caputo (1959). According to Leo (2004), these publications "drew on the rhetoric, symbols, and cultural authority of science to confer legitimacy on controversial police practices" (p. 60). The first manual to appear, *Police Interrogation*, was written by W.R. Kidd, a former police lieutenant. Kidd (1940) strongly rejected third degree tactics, attributing their use to laziness and lack of training in interrogation. He emphasised the gathering of facts about the crime, the recording of interrogations, the use of lie detecting equipment, and highlighted the high standards of investigative practice demonstrated by the Federal Bureau of Investigation, suggesting that police departments should aspire to their professionalism.

Trovillo (1940), a forensic psychologist reviewing Kidd's manual, noted that it advocated the use of various manipulative techniques aimed at securing admissions from suspects. These included: the use of sympathy, flattery, "kidding", face-saving, the suggestions of extenuating circumstances (justification), bluffing, plays on shame and embarrassment, plays on love, jealousy, hate, revenge, appeals to the suspect's self-interest, interrogation in the presence of the victim, and threat of confinement in an asylum. Tellingly, Kidd (1940) admonished, "If you resort to torture, you admit your victim is the better man. When you 'break' a man by torture, he will hate you. If you break him by your intelligence, he will always fear and respect you" (p. 49).

Leo (2004) argues that modern psychological interrogation such as that developed by Kidd and others is "logically intertwined" with the third degree and offers a "functional

alternative" to it. It can be argued that, while Kidd was advocating a change in the outward brutality and means of earlier interrogation practices, he was, to some extent, seeking to maintain some of their psychologically coercive measures. It seems clear also that the drive to force confessions from criminal suspects remained central to the purpose of interrogation. Evidence of a continuing concern about coercive interrogation tactics was to arise again the United States Supreme Court in Miranda v. Arizona (1966). This landmark ruling sought to protect citizens against the possibility of falsely confessing and it substantially impacted police interrogation practices (Leo, 1996a; Leo, 1996b).

## 1.4 The Reid Technique

Among the police interrogation manuals published in the USA, the most influential is *Criminal Interrogation and Confessions*. Originally published under the title *Lie Detection and Criminal Interrogation* (Inbau, 1942, 1948; Inbau & Reid, 1953), the book has since been revised and expanded (Inbau & Reid, 1962, 1967; Inbau, Reid & Buckley, 1986; Inbau, Reid, Buckley & Jayne, 2001). Termed "the bible of the American interrogator" by McKenzie (1994, p. 249), Inbau et al.'s manual has influenced the publication of other police interrogation texts (e.g. Gordon, Fleisher & Weinberg, 2002; Hess, 1997), including a British manual (Walkley, 1987). That the second series of Inbau et al.'s manual is now in its fourth edition after sixty five years in print is testament to its persistent and continuing influence.

Until relatively recently, there was a striking absence of guidance or training in

interrogation for police officers (Bull, 1999) and it was this vacuum of skills and knowledge that prompted the widespread adoption of Inbau and associates' method of interrogation known as the "Reid Technique". The Reid Technique is essentially a set of tactics aimed at weakening the resistance of criminal suspects and eliciting confessions from them. It is characterised by deliberately tough and deceptive strategies. Inbau and Reid (1962) themselves state, "Deceit is inherent in every question asked to the suspect, and in every statement made by the interrogator" (quoted in Leo, 1992).

### 1.4.1 The technique in detail

Inbau et al. (2001) suggest that, before formal interrogation of suspects begins, they should be taken through the Behavioural Analysis Interview (BAI). The BAI, designed to last 30-40 minutes, begins with a series of standard investigative questions to establish the personal details of the suspect and to help build rapport and trust between the suspect and the interviewer. Unlike the ensuing interrogation, the BAI is intended to be dialogical and non-accusatory in nature. It is argued that initial questioning allows investigators to evaluate the suspect's "normal" verbal and nonverbal behaviour. There then follows a series of "behaviour-provoking questions" intended to elicit different verbal and nonverbal responses from truthful and deceptive suspects. Importantly, Inbau et al. state that the BAI provides "objective criteria" on which to base a judgement of the suspect's truthfulness and their innocence or guilt. Where an investigator is "reasonably certain" of a suspect's guilt, it is suggested that interrogation should commence and, preferably, should be conducted by the same investigator.

Inbau et al. (2001) assert the importance of controlling the ecology of the interrogation. The interrogation room should be soundproofed, should contain no objects which might distract the suspect's attention, and should be physically isolated from the rest of the police station in order to deprive the suspect social support. The interrogator should sit close to the suspect to invade his personal space and should maintain eye contact with him.

The first recommended tactic is to directly confront the suspect with evidence of his guilt. Where no evidence exists, false evidence may be presented such as the testimony of an accomplice, witness, or victim. Alternatively, it is stated that incriminating physical evidence such as fingerprints or blood stains have been found and that this places the suspect's guilt beyond doubt. Kassin and McNall (1991) describe this as "maximization, a "hard sell" technique in which the interrogator tries to scare and intimidate the suspect into confessing" (pp. 234-235). The behavioural response of the suspect is then watched intently before the evidence is presented to the suspect a second time.

Inbau et al. advocate that the moral seriousness of the crime be played down through the use of excuses or moral justifications. Inbau and Reid (1962) cite an example of a 50 year old man who had been accused of taking "indecent liberties" with a 10 year old girl. He was told: "This girl is well developed for her age. She probably learned a lot about sex from the boys ... and from movies and TV; ... she may have deliberately tried to

excite you to see what you would do" (quoted in Zimbardo, 1967, p. 452). Further, some of the blame for the offence may be placed on someone else, such as the victim or an accomplice. This enables the suspect to falsely exonerate himself of some of the responsibility for the offence and renders confession more likely. Kassin and McNall (1991) refer to this technique as "minimization".

Inbau et al. state that denials are commonplace in interrogation and that the method with which they are handled is critical. It is argued that allowing a suspect to repeatedly deny an accusation of guilt gives the suspect an unacceptable tactical advantage. In contrast, interrogators can gain the upper hand by interrupting and cutting short attempted denials. Inbau et al. assert that the verbal and non-verbal manner in which a denial is expressed indicates whether the suspect is innocent or guilty.

Further ploys of the technique involve overcoming objections advanced by the suspect, retaining the suspect's attention and handling the suspect's passivity. At this stage, the interrogator directs the suspect's attention to the reason for committing the crime. In doing so, sympathy and understanding are shown and the suspect asked to simply tell the truth. The cumulative effect of the above tactics, which may been employed over several hours, is the weakening of the resistance of the suspect, and sets the scene for the final push towards eliciting a confession.

The next tactic involves presenting two possible incriminating scenarios to the suspect and pressuring them to choose one. For example, "Was this your own idea, or did someone talk you into it?" One alternative implies that the suspect was purposive and calculating in his commission of the crime while the other implies diminished responsibility. Although neither may be true, clearly, there is incentive to opt for the apparently less serious alternative. To accept neither option leaves open the possibility that the more serious scenario is true. If the suspect accepts one of the two scenarios presented to him, the interrogation then focuses on the attainment of a detailed oral confession in front of two witnesses. The final tactic is directed towards eliciting the same confession in writing.

## 1.5 Psychological factors influencing confession

The psychological literature has much to say on the reasons why interrogation techniques such as those advocated by Inbau et al. (2001) are effective in influencing the behaviour of suspects. The purpose of this section is to outline the various pressures that suspects face during interrogation and, with reference to psychological research, to consider the processes which underpin them. As well as discussing the situational aspects of interrogation, the next chapter will also consider a number of personal factors relating to suspects which have the potential to influence their responses.

### 1.5.1 Offender self-reports

Gudjonsson and Petursson (1991) investigated the subjective experiences of suspects who had confessed to a range of violent, sexual, property and other offences during custodial interrogation. Their study was conducted on an Icelandic prison sample.

Gudjonsson and Bownes (1992) replicated the study with a prison sample in Northern Ireland and obtained closely comparable results. (See also Gudjonsson & Sigurdsson, 1999; Sigurdsson & Gudjonsson, 1994). These authors identified three principal reasons for the suspects' confessions.

The first of these was perception of proof, i.e. the suspect's belief that, due to the existence of incriminating evidence against him, there is little point in denying guilt. (As noted previously, a tactic employed in the Reid Technique involves confronting the suspect with evidence of his guilt – either real or false – in order to influence the suspect's perception of proof). Moston, Stephenson and Williamson (1992) investigated the effects of various case characteristics on suspects' decisions to admit or deny involvement in an offence. In line with Gudjonsson and Petursson's findings, they reported that strength of evidence was associated with admissions of guilt and weak evidence was clearly related to denials. A recent experimental study by Kebbell, Hurren and Roberts (2006) in which participants commit a mock crime confirmed the relationship between the strength of eyewitness evidence and the likelihood of confession.

The second factor identified as important in motivating suspects to confess was internal pressure. This relates to the internal need of suspects to relieve the pressures of guilt through confessing to another person. It is reasonable to assume that, especially in cases of serious crimes such as murder and sexual offences, some guilty individuals will

experience emotional distress and will seek to relieve this by owning up. This assumption appears to have empirical support as, in both the Gudjonsson and Petursson (1991) and Gudjonsson and Bownes (1992) studies, sexual offenders appeared more inclined than other types of offenders to confess due to internal pressure. The process of confessing to satisfy internal needs is described by Reik (1959) in his psychoanalytic perspective of confession. He emphasises that behind the compulsion to confess often lies an unconscious need for punishment and impulses of self-betrayal. In doing so, Reik compares the act of confessing with the work of patients in psychoanalysis: "... many of the confessions, in which adequate motives like vindictiveness, boasting, jealousy etc. can be recognized easily, find their deepest motivation only in the need for punishment" (p. 262). Reik also argues that the process of confession can be regarded as symbolic of a child confessing to an accusing but forgiving father:

"In his confession, the criminal has admitted his misdeed to the community, as the child once admitted his naughtiness to his real father or to his substitute. As the confession of the child unconsciously represents a new wooing for love, an attempt at regaining the lost object, the criminal shows in his confession his intention to re-enter society by declaring himself deserving of punishment. The outsider is on his painful detour back to the family of man" (p. 279).

The third reason for suspects' confessions identified by Gudjonsson and Petursson was external pressure. This factor relates to suspects coping poorly with the interrogation,

experiencing tension and confusion, confessing too readily or hastily, and the desire to escape the immediate pressures of the interrogation. Although Gudjonsson and Petursson found that perception of proof, internal pressure, and external pressure are the most significant reasons for confessions, they note that other factors are sometimes involved such as the wish to protect another person such as a victim or an accomplice.

In both the Gudjonsson and Petursson (1991) and Gudjonsson and Bownes (1992) studies, external pressure was associated with dissatisfaction on the part of suspects regarding the confession. Suspects who had been subject to external pressure reported an initial unwillingness to confess, feelings of tension, nervousness and confusion, and a sense of not coping well with the interview. It would seem to be of note that the suspects also reported some belief that they may be allowed to return home after confessing and also that they might receive a lighter sentence. This finding raises the possibility that the seriousness of the crime may have been downplayed by the interrogators in these cases, a ploy advocated by Inbau et al. (2001) as part of the Reid Technique. In contrast to external pressure, internal pressure to confess and perception of proof were associated with offenders being pleased that they had confessed and believing it was in their own interest.

This thesis is concerned primarily with aspects of external pressure experienced by forensic interviewees and how these may interact with personality characteristics of interviewees. The following sections will consider the various categories of external

pressure suspects and witnesses may experience and the psychological processes underlying them.

### 1.5.2 Ecological control

A number of authors have argued that the physical context of interrogation can be an important contributory factor in influencing the behaviour of suspects (Driver, 1968; Inbau et al., 2001; Irving & Hilgendorf, 1980; Kassin & Gudjonsson, 2004; Tousignant, 1991; Zimbardo, 1968; Zulawski, & Wicklander, 1993). The physical setting of interrogation is characterised by confinement, physical isolation, unfamiliarity and is generally free from external stimuli and support, except in some cases where a legal representative may be present. It should be acknowledged, however, that legal representatives vary in their active involvement in interviews (Medford, Gudjonsson, & Pearse, 2003) and hence in the degree of social support they offer. Indeed, Cohen and McKay (1984) suggest that for social support to lessen the adverse effects of stress, the resources that are perceived to be available must match the needs elicited by the stressful event. The physical characteristics of interrogation are, of course, under the control of the criminal justice system and, more immediately, the interrogators themselves. By comparison, the suspect has little or no influence over the ecological setting and it is argued here that suggestions that the physical context can be used to exert influence over suspects are psychologically valid.

Sells (1973) proposed a social system model to outline the experience of isolation and enclosed space. He argued that tolerance of such pressures could be understood in terms

of a reward-cost ratio in relation to where a participant was situated along a voluntary-involuntary dimension. Thus, voluntary participation was associated with a high reward-low cost ratio, whereas involuntary participation resulted in a contrastingly unfavourable ratio. As will be discussed below, the effects of isolation are less favourable for many criminal suspects whose participation is involuntary than for experimental participants.

Evidence for the impact of sensory restriction, confinement and social isolation has been obtained from experimental studies. Research examining the effects of sensory deprivation in human volunteers has shown that even relatively short-term reductions in sensory input can lead to negative consequences for many individuals such as impaired vision, difficulties in focusing, and deterioration on cognitive task performance (Bexton, Heron & Scott, 1954; Grassian & Friedman, 1986; Kubzansky, 1961). Zuckerman, Persky, Link and Basu (1968) reported that several hours of confinement produced an arousal of anxiety, negative affect, somatic complaints, tedium stress, unreality stress and activation of the adrenal cortex. These authors conclude that, "The stress effects of confinement are rather massive and are found even when Ss are neither sensorily nor socially isolated" (p. 194). In a similar experiment, Zubek, Bayer and Shepherd (1969) noted that prolonged confinement produced a variety of stress effects as measured physiologically and psychologically. In their mock prison study, Haney, Banks and Zimbardo (1973) reported that, in reaction to confinement pressures, prisoner-subjects exhibited negative emotional reactions such as emotional depression, crying, rage and acute anxiety (although it should be noted that the prisoners were subject not only to the

pressure of confinement, but to verbal insults from the guard-subjects). More recently, Custaud et al. (2004) found that long-term confinement and inactivity resulted in considerable physiological changes, specifically cardiovascular and hormonal, in a group of male volunteers.

Physical isolation inevitably means social isolation and, in this context, the influencing of attitudes and behaviours can take place. To maintain our attitudes, beliefs, and opinions, we rely to an extent on the support of others. On this matter, Cartwright (1965) notes that, "an opinion, a belief, an attitude is 'correct,' 'valid,' and 'proper' to the extent that it is anchored in a group of people with similar beliefs, opinions, and attitudes" (p. 27). It follows that the denial of social support and the absence of consensual validation, particularly over a prolonged period as occurs in some interrogations, makes the maintenance of beliefs and opinions more difficult.

Studies of proxemics have revealed that manipulations of interpersonal distance can be stressful, particularly when interpersonal distance is inappropriately close (Hall, 1966; Sommer, 1969). For example, Baxter and Rozelle (1975) found that, in a simulated police interview, interpersonal crowding caused discomfort to experimental participants as indicated by their distinct behavioural reactions. Given that a criminal suspect has no means of escape, the interrogator positioning himself physically close to a suspect, as recommended by Inbau et al. (2001), puts the suspect at a psychological disadvantage.

The above experimental studies offer some examples of how confinement pressures, isolation and physical positioning can have adverse effects on the physiology, cognitive functioning, affective state and behaviour of human volunteers. The question is whether or not the results from these studies can be generalised to real interrogations in police custody. As noted previously, it must be borne in mind that there are some clear differences between the experiences of suspects and those of experimental participants. The latter participate on a voluntary basis with the knowledge that they can leave at any time and that they are very unlikely to come to harm. Indeed, experimental participation may be positively reinforcing on the basis of financial reward or, in the case of undergraduates, course credits (Kimmel, 1996; Rubenstein, 1982). In contrast, the experience of criminal suspects is qualitatively different and, in the terms defined by Sells (1973), occurs on a low reward-high cost ratio. For the suspect, police custody is characterised by involuntariness, formality, physical control, and uncertainty (Gudjonsson, 2003; Irving, 1980). As such, it can be argued that the effects of ecological control are considerably more serious for suspects.

### 1.5.3 Attitude change

With direct reference to Inbau et al.'s (2001) interrogation techniques, Memon, Vrij and Bull (2003) offer a social-psychological explanation for confession. They suggest that theories of attitude change (Ajzen, 1985; Ajzen & Madden, 1986; Schifter & Ajzen, 1985) are useful in understanding the effectiveness of interrogation. Attitudes are associations between attitude-objects (virtually any aspect of the social world of

personal significance to the perceiver) and evaluations of those objects (Fazio & Roskos-Ewoldsen, 1994). Attitudes influence and are influenced by the way in which social information is processed and, importantly, they often influence behaviour. Memon et al. suggest that, for criminal suspects, the act of confessing is an attitude-object and, to increase the likelihood of confession, the task of the interrogator(s) is to enhance the perceived positive aspects of that attitude-object. Accordingly, it is first necessary to reduce the perceived negative evaluations of confession, namely that admission of guilt will lead to conviction while denial of guilt will avoid conviction. This can be achieved by presenting the suspect with evidence (false or otherwise) of their guilt and communicating that denial of guilt too will lead to conviction. To enhance the perceived positive aspects of the attitude-object, interrogators can communicate the benefits of admission, e.g. a reduced sentence. Finally, the negative aspects of the attitude-object are emphasised by highlighting the penalties associated with denial. The result of these communications by the interrogator(s) can be a changed attitude toward confession and behavioural change. As Memon et al. state, "The more positive the attitude towards confession, the more likely that the suspect will finally confess" (p. 63).

### 1.5.4 Conformity

Social psychological research has shown that groups can bring considerable pressure on individuals to conform to the group view against their privately held beliefs. Asch (1951, 1955) found that, when making a judgement of an unambiguous visual stimulus, individuals conformed to group opinion, even when the group's opinion was blatantly wrong. Deutsch and Gerard (1955) argued that such conformity is caused by normative

influence, the need to gain social approval or to avoid social disapproval. Asch's conformity studies involved visual perception rather than recall of past events, but this distinction may not be significant. As Driver (1968) noted: "Memory is a compound of perceptions, attitudes and beliefs and does not well occupy a category distinct from them" (p. 53). In real police interviews, the interrogator, or interrogators, may not constitute a group in the conventional sense. Nevertheless, it is not difficult to draw a parallel between the experimental work on conformity and the process of interrogation. From the viewpoint of the suspect, the "group" is the prosecution system itself, the interrogators being its most visible members.

### 1.5.5 Obedience to authority

Irving and Hilgendorf (1980) note that, in all police interrogations, whether with witnesses or suspects, the interrogator carries a socially legitimate authority that is clearly recognised by all citizens. They argue that the behaviour of witnesses or suspects can be manipulated by virtue of subordination to police authority and power and they cite Milgram's (1963, 1974) laboratory studies of obedience in support of their argument.

These studies are significant in demonstrating that "normal" individuals will readily yield obediently to the direct social influence of authority figures even when such persons have little real power to enforce their orders. Tendencies towards obedience stem from a number of causes. In particular, people may implicitly transfer responsibility for their actions to the persons in authority. In Milgram's (1974) terms,

people enter an "agentic state ... the condition a person is in when he sees himself as an agent for carrying out another person's wishes" (p. 135). Obedience is made more likely where authority figures gradually escalate the scope of their orders (Gilbert, 1981) and in instances where there is little time for target persons to consider their actions in detail. The symbolism of authority and visible signs of power appear also to be important in determining degrees of obedient behaviour (Bickman, 1974; Bushman, 1984; Haney et al., 1973). Numerous replications of Milgram's studies have shown that the tendency of most people to obey authority figures is not limited to American cultures (e.g. see Burley & McGuiness, 1977; Kilham & Mann, 1974; Meeus, & Raaijmakers, 1986; Shanab & Yahya, 1978).

As with all laboratory-based research, it can be questioned to what extent Milgram's procedures possess mundane realism, i.e. a similarity between the circumstances surrounding the experiments and those in real situations, and to what degree the findings can be generalised to real-life contexts such as forensic interviews (Orne & Holland, 1968). Irving and Hilgendorf point out that the connection lies in the proneness of both experimental participants and witnesses and suspects in interrogation to obey instructions which they would ordinarily dismiss. For the experimental subject, it is the infliction of pain on another human being, a behaviour that runs counter to societal norms. For the witness or suspect, it is to obey instructions to provide information or to confess, behaviours which, in the immediate absence of legitimate authority, might not normally occur. Irving and Hilgendorf further argue that, in general, there is little doubt

that police investigators will be accorded more legitimate authority than experimenters.

## 1.6 Criticisms of the Reid Technique

Referring to the Inbau and Reid (1962) edition of their manual and the techniques described in other police interrogation manuals, Zimbardo (1967, 1968) questions the "voluntariness" and validity of confessions obtained by them. Unequivocal in his criticism, he notes:

"I am convinced that these methods are psychologically coercive; that they deprive the individual of his human dignity and fundamental rights; and that they debase the police who use them even though the police are trying to be fair as well as efficient. Let the reader judge my contention that they do not serve justice" (1967, p. 450).

Zimbardo highlights Inbau and Reid's assertion that none of their methods are "apt to induce an innocent man to confess" noting that they present no supporting evidence for this claim. Since the publication of Zimbardo's viewpoint forty years ago, it seems that psychological researchers have maintained their critical stance towards the Reid Technique and, despite revisions to Inbau et al.'s interrogation manual in the interim, the fundamental objections remain the same (Meissner & Russano, 2003).

### 1.6.1 Gudjonsson's criticisms

As noted previously, a key aspect of the Reid Technique involves playing down the moral seriousness of the crime through the suggestion of certain excuses or moral

justifications to the suspect. One such "theme" is the suggestion that there was a non-criminal intent behind the suspect's actions. For example, it might be suggested that the suspect, although physically responsible for their involvement in a crime, acted in an accidental manner without premeditation to commit a criminal offence. Gudjonsson (2003) criticises this tactic on the basis that it falsely suggests that admission of involvement will carry reduced negative consequences for the suspect. He cites the case of a Mr R who falsely confessed to the murder of a woman, claiming he had acted in self-defence. The man was sentenced to death, but later released when DNA evidence found someone else to be guilty. According to Gudjonsson, among the factors leading to Mr R's false confession was the alleged suggestion by the police that the killing had been unintentional and that admission of involvement would lead to his release within hours.

Gudjonsson (2003) is critical too of Inbau et al.'s alternative question tactic. This involves presenting to the suspect two incriminatory choices concerning their involvement in the crime. For example, the question might be presented, "Is this the first time something like this has happened, or have you done this kind of thing hundreds of times before?" Importantly, either choice is an admission of guilt. The alternative question may be followed by a supporting statement which further suggests the suspect should choose one answer over another, "I think this is the first time, isn't it, Joe?" As can be seen, an alternative question together with a supporting statement makes it easier for a suspect to admit his guilt. Presumably, a simple "yes" or nod of the head might be

enough.

Gudjonsson regards this tactic as "highly coercive" and particularly dangerous when used with suspects of low intelligence. He notes Inbau et al.'s claim that, "... no innocent suspect, with normal intelligence and mental capacity, would acknowledge committing a crime merely because the investigator contrasted a less favourable circumstance to a more desirable one and encouraged the suspect to accept it" (p. 365, quoted in Gudjonsson, 2003, p. 20). According to Gudjonsson, this statement is misleading as it disregards the overall context in which the alternative question tactic is used. It must be borne in mind that the two incriminating choices are presented to the suspect following an often lengthy period of confrontation and in the context of other techniques used to weaken the resistance of the suspect.

In addition, Gudjonsson (2003) argues that Inbau et al. (2001) are misleading in their suggestion that their methods "may" involve trickery or deceit. In fact, there is no doubt that trickery and deceit are essential to the Reid Technique. Gudjonsson also suggests that the psychological manipulation and interpersonal pressure characteristic of the Reid Technique contradicts somewhat Inbau et al.'s statement that they oppose "the use of force, threats of force, or promises of leniency" (p. xii).

Like Zimbardo (1967, 1968), Gudjonsson notes that Inbau et al. have not published any

data or studies of their observations: "...their book is full of assertions and generalisations about their technique without supporting empirical evidence" (p. 21). As a result, it is not known how successful, relative to other interrogation methods, the Reid Technique is in obtaining confessions. Crucially, the false confession rate following use of the Reid Technique is not known.

### 1.6.2 Kassin and associates' criticisms

The Reid Technique has also been criticised on the basis that investigator response biases during the pre-interrogation interview may unduly influence the conduct of the interrogation itself. Meissner and Kassin (2004) have argued that wrongful presumptions of guilt at an early stage may lead interviewers to adopt an overly confrontational questioning style with innocent suspects. Kassin, Goldstein and Savitsky (2003) investigated this possibility experimentally. Following previous work, a sample of students either committed a mock crime or took part in a related innocent activity and all were instructed to convince the interrogator of their innocence. A group of student interrogators were told to secure a confession from the suspects and also to attempt to determine the innocence or guilt of the suspects. The interrogators were supplied with material on the Reid Technique from Inbau et al.'s (2001) *Criminal Interrogation and Confessions* and invited to devise an interviewing strategy. The interrogators were given either guilty expectations (they were told that 80% of the suspects committed the crime) or innocent expectations (told that 20% of the suspects committed the crime). It was found that interrogators in the guilty expectations condition chose more guilt-presumptive questions than those in the innocent expectations condition, used more

techniques from Inbau et al.'s manual at the outset of the interrogation and judged the suspects as guilty at the end of the interview. It was also found that the interrogators applied most interrogative pressure to the innocent suspects. These findings strongly supported Kassin et al.'s hypothesis that interrogator expectations can lead to behavioural confirmation effects and bias perceptions of guilt.

In an earlier study, Kassin and McNall (1991) used experimental methods to examine two other aspects of the Reid Technique: maximization (an intimidation ploy which presents false incriminating evidence and exaggerates the seriousness of the offence and the magnitude of the charges), and minimization (false offers of sympathy and tolerance, playing down the seriousness of the crime, and implications of a lenient charge). Participants read transcripts of interrogations in which a range of interrogative techniques were employed. They then reported their beliefs regarding sentencing expectations, likelihood of confession, degree of interrogative pressure applied to the suspect and so on. Kassin and McNall concluded that, far from being subtle and noncoercive, as Inbau et al. (2001) would claim, maximization and minimization techniques are dangerous to the extent that they may, through pragmatic implication, alter the sentencing expectations of suspects and lead innocent suspects to make self-incriminating statements. Thus, they claim that confessions obtained by these techniques may not be so different to those obtained by explicit promises and threats. Further, they argue that while confessions elicited by promises of leniency are considered inadmissible by most courts, confessions obtained by minimization may be judged as

satisfactory by jurors.

Russano, Meissner, Narchet and Kassin (2005) developed an experimental paradigm to test the effects of interrogative pressure on the likelihood of both true and false confessions. Participants were instructed to solve logic problems jointly with a partner. The partner was a confederate of the experimenters and, in the "guilty" condition, the confederate asked participants for help. As a result, a proportion of the sample was induced to violate the rules of the experiment, i.e. to "cheat". Both "guilty" and "innocent" participants were interrogated using either an explicit offer of leniency (a "deal") or minimization tactics and asked to sign a confession. As predicted, an explicit offer of leniency increased the likelihood of both true and false confessions. Importantly, minimization tactics, as advocated by Inbau et al. (2001), produced the same result. It seems the participants associated the minimizing statements with the offer of a "deal". Russano et al. conclude that, contrary to Inbau et al.'s claims, minimization has the potential to compel innocent suspects to confess. The experimental findings cited above confirm the views of Gudjonsson that interrogative pressure, inappropriately applied, can lead to false confessions. (For further criticisms of the Reid Technique, see Ofshe and Leo (1997)).

## 1.7 Investigative interviewing approaches

The discussion thus far has, for the most part, considered the questioning of criminal suspects and, in line with the literature and police practice, the term "interrogation" has been used. The term "interviewing" was originally considered to apply only to the

questioning of witnesses, victims and complainants (McKenzie, 2001). Gudjonsson (2003, 2006) notes that such a distinction is an arbitrary one, however. Indeed, in the UK especially, as a result of legislative changes, concerns over miscarriages of justice and reviews of police questioning methods by police officers, lawyers and psychologists, the term "investigative interviewing" has been adopted to refer to all categories of interviewee including those with suspects. This new term is intended to reflect an improved ethos in police interviewing and a move away from confession seeking to a search for information (Clarke & Milne, 2001; Williamson, 1993, 1994). Bull and Milne (2004) describe the change as one "from a blinkered, closed-minded, oppressive, and suggestive interviewing style to one involving open-mindedness, flexibility, and the obtaining of reliable evidence" (p. 186). In this section, attention is turned to the questioning of witnesses as well as suspects and three strands of investigative interviewing are discussed: the Cognitive Interview, Conversation Management and the PEACE model.

### 1.7.1 The Cognitive Interview

A number of authors examining police interviewing practice in the US and UK have noted that police officers generally receive little training on how to conduct interviews with cooperative eyewitnesses (e.g. Fisher, Geiselman & Raymond, 1987; George & Clifford, 1992). As a result, witness interviewing is too often conducted with a less-than-rigorous attitude, interviewers make avoidable errors and interviews fail to elicit a sufficient amount of relevant information. It was to address these perceived shortcomings that the Cognitive Interview (CI) was developed (Geiselman, Fisher,

Cohen, Holland & Surtes, 1986; Geiselman et al., 1984; Geiselman, Fisher, Mackinnon & Holland, 1985). The CI is a series of memory-enhancing techniques based on theoretical, laboratory research in cognitive psychology intended to improve the completeness and accuracy of eyewitness accounts. Designed primarily for use with cooperative victims and witnesses, Bull and Cherryman (1995) suggest that the CI may also be of use with cooperative suspects, i.e. those making admissions of guilt.

The first technique is termed "reinstating the context". It is based on the understanding that memory is context-dependent (Godden & Baddeley, 1975), and so asking a witness to think about how they were feeling just before and during the event to be recalled, perhaps evoking the sounds and smells relating to the event, should facilitate retrieval. The second technique, "report everything", involves having the witness concentrate very hard on the task and focussing on all the sensory details, even those considered trivial or irrelevant, in order to improve recall. These two techniques rely on the concept of encoding specificity, i.e. maximising the relationship between the encoding operations at the time of the crime and the cues available at the time of recall (Tulving & Thomson, 1973). Cognitive interviewers also encourage "multiple retrieval attempts". It is argued that inviting a witness who feels they have recall everything about an event to have another attempt can unlock previously unrecovered detail simply because of a confident assertion that there is more to be retrieved. Unless prompted otherwise, witnesses will often recall events in chronological order, but if they are asked to recall details in a different order, or from a different perspective, techniques known as "varied retrieval",

this may elicit additional information (Baddeley, 1999).

In recognition that many eyewitnesses become anxious during police questioning and that they may have confused expectations about their role, the original CI was revised (Fisher & Geiselman, 1988; Fisher, Geiselman, Raymond, Jurkevich & Warhaftig, 1987). The resultant enhanced CI was informed by the content analysis of tape-recorded field interviews in which a range of communication difficulties typical to police interviews were identified (Fisher, Geiselman & Raymond, 1987). The enhanced CI places more emphasis on the development of rapport and generally a transfer of more control to the witness. Altogether, the revised CI includes thirteen basic skills: establishing rapport, listening actively, encouraging spontaneous recall, asking open-ended questions, pausing after witness responses, avoiding interrupting, requesting detailed descriptions, encouraging intense concentration, encouraging the use of imagery, recreating the original context, adopting the witness's perspective, asking compatible questions, and following the sequence of the CI (Fisher & Geiselman, 1992).

Fisher, Geiselman and Amador (1989) tested the enhanced CI in the field with real victims and witnesses of crime. A group of experienced police officers from the Metro-Dade Police Department were trained in the CI and subsequently tape-recorded a set of interviews over several months with victims of either commercial robbery or purse-snatching. Transcriptions of these interviews were analysed and compared to pre-training interviews conducted by the same officers, and to interviews conducted by a

control group of untrained officers. Comparison of group means showed that the trained interviewers elicited 47% more information after than before training, and that they obtained 63% more information than the untrained interviewers. Fisher et al. also reported high rates of corroboration between interviewee statements and those taken from other witnesses and victims and argued, in line with laboratory findings, that the added information elicited by the CI does not come at the expense of increasing incorrect information. (See also George & Clifford, 1992).

Aside from theoretical and technical considerations of investigative interviewing practice, it is important to assess the acceptability of techniques and their applicability as judged by real forensic interviewers. Towards this aim, Kebbell, Milne and Wagstaff (1999) surveyed UK police officers on how useful they found the CI in practice and how frequently they used its various components. The survey sample comprised officers trained in the CI techniques as well as untrained officers. Trained officers claimed more than the untrained officers to invite witnesses to use context reinstatement, imagery, changed order and changed perspectives. The techniques found to be most useful and most frequently used by the trained officers were establishing rapport, encouraging concentration, reporting everything, context reinstatement, and asking compatible questions. CI techniques such as changed order, changed perspectives and transfer of control were rated as less useful. Overall, the survey showed that police officers regarded the CI in positive terms, believing it to produce more information from witnesses than the standard interview.

Fisher et al. (1989) acknowledge that in the wider investigative interviewing context, the CI has a number of limitations. They note that it has most practical utility in cases in which the bulk of evidence results from eyewitness reports and least in cases where there is a lot of physical evidence. As already noted, the usefulness of the CI does not extend to all witnesses, only those who are forthcoming and cooperative with the police. Fisher et al. concede that the CI often takes longer to conduct than a standard interview and, as such, is less suitable in instances where there is limited time to conduct the interview. They also acknowledge that the CI takes more concentrated mental effort on the part of the interviewer than does the standard interview.

### 1.7.2 Conversation management

Unlike the CI, the conversation management approach is intended for suspects or witnesses who are uncooperative or resistant. It was introduced by Eric Shepherd, a forensic and counselling psychologist, following a request by Merseyside Police to develop a method of interviewing which was in close accordance with the Police and Criminal Evidence Act (PACE, Home Office, 1985a). Conversation management is a non-directive style of interviewing based on a set of "principles of ethical conversation" according to which the interviewee should be respected, empathised with and treated as a conversational participant rather than a "question-answering machine" (Mortimer & Shepherd, 1999; Shepherd, 1991). The approach is summarised by the mnemonic GEMAC (Greeting, Explanation, Mutual Activity, Close).

According to Shepherd, interviewers need to be mindful of the potential for resistance not only from suspects, with whom it may be more likely, but also from some witnesses. Resistance is described as an interplay of two dimensions: "ability to tell-inability to tell", and a motivational decision-making dimension, "willingness to talk-unwillingness to talk" (Shepherd, 1993). The task of the interviewer is to manage the interaction with the interviewee using behaviours which "inspire the confidence, trust and respect essential to facilitate disclosure and to overcome resistance" (Shepherd & Milne, 1999, p. 135).

Accordingly, the interviewer should communicate a sense of equality with the interviewee during the initial greeting and thereafter. The interviewer should be considered in their explanation of the reason for the interview, the "route map" (agenda of actions and topics), the routines (e.g. interviewer's note taking), and expectations (i.e. preparing the interviewee to view silence as constructive rather than pressuring). Shepherd likens the principle of explanation to the activities of counselling and psychotherapy in which therapists aim to define the working relationship with their clients.

The next step in the conversation management approach involves two "mutual activities": monitoring and assertion. Monitoring activities include active listening and focused observation of the non-verbal behaviours of the interviewee. Shepherd warns against the often typical practice of "inappropriate pacing", i.e. rushing in immediately

after the interviewee's response. Instead, he argues that deliberate pausing throughout the interview allows for more efficient cognitive processing on the part of the interviewee: "working memory can operate efficiently, emergent information and prior knowledge (from long-term memory) can be compared, contrasted, synthesized and stored" (Shepherd, 1993, p. 8). The activity of assertion relates to the use of productive questioning, i.e. open-ended style questions which invite full disclosure of information from the interviewee. The final step involves closing the interview by summarising the account provided by the interviewee. The emphasis here is on constructing a highly accurate record of the interviewee's disclosure.

There are no published studies on the effectiveness of exclusive use of the conversation management approach. However, both conversation management and the CI form part of the PEACE model of investigative interviewing which has, as detailed below, been subject to evaluation.

### 1.7.3 The PEACE approach

The development of the PEACE model of interviewing in the UK was prompted by an array of factors. A number of miscarriages of justice including high profile cases such as the Guildford Four and the Birmingham Six focused public and judicial awareness on the use by some police officers of coercive interrogation techniques. Research into false confessions and psychological vulnerability increased understanding of the role of manipulative interrogation tactics in wrongful convictions (Gudjonsson, 2003; Gudjonsson & MacKeith, 1988). Introduction of the Police and Criminal Evidence Act

(PACE, Home Office, 1985a) and accompanying Codes of Practice (Home Office, 1985b) instituted significant changes in police practice in the UK which remain to this day. Among other things, the Codes set out procedures for the interviewing of suspects and require that interviews with suspects are tape recorded. As Gudjonsson (2003) notes, this "reduced the scope for coercive questioning, and the use of deception, trickery and psychological manipulation" (p. 52). The tape recording of interviews has allowed researchers to more easily assess police interviewing techniques. Some of this research is described in detail in the next section of the current chapter and, together with data from observational studies, has served to highlight a number of deficiencies in police interviewing methods with suspects.

Two Home Office publications detailed a new set of principles for investigative interviewing and set out a training programme for interviewing skills (Home Office Circular 22/1992; Home Office Circular 7/1993) then, following publication of two booklets on investigative interviewing (CPTU, 1992a, 1992b) by the Association of Chief Police Officers and the Home Office, a week-long nationwide training course was developed to train as many police officers as possible in the new interviewing approach (Williamson, 1993, 1994). The name of the new approach, PEACE, is an acronym of the five stages of the interview. (For a more detailed description of the PEACE model, see Milne and Bull, 1999).

*Planning and preparation*. The importance of adequate planning and preparation before

the interview begins is highlighted. This includes understanding the purpose of the interview, obtaining background information on the case, and defining aims and objectives.

*Engage and explain*. This refers to engaging the interviewee in conversation in a respectful manner and explaining the reason for the interview and its format.

*Account*. This element of the approach incorporates the CI (Fisher & Geiselman, 1992) and conversation management (Shepherd and Milne, 1999) and is aimed at allowing the interviewee to provide an accurate account.

*Closure*. Interviewers are taught to conclude the interview effectively by reviewing and summarising the account to the satisfaction of both the interviewer and interviewee.

*Evaluate*. The importance of evaluating the interview and the performance of the interviewers is emphasised. Williamson (1994) suggests that reviewing tape-recorded interviews would be a useful learning experience for police officers.

Before the PEACE model was introduced nationally, a five day experientially based pilot training course was run in 1992 and involved practice in planning and carrying out interviews with both witnesses and suspects. An independent evaluation of the course conducted by McGurk, Carr and McGurk (1993) reported the success of the course in

these terms: "Students knowledge increased and their interview skills were significantly enhanced. The improvement was sustained over a six month period compared to the performance of the control group which, over the same period, was subject to only routine work experience" (p. v). The success of the pilot course led to the introduction of PEACE training nationally. A further evaluation of the PEACE interviewing approach by Clark and Milne (2001) produced a mixed assessment. While the standard of interviewing appeared to have improved since pre-PEACE with regard to the general ethos, there was clear evidence of shortcomings, particularly in relation to witness interviewing (Bull & Milne, 2004; Milne & Bull, 2003). Some of Clark and Milne's key findings are reported in the following two sections of this chapter.

In the last five years, a new strategy for police interview training has been implemented in the UK. Building on the foundation of the PEACE model, the new strategy is a five-tier system intended to accommodate police officers at different career stages and to deal with a range of different crimes (Griffiths & Milne, 2006). The new approach to interview training stems from recognition that a "one size fits all" approach to interviewing is not appropriate and from a recommendation made by Clark and Milne (2001) - and also earlier by Shepherd (1988) - that interviewing skills are best developed systematically across the career span.

## 1.8 What happens in police interviews with suspects

At the outset of this chapter, it was noted that, until recently in some countries, police interviewing was largely an unseen activity, unavailable for empirical study. The

introduction in the UK of mandatory recording of interviews with suspects changed that and has allowed researchers to assess which interviewing tactics are in use, the extent to which psychologically manipulative tactics might be in use, and the impact of legislation and training initiatives on interviewing. The research studies included here, all conducted in the UK, are presented in chronological order and cover the period prior to the introduction of PACE through to the promotion of the investigative interviewing ethos and evaluation of the PEACE interviewing model in the field. Data were drawn from observational study of real interviews with suspects, analysis of audio and video recordings, as well as typed transcripts of interviews.

### 1.8.1 Irving's studies

Irving (1980) conducted an observational field study of police interrogation practice and procedure at the behest of the Royal Commission on Criminal Procedure. The study was exploratory in nature and attempted to answer the question "what actually goes on" in police interviews with suspects. Irving was interested in whether any special techniques were being utilised and, if so, how these might influence the outcome of the investigations. Seventy six custodial interviews were observed over a six month period at a single English police station. The sample of cases covered a wide range of types of crime, although serious crimes were most strongly represented and Irving made an attempt to ensure that a wide selection of police officers was observed.

The observer sat in the interview room for the full duration of the interviews, with full view of the suspect and the interviewers. Irving fully acknowledged that the presence of

an observer was potentially problematic in terms of the risks of affecting the suspect's and/or interviewers' behaviour, and with regard to observer bias. Any assessment of the reliability and validity of the findings must take these factors into account. Nevertheless, given the hitherto closed and self-contained nature of Criminal Investigation Departments, the chosen method of participant observation was significant and clearly suited the research objective.

In general, the majority of criminal suspects tend to be interviewed only once (e.g. see Baldwin, 1993). However, where more serious crimes have been committed, there is an increased likelihood of more than one interview being conducted. In Irving's sample of cases, 30% of suspects were interviewed once, 42% twice and the remainder three or more times. The time taken per interview varied from under ten minutes to over two hours. However, the vast majority of interviews (88%) were conducted in less than sixty minutes.

Irving noted that in a small number of cases, interviews were conducted in order to resolve doubts held by the investigating officer. Thus, some interviews were carried out to eliminate a suspect from further investigation, or to establish whether or not a crime had actually been committed. Where such doubts did not exist, a considerable number of interviews (42%) were conducted for the purpose of obtaining a confession, either as main evidence or additional evidence in the case.

Irving describes the relevance of other evidence to the process of interviewing a suspect. He notes that forensic evidence typically takes time to become available and, as such, is considered to be a "back-runner" in relation to confessional evidence in securing a criminal charge. Importantly, however, the promise of the arrival of forensic evidence (fingerprints, blood type, matching of hairs or fabric fibres) is considerable, not only in and of itself, but also to the interviewing process. On this issue, Irving observes, "It is generally agreed that the lay reverence for forensic science created by writers of detective fiction is a considerable help to detectives and it would be a disadvantage to the CID if the real state of affairs was generally known" (p. 117).

Where suspects are found to be in possession of certain types of incriminating property, e.g. stolen goods, weapons or narcotics, this is considered as "real evidence" and, for obvious reasons, deemed to be of primary assistance where interviewing a suspect is concerned. Irving also observed that eyewitness statements were commonly used as part of interviews in an attempt to secure an admission.

Irving observed that, for serious offences, two officers were always present at the interview and this allowed for some variation in how the interview was conducted. In some cases, one interviewer was active and the other passive; in other cases, both officers were active in the interrogation. Irving rarely observed the use of contrasting

styles as in the often mentioned "Good Cop/Bad Cop" routine.

Irving identified four discernible interviewing styles used by the police officers. The first he defined as "business-like", the second as "befriending", and the third "authoritative". The fourth style was described as primarily authoritative with elements of the other two styles. The chosen style seemed to be employed in response to suspect behaviour. Thus, initially aggressive suspects were subject to an "authoritative" style and so forth.

The effects of police custody on the suspects were examined. It was noted that the unfamiliarity of the cell block, confinement and isolation, the threat of harm, and the insistence on obedience to police authority, all had the potential to impair the suspect's ability to deal with the interrogation. Hence, when considering the behaviour of a suspect under interrogation, it would seem circumspect to consider not only the conditions of the interview itself, but also the external conditions which preceded it.

In the course of interviews with suspects, police officers attempt to persuade or influence suspects to confess or to provide information. Irving found that five types of persuasive tactic were used in the interviews he observed. The first and most frequently used tactic was to persuade the suspect that, on the basis of existing evidence against the suspect, the inevitable outcome of the interview would be a criminal charge. The idea was thus communicated that there was only one decision to be made, i.e. when rather

than whether to confess. Such a strategy relied on the informational advantage held by the police over the suspect, but, if not based on the truth, amounted to an "informational bluff". Irving asserted that this tactic was highly effective in obtaining confessions.

The second tactic arose from the police officers' expert status in terms of their knowledge of the criminal justice system. By disclosing key information at certain stages of the interview, the interviewers were afforded the opportunity to persuade the suspect to co-operate. For instance, to elicit further information from a suspect, it might be emphasised that the courts valued co-operation and honesty on the part of suspects.

The third tactic involved the exercise of police discretion. The police have certain powers over the nature of the charge, the duration of custody, and other matters relating to the ongoing investigation. It was communicated to suspects that the officers could exercise their discretion in ways which could be of benefit or disadvantage to suspects. For example, it was suggested that the suspect could go home when certain issues had been "cleared up", or it was implied that the seriousness of the charge could be altered according to whether or not the suspect told the truth.

In the course of an interview, a suspect will most likely make some assessment of the consequences of confessing. The fourth tactic was aimed at influencing such assessments. Police officers were observed maximising the positive consequences of

conviction. For instance, the suspect might be told that a full confession would enable him to "get things off his chest" and thus feel better. Irving also observed the minimisation of the negative consequences of conviction. For example, it might be emphasised that an admission would increase the likelihood of lenient treatment in court, or that the criminal offence was not of a serious nature and thus not worthy of concern. The fifth and least frequently employed tactic was the use of custodial conditions such as confinement. Irving considered the use of confinement between or before interviewing to be "markedly effective" in dealing with uncooperative and aggressive suspects.

The above persuasive tactics were observed in 73% of Irving's sample of interviews. Interviewers generally used two or more tactics in each interview. In the remaining 17% of interviews, no identifiable tactics were used. Of the sixty suspects observed, 65% made admissions either during the interview or subsequent to it. Whilst Irving suggested that persuasive interviewing tactics may have produced these admissions, he acknowledged that it was difficult to assess with any certainty the direct effect of specific tactics on the behaviour of suspects.

Irving (1980) concluded: "The results of our observations at Brighton confirm that the kinds of techniques which were predicted from a review of the psychological literature, are in fact used, and that there are many similarities between what is taught to American detectives and what happens in the interview rooms at Brighton Police Station" (p. 150).

Irving's research and that of Softley (1980) are of significant interest in that both were carried out prior to the introduction of the Police and Criminal Evidence Act and its Codes of Practice (PACE, Home Office, 1985a, 1985b). Subsequent work on police interrogation thus allows some assessment to be made on how legislation has impacted on police questioning techniques.

### 1.8.2 Baldwin's study

In a key study, Baldwin (1993) analysed 400 video recordings and 200 audio recordings of police interviews conducted in three police forces. The aim was to assess police interviewing techniques. Baldwin makes reference to the "mythology" of police interviewing, i.e. persistent popular notions which hold that suspects are likely to be awkward and aggressive and consequently that most interviews are difficult encounters, confrontational in nature. Further myths, according to Baldwin, are that interviews invariably take a long time, and that special psychological skills are necessary to "turn around" the testimony of resistant suspects. The findings contradicted these popular notions.

It was observed that the majority of suspects (73.7%), far from being resistant, presented a "co-operative" or "submissive" attitude during interviewing and that they were willing to answer police questions. The reason for such behaviour is not clear. It may be that most suspects are genuinely obliging in nature. Alternatively, their accommodation may

be more a product of interrogative pressure exerted before the recording device was switched on.

In line with Irving (1980), Baldwin found that interviews were not protracted affairs and did not last long. Almost 75% of interviews were completed within half an hour and nearly 90% of suspects were interviewed only once. Thus, the idea that suspect interviews commonly consist of long-drawn out "grillings" was seriously challenged.

The notion of the need for special interrogatory skills does not accord with Baldwin's data. Suspects are rarely persuaded to change their testimony. Indeed, it was found that only 20 of the 600 suspects, just over 3%, were persuaded to change their story in the course of an interview. Further, over one third of suspects made a straight confession from the outset of the interview. Baldwin states, "In most of these cases the interviewer needed to do little more than ask suspects what happened, keep quiet while they answered, and then follow up with a few sensible, straightforward questions to clarify basic details" (p. 335).

The Police and Criminal Evidence Act 1984 contains Codes of Practice which detail how the commencement of interviews should be conducted. The Codes indicate that suspects should be cautioned that they have the right to remain silent. Baldwin's analysis revealed that many cautions were delivered in a casual, sometimes incomprehensible,

manner and that over 10% of cautions were delivered incorrectly. It seems that little importance was attached to this part of the procedure.

It was also noted that most of the police interviewers made little effort to establish "rapport" with suspects in the early stages of the interview. As a result, the backdrop for most interviews was one of awkwardness, tension and formality. Baldwin expressed uncertainty as to why the opening interview phase was so poorly conducted, but speculated that either police officers unwittingly engaged in poor technique or that they deliberately attempted to unsettle the witness.

Baldwin proposed that four principal flaws emerged from his detailed analysis. First, he argued that the police interviewers demonstrated a "general ineptitude" characterised by a lack of planning and an absence of interview structuring. Second, officers conducted interviews with a ready made "assumption of guilt". This was apparent from the highly repetitive nature of questioning and from the use of leading questions. Third, interviewers were often coercive and exerted undue pressure on the suspect, or otherwise acted unprofessionally. For example, in a number of cases, the interviewer advised, erroneously, that admission to previous offences would not adversely affect the extent of their conviction. Four, more generally, aspects of the interview were conducted poorly. For instance, interviewers interrupted suspects' responses, asked questions in rapid succession, or lost control of the interview.

Baldwin's final assessment is highly critical. He asserts, "A majority of the police officers whose tapes of interview were examined in this study could not be described as good interviewers, and, however volubly they spoke (as very many of them did) about the high-level psychological concepts that they applied in interviewing, their social skills were often limited" (p. 350). Excerpts from a number of interview transcripts are presented in Baldwin's paper and, to the extent that the excerpts may be representative of police interviewing practice more generally, they make unfavourable reading.

### 1.8.3 Moston's studies

Moston and Engelberg (1993) analysed 118 tape recorded interviews with criminal suspects, covering a wide range of offences. They examined the strategies employed at the outset of the interviews as the opening strategies were deemed to be largely indicative of the interviewer's level of competence and reflective of assumptions of guilt made by the police officer. Two main interviewing strategies, each characterised by a particular form of questioning, were identified: information gathering (or inquisitorial) and confession (or accusatorial).

Within the inquisitorial approach, strategies were aimed at either establishing rapport with the suspect, gathering information related to the specific offence, or gathering information of a less specific nature. Only 5.1% of all interviewers attempted to establish rapport at the outset of the interview. This observation is line with that of

Baldwin (1993). However, Moston and Engelberg are not critical of this finding on the basis that the establishment of rapport is often justified as a means of detecting deception, a practice shown to be of dubious reliability (Moston, 1992; Vrij, 2000; Vrij & Mann, 2001). The most commonly used tactic, employed in 43.2% of interviews, was offence specific information gathering. Here, the offence is explicitly mentioned in the question and the suspect asked for their version of events. The next most frequently occurring strategy, occurring in 28% of interviews, was that of non-specific information gathering. This is a more open-ended tactic in that the offence is not mentioned during questioning.

The accusatorial approach is confrontational in nature and focused on eliciting a confession from the suspect. Here, strategies included direct accusations of guilt, presentation of evidence followed by a request for an explanation, and presentation of evidence followed by an accusation of guilt. Taken together, such strategies were employed in almost one quarter (23.7%) of all the interviews. On the basis of Moston and Engelberg's (1993) evidence, it might be concluded that police interviewing practice is primarily non-coercive in nature. However, the picture alters slightly when an examination is made of interviews in which the suspect has made an initial denial.

Moston and Stephenson (1993) describe the persuasive strategies adopted by police officers in such instances. In the face of an initial denial, police officers have been found to adopt negative or critical feedback according to whether or not evidence against the

suspect was available. Where evidence is not available, interviewers may blatantly state that the suspect is lying. Direct statements of this kind carry the clear assumption of the suspect's guilt and pose a direct challenge to the suspect (Moston et al., 1992). Another form of critical feedback used in the absence of evidence is silence. Silence can be used to either suggest that an answer is wrong or to implicitly request that the suspect elaborates on the answer given.

Some officers resorted to simply repeating questions. Such a technique communicates to the suspect that their response is unacceptable and invites an alternative answer to the question. Repetition of questions, a key feature of the Gudjonsson Suggestibility Scale (Gudjonsson, 1984, 1987b), has been reliably shown to increase suggestibility.

Interviewers were found to comment on the suspect's body language on the premise that certain non-verbal cues are indicative of deceit. The notion that deception and lying can be reliably detected through body language is a popular one, held by both the general public and the police alike. However, a number of studies have demonstrated that, with few exceptions, this notion is spurious (Depaulo & Pfeifer, 1986; Moston, 1992; Vrij & Graham, 1997).

Where evidence against the suspect is available, the strategies employed are somewhat different. The strength of evidence against a suspect has been shown to affect the

outcome of interviews. Moston et al. (1992) demonstrated that strong evidence is associated with a high number of admissions, whereas weak evidence is more likely to be associated with denials. In line with the findings of Irving (1980), Moston and Stephenson (1993) highlight the salience of available evidence as a vehicle for persuasive questioning. Direct evidence, such as the possession of stolen goods, clearly indicates guilt and, as such, is highly effective when introduced into questioning. In contrast, the testimony of witnesses, victims, and accomplices is apparently easily countered by the suspect and hence is far less effective when used as a persuasive tactic.

Overall, Moston and Engelberg (1993), and Moston and Stephenson (1993) are less critical of police interviewing practice than Baldwin (1993). They point to the positive impact that the introduction of PACE and Codes of Practice for Police Officers and the tape recording of interviews has had.

### 1.8.4 Pearse & Gudjonsson's study

In a more recent study, Pearse and Gudjonsson (1999) examined interview data from 18 serious criminal cases. The crimes involved were murder, rape, arson, armed robbery, buggery, robbery, burglary, incest, attempt to pervert the course of justice, and blackmail. They sought to establish the extent to which manipulative interrogation techniques of the type recommended by Inbau et al. (1986) were being used in police-suspect interviews in England and Wales.

The authors had access to both audiotaped recordings and typed transcripts of the interviews. In each case, suspects initially denied the allegation put to them, then during the course of the interview, changed their mind and made a confession. Each suspect later retracted their confession. Pearse and Gudjonsson predicted that, where interviewers had overcome initial resistance from the suspect, the use of manipulative and coercive tactics would have been used.

Analysis of the interview data identified 39 interviewing tactics. Factor analysis was conducted on 33 of these tactic variables and yielded six factors. The first and primary factor was labelled "intimidation". Among the tactics loading on this factor were maximising the seriousness of the offence, maximising the anxiety of the suspect, manipulation of detail, and manipulating the self-esteem of the suspect. The second factor was named "robust challenge". The loadings here relate to outright challenges made such as comparing suspect testimony with that of the witness, and asking if the witness was lying. The third factor was described as "manipulation". Tactics loading on this factor were those which minimised the seriousness of the offence and minimised the responsibility of the suspect, and offers of inducement. Pearse and Gudjonsson refer to these first three factors as "overbearing" in nature, whilst the remaining factors are described as more "sensitive".

The fourth factor was labelled "question style". This included a high loading for the use of leading questions and a moderate loading for closed questions. The fifth factor was

named "appeal". Tactics loading on this factor were appeals to the suspect's good character, appeals to tell the truth, reassurance, and suggestions that it would be in the suspect's interest to confess. The sixth factor was described as "soft challenge". Tactics loading here were the introduction of witness testimony, general evidence, questioning in low tones, and shame reduction tactics. Together, the six factors accounted for 38.2% of the variance.

Pearse and Gudjonsson (1999) conclude: "This study has confirmed that in serious criminal cases, where there is an initial resistance to confess, British police officers have resorted to American-style tactics to overcome resistance and secure a confession. Even though in the 18 cases the tactics were successful in terms of obtaining a confession, this was achieved at a considerable risk of the confession being rendered inadmissible by a court and the defendant acquitted" (p. 234).

### 1.8.5 Clarke & Milne's study

Clark and Milne (2001) conducted a comprehensive evaluation of the PEACE model of investigative interviewing. Assessing 177 taped interviews, they found that interviews with suspects had improved since the introduction of PEACE, but that further development was necessary. The "preparation and planning" stage (P) of interviews was rated only as adequate and the following shortcomings were highlighted: "interviewers were often unaware of the full circumstances of the incident, did not seem to be aware of the point to prove an offence, and in a number of cases searched or read from statements during the interview" (p. 34). With regard to "engage and explain" (E), little rapport

building was found and the purpose of the interview was clearly explained in only 12% of cases. The "account" stage (A) was generally conducted well with good evidence of active listening skills, although it was noted that challenges of suspects' accounts tended not to be carried out in a professional way, e.g. conducted using evidence. Importantly, Clark and Milne reported that the question types most often employed were "open questions" and "appropriate closed questions" with "leading questions" occurring only rarely. The "closure" stage (C) was often rushed and interviewers tended not to give a good summary of the account provided by the interviewee. (The "evaluate" (E) stage of the PEACE model was not assessed).

Overall, interviews with suspects were rated in the Clark and Milne evaluation as superior to those conducted with witnesses and victims. However, there were some evident skill gaps and little difference between trained and untrained interviewers. Whilst there was some evidence of transfer of skills from training to practice, this pertained more to those aspects of interviewing relating to legal requirements and less to those involving communication skills and development of the interviewee's account. According to Milne and Bull (2003), "interviewers are therefore learning from the course, but are being rigid in the manner in which they interview" (p.119).

## 1.9 What happens in police interviews with witnesses

Despite calls from a number of authors (e.g. Clarke & Milne, 2001; Heaton-Armstrong, 1995; Milne and Shaw, 1999) for interviews with witnesses to be audio or video recorded, this does not currently happen in any country (Milne & Bull, 2003).

Researchers wishing to learn what occurs during interviews with witnesses must obtain special permission from police authorities to observe and/or record interviews and, perhaps as a result of this, there is very little published field research on witness interviewing. Until relatively recently, research in this area had been conducted on a small scale.

Fisher et al. (1987) analysed eleven tape recorded interviews with witnesses conducted by Florida police officers both at the crime scene and the police station. A tendency was noted for interviewers to interrupt witnesses with, often unrelated, closed questions while they were responding to initial open-ended questions. It was also noted that the interviewers tended not to allow pauses or silences during questioning. A further field study by Fisher et al. (1989), as noted above, demonstrated the superiority of the enhanced CI over standard witness interviewing techniques. Within the UK, two prominent studies on police interviews with witnesses were conducted by McLean (1995) and Clark and Milne (2001), and their findings are reported here.

### 1.9.1 McLean's study

In a small scale study, McLean (1995), a Police Inspector with Yorkshire Police, analysed 16 tape-recorded witness interviews and identified some key aspects of poor interviewing practice. The main purpose of his research was to examine what happens during the social exchange between witnesses and police officers that may affect the written statement. The interviewing style was described as directive and was characterised by interruptions, filling of pauses, and change of topic with the interviewer

occupying a disproportionate amount of talking time. Nearly 66% of all questions asked comprised either leading questions or questions of the closed yes/no type and were considered by McLean as "counter productive" or "risky". On comparing the recorded conversation with the resulting written statement from each interview, he found a considerable amount of information provided by the interviewee was absent from the statement. In addition, information contrary to what the witness had said had been included in the statement. McLean concludes that there is a qualitative difference between suspect and witness investigative interviewing with the latter characterised by a more counter-productive questioning style.

### 1.9.2 Clarke & Milne's study

As part of their comprehensive evaluation of the PEACE model of investigative interviewing, Clark and Milne (2001) assessed 75 taped interviews with witnesses. They found the length of interviews to vary widely and, on average, only a quarter of the time was spent actually "interviewing" the person, the remainder of the time spent writing down what the interviewee was saying. Both the "preparation and planning" (P) and "engage and explain" (E) stages of the PEACE approach were considered to be conducted poorly. For example, 26% of interviewers did not introduce themselves at all, 32% gave no explanation of the interview's purpose, and 53% failed to mention that the interview is an opportunity for the witness to give a personalised account.

The evaluation found that interviews were oriented heavily towards statement-taking as opposed to allowing interviewees to provide a free recall of their experience. As such,

interviewers tended to use a closed-questioning style and there was only limited or poor use of either the CI or conversation management approaches which are intended to form an integral part of the "account" (A) stage of the PEACE model. The closure (C) stage too was rated poorly with 59% of officers failing to summarise the statement and 62% not offering the interviewee the option to add or edit the statement. Clark and Milne judged the overall standard of witness interviewing to be significantly lower than that of the interviewing of suspects and they note that, contrary to the purposes of the PEACE training programme, many interviewers seem not to have adopted new interviewing skills. Instead, they "steadfastly cling to the question and answer routine that was described as the standard police interview back in 1992 prior to the advent of PEACE" (p. 113).

## 1.10 Conclusions

Based on the above review of police interrogation and interviewing, and related empirical research, a number of conclusions can be surely drawn. First, the nature of police interrogation practice has not been unchanging over time. Historical perspectives reveal a shift from physical to psychological coercion and, at least in some jurisdictions, a move from confession seeking approaches towards information gathering ones. A concern remains, however, that the use of coercive interrogation tactics may persist in some instances and, while perhaps necessary with some highly resistant or obstructive cases, such tactics are inappropriate and inadvisable with most criminal suspects. Second, in recent years, psychological principles have been applied to interrogation and interviewing. The result has been a greater understanding of the processes leading to

confession, an evidence-based critique of the assumptions underlying some interrogation techniques, and the development of new, more effective methods of interviewing suspects and witnesses. Third, according to some research, the standard of witness interviewing is poor compared to that of suspect interviewing. The reason may be due to failures in transferring training to practice.

Overall, it can be argued that, to date, the application of psychological knowledge has been of benefit to the practice of police interviewing and that there would seem to be scope for further work. The studies in this thesis are experimental in nature and, following recent studies on interrogative pressure and interrogative suggestibility, aim to assess the effects of situational variations in the interrogative context on patterns of interviewee responding. Before introducing a detailed rationale for the present research, it is necessary to consider in closer focus the types of interrogative pressure brought to bear on suspects and witnesses and how these pressures may interact with personality factors and the vulnerabilities of some interviewees. In Chapter 2, therefore, attention is turned to the Gudjonsson and Clark (1986) model of interrogative suggestibility and associated research.

# Chapter 2: Interrogative Suggestibility and Interrogative Pressure

## 2.1 Defining interrogative suggestibility

Susceptibility to suggestion and interpersonal pressure within an interrogative context has been termed "interrogative suggestibility." Gudjonsson (2003) credits Binet (1900, 1905) with the introduction of the concept and with early attempts at its measurement. He also cites as instrumental the experimental work of Stern (1910, 1938, 1939) and notes that both Binet and Stern presented participants with static pictures and asked leading questions about them. This early work demonstrated the propensity of suggestive questioning to distort memory recall and testimony. In their definition of interrogative suggestibility, Powers, Andriks and Loftus (1979) emphasised processes of memory: "the extent to which they (people) come to accept a piece of post-event information and incorporate it into their recollection" (p. 339). Gudjonsson (2003) argues that this definition is limited in two respects. First, post-event information communicated during questioning may be accepted as credible by individuals, but not necessarily incorporated into their recollection. Second, the definition is not sufficiently detailed to generate testable hypotheses.

Gudjonsson and Clark (1986) defined interrogative suggestibility as: "the extent to which, within a closed social interaction, people come to accept messages communicated during formal questioning, as a result of which their subsequent

behavioural response is affected" (p. 84). Implied in this definition are five interrelated components: (i) a social interaction (ii) a questioning procedure (iii) a suggestive stimulus (iv) acceptance of the stimulus (v) a behavioural response.

The first of these reflects the dynamic social nature of police interviews as a changing sequence of social actions between individuals. The research findings cited in the previous chapter on what happens during police interviews illustrate the socially interactive nature of interviewing. The second component relates to the information gathering process by which one or more individuals seek, through questioning, to elicit the memory recollections of a past event or experience from another individual. Gudjonsson (2003) notes that anything interfering with recall can undermine attempts by the interviewer to obtain valid information from the interviewee.

One part of the suggestive stimulus component of Gudjonsson and Clark's definition refers to leading questions. These are questions phrased in such a way as to communicate expectations and premises to interviewees and hence to suggest a particular answer. Importantly, leading questions have been shown to produce distorted responses (Richardson, Dohrenwend & Klein, 1965) and the introduction of post-event information contained in leading questions can significantly affect eyewitness accounts (Ainsworth, 1998; Kebbell & Giles, 2000; Loftus, 1975, 1979; Memon & Wright, 2000).

The fourth component, acceptance of the stimulus, emphasises that individuals must perceive the message in the suggestive stimulus to be credible. Again, Gudjonsson (2003) notes that acceptance of the suggestive information does not necessarily mean that it will be incorporated into memory. Finally, according to Gudjonsson and Clark's definition of interrogative suggestibility, individuals must respond, verbally or non-verbally, to indicate acceptance or non-acceptance of the suggestion.

It should be noted that interrogative suggestibility bears limited relation to traditional types of suggestibility. On the basis of correlational and factor analytical work, Eysenck and Furneaux (1945) proposed two independent types which they termed "primary" and "secondary" suggestibility. Primary suggestibility has been defined as "the uncritical amenability of an individual to outside influences which intimate that a prescribed course of behaviour or action should be followed" (Trippi, 1973, p 220). It is thought to be associated with motor processes and receptivity to hypnotic induction, and is measured using ideo-motor tests such as Hull's (1933) Body Sway test (Gudjonsson, 2003). Secondary suggestibility appears to encompass a more diffuse set of phenomena than does primary suggestibility. According to Eysenck (1947), secondary suggestibility is associated with the variables "indirection" and "gullibility" and he found that it was negatively correlated with intelligence, but not correlated with primary suggestibility. Eysenck (1947) also proposed the concept of "tertiary suggestibility", relating it to attitude change brought about on the basis of persuasion by a prestige figure. While Weitzenhoffer (1953) refers to the "ambiguous" and "in-between" nature of tertiary

suggestibility, and Evans (1967) notes the lack of empirical support for its existence, Gudjonsson (2002) considers it to be similar to interrogative suggestibility.

Gudjonsson (1987a) argues that the key features of interrogative suggestibility which distinguish it from other types of suggestibility are: (i) it involves a questioning procedure within a closed social interaction; (ii) the questions are primarily concerned with past experiences and events, memory recollections and knowledge states. This makes it different from primary suggestibility which is almost exclusively concerned with motor and sensory experiences of the immediate situation; (iii) it has a strong "uncertainty" component which relates to the cognitive processing capacity and functioning of the individual (Gudjonsson & Clark, 1986). Interrogative suggestibility may, however, be related to secondary suggestibility on the basis that at least some of tests considered by Eysenck and Furneaux (1945) to assess this factor – fidelity of report tests used by early investigators such as Binet and Stern – bear some resemblance to the experience of eyewitnesses facing questions about their memory for an event (Baxter, 1990).

## 2.2 The Gudjonsson-Clark theoretical model

Gudjonsson and Clark (1986) developed a detailed social-psychological model of interrogative suggestibility intended to provide "a framework for understanding the process and outcome of police interviewing" (Gudjonsson, 1991, p 280). According to the model, the outcome of interrogations depends on the conditions of uncertainty, interpersonal trust, and expectations of success experienced by the interviewee and the

coping strategies employed by individuals to manage these variables. Gudjonsson and Clark's approach is concerned in part with individual differences in suggestibility and takes into account that responses to police interrogation can vary significantly across individuals. The Gudjonsson-Clark theoretical model is shown in Appendix A.

In terms of the theoretical model, the interrogation situation involves two participants, a witness and an interrogator. (Gudjonsson and Clark note that, in practice, more than one interrogator may be involved in police interviewers. They do not mention the involvement of other agents in the interrogation situation such as a legal representative). The term "witness" is used to denote any person who is subjected to police questioning. As such, the model does not make an explicit distinction between witnesses, suspects and victims (Gudjonsson, 1997).

Gudjonsson and Clark (1986) suggest that police officers enter an interrogation with a particular cognitive set, "an 'event model' ... and seek to extract information consistent with this model" (p. 88). As such, police interviewers may have expectations and premises which influence the form and content of their questioning. To the extent that these expectations and premises are in error and that the process of questioning derives from a too rigid a cognitive set, there exists the possibility that they may bias or cue a witness to respond inaccurately.

Interviewees too enter an interrogation with a "general cognitive set" comprising mood, thinking and expectations, and this is thought to be influenced to some extent by the past experiences of the witness. For example, those who have experienced police interrogation in the past are likely to have a different cognitive set to individuals who are unfamiliar with interrogation procedures (Gudjonsson & Singh, 1984a). The general cognitive set of a witness is also likely to be influenced by their perceptions of and attitudes towards the police, i.e. these may be negative, suspicious and obstructive on the one hand or positive, trusting and cooperative on the other. Based on their particular general cognitive set, witnesses will adopt a "general coping strategy" to deal with the interrogation which can facilitate either a suggestible or resistant set of responses.

According to Gudjonsson and Clark, three components are essential to the process and mechanism of interrogative suggestibility. The first of these, uncertainty, occurs when the interviewee is not sure of the correct answer to a question. This is more likely in cases where an interviewee's memory for an event is poor, perhaps because of a long delay between encoding and retrieval, and/or due to limited cognitive processing capacity. The second necessary component is interpersonal trust. This reflects the degree to which the interviewee perceives the intentions and behaviour of the interrogator to be genuine and without trickery. Interviewees who are suspicious of the interrogator's motives (e.g. they may perceive questions to be overtly misleading) are less likely to yield to suggestions.

The components of uncertainty and interpersonal trust may not be sufficient on their own to elicit suggestible responses from interviewees. This is because, faced with uncertainty over an answer, interviewees can simply state that they "don't know". Hence, the third key component of interrogative suggestibility is expectation of success. Interviewees may believe that it is necessary to provide an answer and/or that they should know the answer to questions asked by the interrogator. This aspect of suggestibility relies on the reluctance of interviewees to explicitly state or acknowledge to themselves their lack of certainty. It is the manipulation of these three components by an interrogator together with an interviewee's cognitive appraisal and coping strategies that influences an interviewee's susceptibility to suggestions.

Finally, Gudjonsson and Clark's model incorporates both the "leading questions" and "negative feedback" aspects of suggestibility (Gudjonsson, 1983, 1984). The effects of leading questions have been extensively researched (e.g. Loftus, 1979) and, as previously mentioned, are those which have the potential to distort interviewee responses and hence make testimony less accurate. Within the model, negative feedback is, "a signal communicated by an interrogator to a witness, after he/she has responded to a question or a series of questions, intended to strengthen or modify subsequent responses of the witness" (Gudjonsson & Clark, 1986, pp. 93-94). Because the studies presented in this thesis are concerned with the negative feedback aspect of interrogative suggestibility, this will be considered in more detail in subsequent chapters.

## 2.3 Measuring interrogative suggestibility

### 2.3.1 Rationale for the development of an Interrogative Suggestibility Scale

Courts of law have reasonable concern for the accuracy of verbal accounts they receive from victims, witnesses and suspects. One threat to the quality of such evidence is the fallibility of human memory, a phenomenon well documented in the literature (e.g. see Bartlett, 1932; Clifford & Bull, 1978; Crombag, Wagenaar & van Koppen, 1996; Cutler & Penrod, 1995; Jelicic et al., 2006; Johnson, Hashtroudi & Lindsay, 1993; Loftus & Palmer, 1974; Smith, 1930). Another important issue is the reliability of evidence given during police questioning by individuals who may be particularly susceptible to suggestion. For example, Gudjonsson and Gunn (1982) reported on the case of a mentally handicapped victim of a sexual assault. The authors were asked by the Director of Public Prosecutions to assess whether the victim was competent as a witness and, if so, how reliable she was, i.e. what was the victim's general level of suggestibility. Their psychological assessment showed a high level of suggestibility and it was concluded that this would make some of her evidence unreliable. Experience of this and similar cases prompted Gudjonsson to develop a clinical instrument to "measure objectively the vulnerabilities or proneness of people to give erroneous accounts when interviewed" (Gudjonsson, 1997, p ix).

### 2.3.2 Requirements of the Gudjonsson Suggestibility Scale

The Gudjonsson Suggestibility Scale (GSS 1) (Gudjonsson, 1983, 1984) was developed for clinical and forensic purposes to measure interrogative suggestibility in potential witnesses. It was intended that the scale would primarily measure the extent to which

interviewees gave in to leading questions, and the extent to which they gave in to interrogative pressure in the form of negative feedback. It was considered important that the true purpose of the test was not immediately apparent and, to increase the scale's reliability, that a broad range of scores would be obtained from the general population, forensic populations and people with intellectual disabilities. Finally, given that testing would often take place in forensic settings, it was deemed necessary that administration of the scale was quick and simple (Gudjonsson, 1997). A parallel form of the scale (GSS 2) (Gudjonsson, 1987b) was later introduced to allow repeated assessments and examination of the test-retest reliability of suggestibility. The two versions of the scale differ only in the content of their narrative paragraphs and interrogative questions. They are identical with regard to format, administration and scoring. Normative data have been presented for UK subjects (Gudjonsson, 1997) and more recently for a US sample (Pollard et al., 2004). Inter-scorer reliability has been shown to be very high for both the GSS 1 (Richardson & Smith, 1993) and the GSS 2 (Clare, Gudjonsson, Rutter & Cross, 1994). Recent research has investigated the viability of a shortened version of the GSS (Smeets, Leppink, Jelicic & Merckelbach, in press) and a computer-administered version (Gorassini, Harris, Diamond & Flynn-Dastoor, 2006).

### 2.3.3 GSS procedure and specific measures obtained

In terms of Gudjonsson and Clark's (1986) model of interrogative suggestibility, the key outcome measures of the GSS 1 and GSS 2 are "Yield", the number of suggestive questions that interviewees give in to, and "Shift", an interviewee's response to negative feedback and repeated questioning. In the GSS procedure, a narrative paragraph is read

out to the interviewee or played from a tape recorder. According to Gudjonsson (1997), the stories were constructed in a similar way to those used in memory tests such as the Wechsler Memory Scale (Wechsler, 1981). The interviewee is asked to verbally recall as much as possible about the narrative and the number of "ideas" recalled gives an Immediate recall score. A further Delayed recall score is obtained 50 minutes after initial recall.

The interviewee is then asked 20 specific questions about the narrative, 15 of which are suggestive questions of three types. The first type is "leading questions". For example, question 20 on the GSS 1, "Did the woman's clothes get torn in the struggle?" includes the premise of a struggle and makes an affirmative answer plausible. The second type is "affirmative questions". These questions mislead by introducing sufficient doubt about an aspect of the story and by prompting an affirmative response. Question 14 on the GSS 2, "Did the boy commonly ride the bicycle to school?" is an affirmative question. The third type of suggestive question used on the GSS 1 and 2 is "false alternative questions". These imply the presence of information not contained in the story and suggest that interviewees choose between two false answers. An example is question 19 on the GSS 1, "Were the assailants armed with knives or guns?" Gudjonsson (1997) notes that the distinction between "leading questions" and "affirmative" types of suggestive questions is of limited value and, in practice, the term "leading question" is commonly used to refer to each of the three types. Yield 1 is the number of leading questions that the interviewee yields to on the first round of questioning.

Following this initial questioning phase, negative feedback is communicated to the interviewee as follows: "You have made a number of errors. It is therefore necessary to go through the questions once more, and this time try to be more accurate". The 20 questions are asked again allowing a further score, Shift, to be calculated. This is the number of answers changed from the initial questioning phase. Gudjonsson (1997) states that a change in answer must be "distinct" in order to be scored as Shift. Examples include: "Yes" to "No" or vice versa; "One child" to "Two children"; "Tall" to "Medium". According to scoring guidelines for the GSS 1 and 2, the following would not be scored as Shift: "Not sure" to "No"; "Would have" to "Yes"; "Think so" to "Yes".

Yield 2 refers to the number of leading questions which the interviewee gives in to after the negative feedback has been administered. Total Suggestibility is the sum of Yield 1 and Shift and is assumed to indicate the interviewee's overall level of suggestibility. Lastly, a Confabulation score can also be calculated comprising the number of "distortions" in the story's content and the number of pieces of information or "fabrications" which have been added (Clare et al., 1994). An example of a distortion to the GSS 1 narrative is: "Some men came along". This represents a major change in the details of an existing "idea" in the original story. An example of a fabrication is: "The assailants hit her". Such a statement introduces new information actually absent from the narrative. A tabular illustration of the GSS 1 and 2 procedure and the scores derived

from the scales is provided in Appendix B.

## 2.4 Research on interrogative suggestibility and interrogative pressure

As well as being developed for clinical and forensic purposes, the GSS 1 and GSS 2 were designed as research instruments to investigate in detail the nature and mechanisms of interrogative suggestibility. Much of the research using the scales has aimed to establish the relationship between interrogative suggestibility and a range of other psychological constructs. For example, scores on the GSS 1 and 2 have been shown to correlate negatively with intelligence. Gudjonsson (1983) found this relationship between GSS 1 scores and full-scale IQ as measured by the short Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1981). In a later study, Gudjonsson (1988a) reported that intelligence correlated better with suggestibility among participants of below average intelligence (IQ below 100) than those above (IQ of 100 or above). Testing a US sample on the GSS 2, Pollard et al. (2004) also found a negative relationship between suggestibility and intelligence as measured by the Shipley Institute of Living Scale (Zachary, 1986). Suggestibility scores tend also to correlate negatively with immediate and delayed memory recall (e.g. Gudjonsson, 1983). In a recent evaluation of Polish versions of the GSS 1 and 2, Polczyk (2005) found negative correlations between suggestibility scores and memory as well as suggestibility and general intelligence as measured by Raven's Progressive Matrices (RPM; Raven, Court & Raven, 1983).

Consistent with the Gudjonsson and Clark (1986) theoretical model, Gudjonsson

(1988b) found a positive correlation between anxiety and interrogative suggestibility. In that study, high scores on the GSS 1 were associated with high scores on the Spielberger (1969) State Anxiety Inventory. Gudjonsson argues that state anxiety, i.e. transitory, situational stress, is more relevant to suggestibility than general, or trait anxiety. Indeed, a number of studies have found trait anxiety to correlate poorly with suggestibility (e.g. Gudjonsson, 1983; cf. Gudjonsson, Rutter & Clare, 1995). Gudjonsson (1988b) notes also that correlations between the suggestibility measure Shift and state anxiety were particularly high after the delivery of negative feedback. This supports Gudjonsson's (1984) view that the Shift aspect of interrogative suggestibility is more associated with anxiety and coping strategies than is Yield 1. (See also Bain & Baxter, 2000).

Studies have found high scores on the GSS to be associated with low self-esteem (Singh & Gudjonsson, 1984), emotion-focused, as opposed to problem-focused, coping styles (Howard & Hong, 2002), low assertiveness (Gudjonsson, 1988b), external locus of control (Gudjonsson & Lister, 1984), field dependence as opposed to field independence (Blagrove, Cole-Morgan & Lambe, 1994; Singh & Gudjonsson, 1992), and most recently, high self-monitoring, as opposed to low self-monitoring styles (Bain, Baxter & Ballantyne, 2007). As Eysenck and Furneaux (1945) would have predicted, Register and Kihlstrom (1988) found the interrogative suggestibility of a group of college students was not related to their level of hypnotizability as measured by the Stanford Hypnotic Susceptibility Scale, Form C (SHSS-C; Weitzenhoffer & Hilgard, 1962). This finding supports Gudjonsson's (1987a) argument that interrogative suggestibility is not related

to other types of suggestibility.

### 2.4.1 Additional influences on interrogative suggestibility

As noted, according to the Gudjonsson and Clark model, levels of suggestibility measured by the GSS depend upon three interrogator-led influences: the delivery of negative feedback, the use of leading questions, and the repetition of questions. Register and Kihlstrom (1988) highlighted the importance of repeated questions in particular as an implicit form of negative feedback. Their study used an amended version of the GSS 1 procedure in which no negative feedback was delivered. Despite the absence of feedback, participants changed a number of their answers during requestioning. It may be that the repetition of questions without explanation communicates to participants that their previous responses are in some sense incorrect or inappropriate and should be changed. Another possibility is that inconsistencies between first and second answers occur simply as a result of memory failure (Gudjonsson, 2003).

Other research studies using the GSS have identified additional influences beyond those directly cited in the model which affect interviewees' susceptibility to interrogative pressure. One such influence is the firmness with which negative feedback is communicated. In a study by Baxter and Boon (2000) using the GSS 2, participants received negative feedback delivered by interviewers adopting one of three demeanours: friendly, firm or stern. Scores on the measures Yield 2 and Shift increased across the three conditions according to whether the delivery style was a friendly, positive demeanour, a neutral demeanour, or one which was unfriendly and critical.

Another important influence on interrogative suggestibility, therefore, is the behaviour of the interviewer. Bain and Baxter (2000) tested a sample of students on the GSS 1 under two different conditions of interviewing style. Participants interviewed in an "abrupt" manner scored higher on the measures Shift and Total Suggestibility than those interviewed in a "friendly" manner. Baxter, Jackson and Bain (2003) and Bain, Baxter and Fellowes (2004) manipulated interviewer behaviour in a similar way and achieved broadly comparable results. While abrupt or stern interviewing styles raise suggestibility scores above population norms for the GSSs, relaxed and friendly styles lower scores below the norms and the GSS measure most affected by these variations tends to be Shift. That Shift rather than Yield is influenced by interviewer behaviour confirms Gudjonsson's (1984) view that the Shift measure is more sensitive to situational influences while Yield is related more to cognitive processes (see Singh & Gudjonsson, 1984).

The authors suggest that many interviewees facing an abrupt as opposed to a friendly interviewer will perceive a degree of "psychological distance" (cf. Gudjonsson & Lister, 1984) between themselves and the interviewer and will experience increased uncertainty and anxiety. In an attempt to reduce psychological distance and maintain self-esteem, they will direct attentional resources towards their feelings and away from the cognitive processes involved in accurate recall. The significance of these findings is firstly that they highlight the importance of the social dynamics of interrogative suggestibility

which are implicit in the theoretical model, and secondly that severe or unfriendly interviewing styles have the potential to bias or distort the responses of real witnesses by pressuring them to change what may be "true" answers, towards inaccuracy.

The suggestive effect of leading questions has been shown to be reduced by exposure to a warning about the possible presence of misinformation in the questioning procedure. Greene, Flynn and Loftus (1982) presented participants with slides depicting a crime followed by written post-event information which was misleading. Some of the participants were warned that they might receive misinformation. When recall accuracy was later tested, participants who had been exposed to a warning prior to the misinformation exhibited more resistance to the suggestive stimuli than participants who had received no warning. The authors found that the warning did not increase depth of processing or rehearsal of the event. Rather, participants receiving a warning appeared to more closely scrutinise the post-event information and, as a result, were more accurate when recalling details of the event.

Similar work has been conducted using the GSSs. Boon and Baxter (2000) found that all four key measures on the GSS 2 (Yield 1, Yield 2, Shift and Total Suggestibility) were significantly reduced by warnings about the presence of misinformation in the GSS questions. Participants in their study did not exhibit complete resistance to the leading questions and interrogative pressure, however, suggesting that beyond the situational influence of the interviewer, there may exist some "core" suggestibility within

individuals related to cognitive processes such as attention and memory. Bain et al. (2004) found that two measures on the GSS 1 (Yield 1 and Total Suggestibility) were affected by warnings. They note that, of the GSS measures, Yield 1 is likely to be the most sensitive to warnings about misleading information (Boon & Baxter, 2000; Boon & Baxter, 2004).

The discussion thus far has focused on interrogator-led influences on interrogative suggestibility as well as individual differences in the personality and ability of interviewees which may shape patterns of responding. It should not be overlooked that levels of suggestibility as measured by the GSSs are also under direct control of the participant. Baxter and Bain (2002) demonstrated this experimentally. In their study, half of the participants were administered the GSS 1 in the standard manner. The other half of the sample was instructed to appear gullible or susceptible to interrogative pressure, i.e. to fake suggestibility. As predicted, participants in the faking condition obtained a pattern of scores on the GSS 1 which distinguished them from the control group: faking participants scored significantly higher on the Yield 1 measure. It seemed that the faking participants recognised the need to give in to the suggestions contained in leading questions. They did not, however, identify the need to shift their responses. Baxter and Bain argued that people who are genuinely vulnerable to interrogative pressure would be likely to show an increase in both Yield 1 and Shift and hence they concluded that the measure of Shift is the most sensitive in the detection of genuine vulnerability.

In a further study, Woolston, Bain and Baxter (2006) replicated the above finding. They also found that participants who were instructed to be generally compliant with the demands of the interview, i.e. to give in to whatever the interviewer seemed to want, scored significantly higher on all of the GSS 1 measures. Hence, Woolston et al. argued that compliant patterns of responding may strongly resemble those of genuine vulnerability. Together, these two studies demonstrate that the pattern of suggestibility scores on the GSSs can differ not only in accordance with the vulnerability of participants, but also in line with different faking strategies and differences in motivation held by participants.

### 2.4.2 Studies using adapted GSS procedures

As previously noted, some studies using the GSSs have amended parts of the procedure in order to more closely examine the effects that such changes have on levels of interrogative suggestibility. Register and Kihlstrom (1988) eliminated the negative feedback component in their study and, as a result, demonstrated the pressuring effects of requestioning alone on interviewees' tendencies to shift their initial answers. Boon and Baxter (2000) showed that removal of negative feedback lowers Total Suggestibility scores significantly below those obtained using the standard GSS 2 procedure.

Adapting the GSS procedure in certain ways may improve its ecological validity as a research tool. From the perspective of the person being administered the scale, the GSS

procedure is analogous, at least to some extent, to the experiences of witnesses, suspects or victims who undergo real police interviewing procedures: the scale administrator adopts a role similar in many respects to that of a police interviewer and, as with real interviews, administration of the scale is social and dialogical in nature. It involves a verbal questioning process in which specific memories of interviewees are under examination and they are required to respond verbally to the demands of a questioner. However, it can be argued that the standard GSS procedure, although similar in a number of aspects to real interviewing practice, is sufficiently different in key respects to limit the generalisability of research findings based on its use. In particular, the type of questions employed in the GSS procedure and the use of a verbally presented story may, as discussed below, be regarded as problematic.

In a recent study, Baxter, Boon and Marley (2006) adapted the GSS 2 procedure to comprise only neutral or "minimally leading" questions (as opposed to the overtly leading questions employed in the GSS). These questions addressed the GSS narrative as usual, but did not include the overtly leading information contained in the standard questions. Baxter et al. avoided the term "non-leading" noting that any question may hold implications and so may have the potential to "lead". Following Baxter (2004), they argued that interview-based evidence obtained through the use of leading questions may be challenged by defence lawyers in court on the basis that it has been obtained using manipulative interviewing techniques. (See also Baxter, Bain & McAusland, 2007). To the extent that the use of overtly leading questions is recognised by professional

interviewers as counterproductive and, as a result, less likely to occur in real interviews, it would seem appropriate that applied research designs take this into account.

Another aspect of the GSS procedure which would seem to limit its capacity to simulate real forensic interviews relates to the stimulus about which participants are questioned. The GSS tests memory for a narrative passage which is either read out or played as an audio recording. In contrast, real witnesses are questioned about past events which they have personally experienced. The key distinction is the type of memory under examination in each case. Recall of the verbal material presented in the GSS narrative involves retrieval from semantic memory and so is related to general factual knowledge about the world and language, including memory for words and concepts. In contrast, recall of a past event places demands on autobiographical episodic memory, i.e. the system of long-term memory concerned with specific experiences and events which have occurred in particular places and at particular times (Beail, 2002; Cardone & Dent, 1996; Scullin & Ceci, 2001). (Tulving's (1972) theoretical argument that these two types of memory are distinct is supported by neuropsychological evidence obtained from case studies of patients with neocortical lesions and amnesic syndrome (see Tulving, 1989; Warrington, 1986)).

## 2.5 Aims of the thesis

The experimental studies conducted for this thesis follow those cited immediately above in seeking to further develop the Gudjonsson and Clark theoretical model. The interviewing procedure employed is adapted from and retains the question-feedback-

requestion features of the GSSs. In the interests of ecological validity, however, the research presented here departs from most previous work in two key respects. The questioning procedure turns the focus away from overtly leading questions and towards interpersonal pressures often present in forensic interviews which have received little or no attention in previous studies. In addition, questioning in the present research is based on a more complex stimulus than that used in the GSS procedure and hence is more likely to elicit retrieval from episodic rather than semantic memory. The first aim was to investigate situational determinants of interrogative suggestibility, namely the type of verbal feedback communicated to interviewees and the general manner of the interviewer. A second aim was to examine the relationship, if any, between the tendency to yield to interrogative pressure and the possible pressuring effects of the presence of others in the interrogative context. This aspect of the research was intended to model the involvement of an additional interviewer and interviewee ally, such as a legal representative. The third aim was to explore the influence on suggestible responding of varying levels of interviewee self-esteem. A final aim was to further investigate the relationship between interrogative pressure and anxiety.

# Chapter 3: Study 1

## 3.1 Abstract

Much experimental research on interrogative pressure has concentrated on the effects of leading questions, and the role of feedback in influencing responses in the absence of leading questions has been neglected by comparison. This study assessed the effect of negative feedback and the presence of a second interviewer on interviewee responding in a simulated forensic interview. Participants viewed a videotape of a crime, answered questions about the clip, and were requestioned after receiving feedback. Compared with neutral feedback, negative feedback resulted in more response changes, higher reported state anxiety and higher ratings of interview difficulty. These results are consistent with the predictions of Gudjonsson and Clark's (1986) model of interrogative suggestibility. The presence and involvement of a second interviewer did not significantly affect interviewee responding, although trait anxiety scores were elevated when a second interviewer was present. The theoretical and applied implications of these findings are considered.

## 3.2 Introduction

"Interrogative pressure" (IP) is any influence applied to interviewees during an interrogation which may interfere with their attempts accurately to recall an event. One such influence is negative feedback - a communication perceived by interviewees to mean that their answers, or they themselves, are in some sense "unsatisfactory". IP may imply or require that the interviewee should produce more or different information.

"I think you're, I think you're telling us a lie, John, and it's time, you know, that we get down to the nitty gritty of this thing" (cited by Ofshe & Leo, 1997, p. 224)

"...It's just not on ... Look, I'm getting fed up with this because it just isn't on ... You're not telling me – I don't understand why. Tell me who was with you" (cited by Baldwin, 1993, p. 347)

The above interviewer statements were drawn from interviews with suspects and it may be that such direct negative feedback is less prevalent in witness interviewing. However, negative feedback can be implicit in the repetition of questions (Linton & Sheehan, 1994; Register & Kihlstrom, 1988), or in an unsupportive or disapproving interviewer manner (Bain & Baxter, 2000; Baxter & Boon, 2000). It may be that forensic interviewers should be as aware of the possible distorting effects of negative feedback, whether overt or implicit, as they should be of the effects of leading questions (Baxter et

al., 2006; Loftus, 1979; see also Ainsworth, 1998; Gudjonsson, 2003; Memon & Wright, 2000).

Gudjonsson and Clark (1986) argued that interviewees receiving negative feedback will appraise it, then reject or accept it. Rejection of negative feedback may lead interviewees to resist further suggestions because they distrust their interviewers. There is evidence to suggest that rejection of negative feedback may be more likely in a forensic population. For example, Gudjonsson and Singh (1984a) investigated the relationship between interrogative suggestibility and number of previous convictions among a sample of delinquent boys. They found that number of convictions correlated negatively with the tendency to give in to IP in the form of negative feedback. The authors suggest that those with experience of interrogative procedures may learn to resist attempts to pressurise them and that repeat offenders may be characteristically more prone to resist IP than those who offend infrequently.

According to the model, acceptance of negative feedback commonly causes emotional and physiological reactions, specifically an increase in anxiety and temporarily lowered self-esteem. Interviewees also experience uncertainty: often, they do not know with any certainty the "right" answer to a question. As a result, they subordinate their personal judgement and, instead of relying on their own ability to remember, they rely on external cues such as the perceived demands of the interviewer. Negative feedback is also likely to increase the "psychological distance" between the interviewer and interviewee

(Baxter & Boon, 2000; Gudjonsson & Lister, 1984). Interviewees who feel socially isolated may be more anxious and may try to appease interviewers, at the expense of attending to the task of accurate recall. Negative feedback can evoke a suggestible coping strategy in interviewees (cf. Baxter et al., 2003; Howard & Hong, 2002, Gudjonsson & Clark, 1986), characterised by susceptibility to further IP and a tendency to change initial responses.

Some studies investigating the effect of negative feedback on response change have relied on the Gudjonsson Suggestibility Scales (GSS 1 & 2) (Gudjonsson, 1984, 1987b). In this procedure, interviewees hear a spoken narrative, recall it, and are then asked 20 questions, 15 of which are leading. The scales measure interviewee tendencies to change or "Shift" previous answers in response to negative feedback and to "Yield" to implications embedded in leading questions. Using the GSS, Tata and Gudjonsson (1990) found that the type of verbal feedback used affected the mood of interviewees and their tendency to shift previous responses: verbal negative feedback, shown below, increased response change and acceptance of material presented in leading questions.

"That wasn't very good. You have made quite a number of errors. Your memory was poorer than I had expected and worse than that of most other people. I am rather disappointed in you. I hope you can do better when we repeat the questions now."

The positive feedback was worded as follows:

"You did really well. Your memory was, in fact, much better than that of most people I've seen. I really am very pleased. It's gratifying to find a subject as good as you. You may in fact do even better when we repeat the questions now."

Participants receiving negative feedback showed higher GSS scores than those receiving positive feedback, at initial testing and at a one week follow-up.

Tata and Gudjonsson also found that negative feedback led to higher levels of reported anxiety, depression, and hostility. Singh and Gudjonsson (1984) found that participants' Yield scores on the GSS 1 increased significantly on re-testing one week after listening to the narrative. Interestingly, from the point of view of the present study, they found a slight decrease in response change after this delay, although this may have been because, as Singh and Gudjonsson argued, the demand characteristics of the procedure may have become familiar and so less credible. Singh and Gudjonsson argued that, generally, because participants' memories for the narrative will tend to deteriorate over a week, they will tend to become more susceptible to cues provided by the interviewer, i.e. their uncertainty increases (Gudjonsson & Clark, 1986).

The GSS procedure includes leading questions and so does not assess the impact that

negative feedback alone may have on interviewee responding. Baxter et al. (2006) adapted the GSS 2 to comprise only "minimally leading" questions, i.e. the overtly leading aspects of the GSS 2 questions were removed. They reported that negative feedback alone continued to be associated with response change, although this change was less than the GSS norm as would be expected since the component of IP associated with leading questions was absent.

This finding may be of some significance. Leading questions are sometimes used in police interrogations (Baldwin, 1993; Bull & Cherryman, 1995; Ceci & Bruck, 1995; McLean, 1995; Pearse & Gudjonsson, 1996, 1999), but their use is widely recognised as problematic (Fisher, 1995; Fisher & Geiselman, 1992; Griffiths & Milne, in preparation), particularly if a witness has some psychological vulnerability (Gudjonsson, 2003; Redlich, 2004). If use of leading questions is decreasing in forensic interviewing, there would appear to be value in examining the effect of IP in the absence of such questions. Also, use of a spoken narrative as a stimulus may limit the generalisablity of findings based on the GSS to the dynamics of real police interviews concerned with events.

One aim of the present study was to examine the effects of negative feedback on recall of a videotaped event. Use of such material is common in eyewitness testimony research (e.g. Lane, Mather, Villa & Morita, 2001; Memon, Hope & Bull, 2003), but has not been usual in research specifically concerned with the effects of feedback. In common with

Baxter et al. (2006), the questions asked in this study were not overtly-leading. Otherwise, the basic procedure was similar to that of the GSS. A further aim was to investigate differences in IP associated with two types of verbal feedback: neutral and negative.

Most research on response change in interviews has involved one interviewer and one interviewee. However, police interviews often involve two interviewers and a single interviewee. Surveying 1067 police interviews, Moston et al. (1992) reported that two officers were involved in around 75% of them. Similarly, Baldwin (1993) reported that interviewers generally worked in pairs, suggesting that the psychological impact of a second interviewer is worth investigating. (The situation in the United States is different. In his observation of police interrogations, Leo (1996c) noted that two detectives conducted questioning in only 31% of cases).

The presence of others can facilitate or inhibit task performance, depending on the task. Generally, the presence of others improves performance of well-learned and/or easy tasks but inhibits performance otherwise (see Bond & Titus, 1983; Geen, 1989; Guerin, 1993; Zajonc, 1965). The task of a witness or suspect under interrogation is often difficult and, excepting cases of frequent offenders, unlikely to be well-learned. As such, it seems reasonable to assume that, in an interview with two interviewers, the performance of an interviewee might be inhibited. The presence of a second interviewer may provide a special type of IP by increasing the cognitive demands of the

interrogation to the extent that the interviewee's cognitive resources are diverted from the task of accurate recall: the interviewee might experience increased uncertainty, perceived lack of control, and greater need to succeed (Gudjonsson & Clark, 1986; Gudjonsson & Lister, 1984). The presence and involvement of a second interviewer may also create a social pressure on the interviewee. Through active, visible co-operation with the principal interviewer, the second interviewer may reinforce the demands of the principal interviewer by apparent acquiescence to these.

This study had two main hypotheses. Firstly, it was predicted that negative feedback would be associated with more response change than neutral feedback, and be associated with higher interviewee state anxiety and ratings of interview difficulty. Secondly, it was predicted that participants interviewed by two interviewers would score more highly on these measures than participants interviewed by a single interviewer: i.e. scores on these measures would be lowest for a neutral feedback/one-interviewer group, higher for neutral feedback/two-interviewer and negative feedback/one-interviewer groups, and highest for a negative feedback/two-interviewer group.

## 3.3 Method

### 3.3.1 Design

The study had a two-factor, between participants design. The independent variables were type of feedback (neutral or negative), and number of interviewers (one or two).

### 3.3.2 Participants

Eighty eight students participated (48 males, 40 females), including Arts and Engineering students, and mature students attending evening programmes. Ages ranged from 18 to 72 years (Mean = 31.49, SD = 17.75). Twenty two participants were randomly assigned to each of the four experimental conditions. Age groups were equally represented in each condition. Participants' names were entered into a draw for a cash prize.

The primary interviewer was a 48 year old male; the second was a 34 year old male. Both dressed formally.

### 3.3.3 Materials

#### 3.3.3.1 Videotaped event

The stimulus was a 77 second videotaped clip, designed to contain approximately equal amounts of event, speech, and appearance information. It showed the theft of a briefcase in a street, involving three main characters, a woman and two men, who conversed, and two male passers-by, one of whom took the case without its owner noticing.

#### 3.3.3.2 State and Trait Anxiety Inventory (STAI)

This instrument consists of two separate 20-item self-rating scales, one measuring state anxiety and the other trait anxiety (Spielberger, 1983; Spielberger, Gorsuch & Lushene, 1970). Trait anxiety has been conceived as a relatively stable personality trait consisting of feelings of apprehension, tension, and increased activity of the autonomic nervous

system. Individuals high in trait anxiety perceive more situations as threatening or dangerous than persons low in trait anxiety (Spielberger, 1972). In contrast, levels of state anxiety are thought to fluctuate within individuals according to the stressors perceived by them. State anxiety is high when individuals perceive subjectively a threat to self, and lowest in the absence of perceived stress or danger (Barnes, Harp & Sik Jung, 2002).

#### 3.3.3.3 Interview Rating Form

This 10-item questionnaire asked participants to rate their: ease of recall, distraction, confidence, influence, comfort, concentration, reasonableness of questions, confidence in answers, experience of pressure, and overall task difficulty - experienced during the questioning phase of the experiment - on a 5-point Likert scale, ranging from 1 – "not at all" to 5 – "very". The Interview Rating Form is shown in Appendix I.

### 3.3.4 Procedure

Participants were tested individually in a room containing a desk and three chairs, a television and video recorder, and a video camera. They were first asked to read and sign a consent form. They were then asked to watch the videotaped event, after which they sat at the desk, facing the interviewer across from them. Each interview was video recorded.

After playing the video the interviewer said to participants: "Tell me everything you remember about the scene you witnessed on the video". Following this free recall,

twenty nine questions were asked about the videotaped scene. The questions, shown in Table 3-1 below, addressed those details of the event likely to be forensically relevant. Twenty-two of the questions were closed (e.g. "Was he wearing a jacket?"; "Did he leave the scene to the left or to the right?"). The remaining seven questions were general (e.g. "What did he say?") and open-ended specific questions (e.g. "How tall was he?"). None of the questions was overtly-leading.

**Table 3-1.** Questions asked after participants had watched the video clip.

| |
|---|
| **Second man** |
| 1. The second man to appear on the scene: What age was he? |
| 2. Was he wearing a jacket? |
| 3. Did he have dark or light hair? |
| 4. How tall was he? |
| 5. * What did he say? |
| 6. What accent did he have? |
| 7. Could he have seen who took the case? |
| 8. Could he have taken the case? |
| 9. Did he leave the scene to the left or to the right? |
| **By-passers general** |
| 10. * Did you see anyone walk past in the background? |
| **By-passer 1** |
| 11. The (first) by-passer in the background: What age was he? |
| 12. Was he wearing a jacket? |
| 13. Did he have short or long hair? |
| 14. Was he carrying anything? |
| 15. * Did he say anything? |
| 16. Could he have taken the case? |
| 17. Did he leave the scene to the left or to the right? |
| **By-passer 2** |
| 18. The second by-passer in the background: What age was he? |
| 19. Was he wearing a jacket? |
| 20. * Did he have short or long hair? |
| 21. How tall was he? |
| 22. Did he say anything? |
| 23. Could he have taken the case? |
| 24. Did he leave the scene to the left or to the right? |
| **The case** |
| 25. * Did you see the case? |
| 26. Did you see the case being removed? |
| 27. Where was the case when it was stolen? |
| **The woman** |
| 28. Could the woman have seen who took the case? |
| **The thief** |
| 29. Who is most likely to have taken the case? |

Note: Asterisk denotes questions asked by second interviewer.

As would be the case in a real forensic interview, participants were asked only those questions relevant to their accounts. Most participants (N = 53) reported that they saw two by-passers in the background and so were asked the questions relating to them. The remaining participants (N = 35) stated that they saw only one by-passer or none. In these instances, the interviewer proceeded to the next appropriate question in the question set.

In the conditions with two interviewers, the second interviewer sat next to the primary interviewer and faced the participant at an angle. The second interviewer asked every 5$^{th}$ question. The interviewers' manner throughout the experiment was intended to be formal; neither overtly friendly, nor too abrupt. Apart from instructing and questioning them, interviewer communication with participants was minimal.

After questioning, participants received either negative or neutral feedback. In the negative feedback conditions, the principal interviewer briefly consulted some papers and said firmly, "From my records here I see that others we've asked about this have done better than you. I'd like you to try again, to see if you can do better." In the neutral feedback conditions the interviewer said, "Thank you for answering these questions. To ensure we have your answers recorded correctly, we'll run through the questions once more." The questions were then repeated. Participants then completed the State Anxiety Inventory, the Interview Rating Form, and the Trait Anxiety Inventory.

Finally, participants were advised that a full written debriefing would be made available once all the data had been collected. It was explained to participants in the conditions involving negative feedback that this did not reflect on their performance and was merely a standard feature of the experiment.

#### 3.3.4.1 Scoring

Recall was a secondary measure, intended only to allow a basic comparison between the event memory available to each group. For simplicity of scoring, it was based on answers given during the first round of questions, rather than on participants' free recall. For the majority of questions, participants received one point for each correct answer. Five of the twenty nine questions asked for an estimate of either height or age. Scores on these questions ranged from 0 to 10 according to accuracy of the estimate, 10 being completely accurate. It was not possible to calculate the difference between actual and estimated age due to the manner in which the majority of participants answered such questions. With few exceptions, participants tended to give rough estimates of ages (and heights) rather than stating a precise figure, e.g. "Early 20s" or "Mid to late 30s".

One question asked what one of the actors had said. This question was scored from 0 to 6 according to how many elements of detail were recalled from the dialogue. To the question "Did you see anyone walk past in the background?" a score of 0 was given for the responses "No" and "Don't know"; 1 for, "Yes, one person", and 2 for, "Yes, two people". The maximum possible recall score was 80. To assess the reliability of recall scoring, specifically of the general and open-ended questions, an independent rater

scored all answers to the first round of questions across all 88 interviews. The rater was provided with the video recorded interviews and coding frame. Otherwise she was experimentally blind.

The possible post-feedback response change scores ranged from 0 to 29. Responses were considered "changed" if the second answer was markedly different from the first. Examples of such changes are: "Yes" to "No" and vice versa; "Scottish" to "English"; "Don't know" to "Yes"; "Right" to "Left"; "He thought it was the gasket" to "It was an electrical fault"; and "Early thirties" to "Between twenty and thirty". The following are examples of second-round answers not considered to be response changes: "Can't remember" to "I think so"; "Twenty five to thirty" to "Twenties"; "5'6, 5'7" to "5'8"; "Shortish" to "Medium"; and "Left" to "Not sure, possibly right".

The number of responses changed in the direction of inaccuracy was also scored. Participants received one point each time this occurred. Examples of responses changed toward inaccuracy are: "Dark" to "Light" (where the correct answer was "Dark"); "Right" to "Left" (where the correct answer was "Right"); and "One of the by passers" to "Don't know" (where the correct answer was "One of the bypassers"). Further examples are provided in Appendices E and G.

## 3.4 Results

### 3.4.1 Memory recall

Inter-rater reliability was assessed for the general and open-ended items (7 of the 29 questions). Cohen's Kappa coefficients for each question indicated a high level of agreement between the recall ratings: Q1 = .9; Q4 = .91, Q5 = .93; Q10 = .96; Q11 = .96; Q18 = .87; Q21 = .94. The mean recall scores in Table 3-2 are for both closed and general/open-ended items. Separate analyses of these measures, using two-way between-subjects ANOVAs, showed no significant differences or interactions between groups.

### 3.4.2 Anxiety

The alpha coefficients of the state and trait anxiety assessments were .92 and .89, respectively. Mean state and trait anxiety scores are shown in Table 3-2. A two-way between-subjects ANOVA performed on the state anxiety scores found a main effect of feedback ($F(1,84) = 12.06$, $p < 0.01$). Participants receiving negative feedback reported significantly higher levels of state anxiety (mean = 44.91) than those receiving neutral feedback (mean = 37.50). The number of interviewers did not significantly affect this measure (two-interviewers, mean = 42.93; one interviewer, mean = 39.47). There was no interaction between feedback and number of interviewers. A significant positive correlation was found between state anxiety and response change ($r = .31$, $n = 88$, $p < 0.01$). Analysis of the trait anxiety scores found no main effect of feedback and no interaction. Participants interviewed by two interviewers reported significantly higher levels of trait anxiety (mean = 40.07) than those interviewed by one interviewer (mean = 34.86, $F(1,84) = 9.82$, $p < 0.01$). Reanalysis of the state anxiety scores, holding trait

anxiety as a covariate, showed the same main effect of feedback ($F(1,84) = 9.83$, $p < 0.01$). State anxiety and trait anxiety correlated weakly ($r = .28$, $n = 88$, $p < 0.01$) and no correlation was found between trait anxiety and response change.

**Table 3-2.** Mean (standard deviation) memory recall, anxiety scores, and interview ratings by condition.

| No. of interviewers | Feedback Condition | | | |
|---|---|---|---|---|
| | Neutral | | Negative | |
| | One ($N = 22$) | Two ($N = 22$) | One ($N = 22$) | Two ($N = 22$) |
| Recall (closed items) | 10.23 (3.96) | 10.36 (4.75) | 13.00 (3.49) | 10.68 (3.33) |
| Recall (open-ended items) | 17.05 (8.17) | 14.86 (7.85) | 18.14 (7.49) | 15.55 (6.35) |
| Total recall | 27.27 (10.68) | 25.23 (11.08) | 31.14 (9.42) | 26.23 (8.25) |
| State anxiety | 36.27 (7.50) | 38.73 (9.52) | 42.68 (11.54) | 47.14 (10.98) |
| Trait anxiety | 34.09 (6.18) | 38.00 (7.32) | 35.64 (8.02) | 42.14 (9.30) |
| Rated difficulty | 22.00 (4.25) | 22.91 (6.52) | 25.23 (5.89) | 28.32 (6.05) |

### 3.4.3 Interview ratings

A reliability analysis showed the Interview Ratings form to have high internal consistency (Cronbach's alpha = .84). The item ratings were summed to give an overall measure of interview difficulty. The range of scores was 10 to 50. Table 3-2 shows the mean total interview difficulty scores.

A two-way between-subjects ANOVA showed a main effect of feedback ($F(1,84) = 12.44$, $p < 0.01$) on interview difficulty scores. Participants receiving negative feedback reported higher levels of interview difficulty (mean = 26.77) than those receiving neutral feedback (mean = 22.45). Number of interviewers did not affect perceived interview difficulty or interact with feedback.

### 3.4.4 Response change

Because the number of questions asked varied between participants, response change scores, shown in Table 3-3, were computed as a percentage of the number of questions answered.

A two-way between-subjects ANOVA found a main effect of feedback on response change ($F(1,84) = 18.89$, $p < 0.001$). Participant groups receiving negative feedback made significantly more response changes (mean = 10.81) than those receiving neutral feedback (mean = 3.88). There was no effect of number of interviewers and no interaction between feedback and number of interviewers.

Normality checks on the response change scores revealed a positively skewed distribution. Therefore, as a precaution, the response change data were re-analysed using a Kruskal-Wallis test. In line with the ANOVA, this showed a significant effect of feedback on response change ($\chi^2 = 25.31$, $df = 1$, $p < 0.001$), and no effect of number of interviewers.

**Table 3-3.** Mean (standard deviation) percentage response change by condition.

| | Feedback Condition | | | |
|---|---|---|---|---|
| | Neutral | | Negative | |
| No. of interviewers | One ($N = 22$) | Two ($N = 22$) | One ($N = 22$) | Two ($N = 22$) |
| Response change | 4.43 (7.97) | 3.34 (6.61) | 12.11 (7.92) | 9.51 (7.32) |

Because the experimental sample comprised a relatively wide age range (18 – 72 years), a correlational analysis was conducted on the age of the participants and percentage response change. A Pearson's r calculation showed no relationship between these variables ($r = 0.05$, $n = 88$, n.s.)

### 3.4.5 Response change toward inaccuracy

Across the four experimental groups, a total of 175 responses were changed. Of these, 65 were changed towards accuracy and 70 away from accuracy. The remaining 40 responses were changed neither towards nor away from accuracy. Table 3-4 summarises the direction of response change as it relates to type of feedback.

**Table 3-4.** Total response change (RC) away from and toward accuracy by type of feedback.

| | Feedback Condition | |
|---|---|---|
| | Neutral ($N = 44$) | Negative ($N = 44$) |
| RC away from accuracy | 13 | 57 |
| RC toward accuracy | 24 | 41 |

The response change toward inaccuracy scores were computed as a percentage of the number of questions answered (see Table 3-5). A two-way between-subjects ANOVA showed a significant difference between the neutral and negative feedback groups ($F(1,84) = 21.37$, $p < 0.001$), i.e. negative feedback groups (mean = 5.19) were less accurate than the neutral feedback groups (mean = 1.02). Number of interviewers did not affect this measure and no interactions were found.

Also, the response change toward inaccuracy scores were computed as a percentage of the number of response changes. Again, a significant difference was found between the neutral and negative feedback groups ($F(1,84) = 22.25$, $p < 0.001$). No difference was found between the one and two interviewer groups and there were no interactions. Overall, participants receiving negative feedback changed significantly more responses toward inaccuracy (mean = 40.98) than participants receiving neutral feedback (mean = 9.30).

**Table 3-5.** Mean (standard deviation) response change (RC) toward inaccuracy by condition.

| | Feedback Condition | | | |
|---|---|---|---|---|
| | Neutral | | Negative | |
| No. of interviewers | One ($N = 22$) | Two ($N = 22$) | One ($N = 22$) | Two ($N = 22$) |
| RC as % of no. of questions | 1.10 (3.08) | 0.94 (2.42) | 5.80 (6.15) | 4.57 (4.27) |
| RC as % of total changes | 9.55 (24.39) | 9.06 (24.52) | 40.17 (38.42) | 41.79 (36.01) |

## 3.5 Discussion

Recall scores did not vary as a function of experimental condition, i.e. each group had comparable memories for the event prior to administration of feedback (see Table 3-2). Table 3-3 shows that negative feedback was associated with more response change than was neutral feedback, supporting the first hypothesis and the findings of Tata and Gudjonsson (1990) and Baxter et al. (2006). That some response change was observed in each of the neutral feedback conditions was not unexpected. The repetition of questions may, in itself, have acted as an implicit form of negative feedback (Gudjonsson, 2003; Gudjonsson & Clark, 1986; Register & Kihlstrom, 1988). Also, memory for some details may have been too weak to be consistent on requestioning.

It seems likely that negative feedback increased participants' uncertainty about their previous answers, distracting them and decreasing the reliability of their recall. According to Gudjonsson and Clark (1986), participants who accept the message communicated in the negative feedback are prone to strong subjective and physiological reactions. Important among these are temporary lowering of self-esteem, "cognitive avoidance" and "emotion-focused coping" (Carver, Scheier & Weintraub, 1989; Folkman & Lazarus, 1980; Howard & Hong, 2002) with the likelihood that they question their own judgement and "go along" with the interviewer in order to reduce their level of arousal. In the present study, negative feedback may also have affected participants' expectations, such that they attended less to recall and more to managing the interpersonal situation and meeting the perceived demands of the interviewer (Bain

& Baxter, 2000; Baxter & Boon, 2000; Gudjonsson & Clark, 1986; Gudjonsson & Lister, 1984). These interpretations are supported by the interview ratings (see Table 3-2). Participants receiving negative feedback reported higher overall difficulty with the interview than those receiving neutral feedback.

That negative feedback increased response change in the absence of leading questions supports Gudjonsson's argument that the IP and suggestive questioning aspects of interrogative suggestibility are relatively independent (Gudjonsson, 1984, 1991, 2003). Some evidence suggests that reactions to IP depend primarily on social factors (Bain & Baxter, 2000; Bain et al., 2004; Baxter & Boon, 2000; Baxter et al., 2003; Gudjonsson, 2003; Gudjonsson & Lister, 1984), whereas the acceptance or rejection of misleading information contained in leading questions is mediated more by cognitive mechanisms such as discrepancy detection (Schooler & Loftus, 1986; Tousignant, Hall & Loftus, 1986). Negative feedback also appeared to affect interviewees' reported anxiety (see Table 3-2). Participants receiving negative feedback reported higher levels of state anxiety than those receiving neutral feedback, supporting the Gudjonsson and Clark (1986) prediction. Reported state anxiety and interviewees' tendencies to change responses were positively correlated.

Contrary to the hypothesis, the presence and involvement of a second interviewer had no significant effect on any measure except trait anxiety: participants interviewed by two interviewers reported significantly higher levels of trait anxiety than those interviewed

by a single interviewer (see Table 3-2). That trait rather than state anxiety appeared affected by the number of interviewers was unexpected.

One explanation for this is that two-interviewer group was generally more anxious than the one-interviewer group. Although this might seem unlikely, the absence of a pre-interview measure of trait anxiety means that this possibility cannot be ruled out. It did not seem desirable in planning this study to risk detracting from the attempt to simulate a forensic interview, i.e. the main purpose of the study, by testing participants on the STAI prior to the main procedure. Clearly, however, pre- and post-interview anxiety comparisons would be useful in their own right.

Another possibility might seem to be that the method used in the present study influenced trait anxiety scores. The STAI was administered in this study according to the instructions in the STAI manual (Spielberger, 1983; Spielberger et al., 1970), i.e. administer the state scale first followed by the trait scale. However, participants in this study had been interviewed immediately before completing the scales and may have been too distracted to attend carefully to the distinctions made in the scales. A difficulty with this explanation is that the state scale itself failed to discriminate between the one and two-interviewer groups. Had the two-interviewer groups scored significantly higher on the state scale than the one-interviewer groups, it could be argued that participants did not distinguish between the scales. Because there were no group differences in state anxiety with regard to the number of interviewers, however, this explanation seems

unlikely.

Studies examining the trait scale of the STAI have questioned its construct validity. Bieling, Antony and Swinson (1998) argued that the trait scale of the STAI assesses depression, as well as anxiety, and Kelly (2004) identified a "worry component" in the trait scale, i.e. the scales may not always measure distinct constructs. Future work might test each of these possible explanations of the present findings by systematically varying the use of the STAI scales, comparing the effects of using either the state or the trait scale, or both, before and after interviews.

The second interviewer took an active role in the interrogation, asking every fifth question. That interviewees were nonetheless no more likely to change their answers in this condition may be because even this level of the second interviewer's involvement was not sufficient to impact on them in the predicted ways. Studies of police interviews have shown that one officer tends to have a dominant role. Baldwin (1993) noted that second interviewers often say little, asking only supplementary questions. In this respect, the present study reflected actual interviewing practice. However, interviewer behaviour will often be determined by characteristics of the case such as strength of evidence and offence severity (Moston et al., 1992) and so in some cases the second interviewer will play a more prominent role.

A measure was taken of the tendency of participants to change their answers toward or away from accuracy in response to negative feedback, a measure neglected in previous work on this topic. Where response change occurred, responses were changed both toward and away from accuracy. Negative feedback increased the tendency for answers to be changed toward inaccuracy, and led to more responses away from accuracy than toward it. However, some interviewees' responses were more accurate following negative feedback. This latter effect may be related to the tendency of some interviewees to have, or adopt, a resistant cognitive set when faced with criticism, which motivates them to try harder to be right during requestioning (cf. Baxter et al., 2003; Gudjonsson & Clark, 1986). Alternatively, it may be an artefact of the fairly simple choices with which interviewees were faced: some interviewees may simply have changed an answer - such as "left" to "right" - and so become more "accurate", as a response to IP, without any added recall effort or a belief that their new answers were "better".

A related question left unanswered by this study concerns the degree to which the response change observed was motivated by a desire to appease the interviewer, without acceptance that the initial answer was wrong, as opposed to a genuine private conviction that the second answer provided was, in fact, the correct one. The difficulty here, and in related studies, involves the distinction between compliance and suggestibility. Compliance is a tendency to yield to the direct requests or instructions of another without private acceptance of the other's ideas. Suggestibility also involves yielding to interpersonal influence, with the difference that suggested ideas are also privately

accepted. However, these definitions may often be more theoretically than actually distinct (Woolston et al., 2006). Gudjonsson (1990) found a significant correlation between compliance and suggestibility, and Gudjonsson (2003) argues that the two constructs overlap to some extent. It seems likely that the motivations underlying compliance, such as eagerness to please the questioner and attention to demand characteristics (Orne, 1959), are similar to those thought to elicit suggestibility. Compliance in interviews may often be a compound response which also involves and is triggered by the failures of critical thought characteristic of suggestibility, complicating analyses of the relative contribution of each effect in any particular case (cf. Gudjonsson, 2003; Woolston et al., 2006).

The present findings are based on artificial procedures which limit their relevance to real forensic interviews. Firstly, real-life witnesses and suspects face various pressures which cannot be applied to participants in an experiment. Secondly, real-life interviewees may be asked to recall events after long periods, during which their recall is likely to deteriorate, increasing any tendency to rely on cues present in the interview, as Singh and Gudjonsson (1984) argued. Those authors' findings on response change after a delay, noted above, are not conclusive for reasons which they themselves noted. Indeed, it is possible that the present results are a conservative estimate of average response change following negative feedback, since memories for the event should have been at their strongest immediately after viewing it.

The present findings may have some implications for forensic interviewing practice. Even when overtly leading questions are avoided, some interviewees may alter their answers in response to disapproval communicated by the interviewer, whether this is verbal, non-verbal, or both (see Baxter & Boon, 2000; Bain & Baxter, 2000, Baxter et al., 2006). As such, professional forensic interviewers should be aware of this potential influence in order to minimise the possibility of inadvertently pressurising a witness or suspect. It is not suggested that all forms of IP should always be avoided. In cases where an interviewee appears hostile or evasive, IP which involves confronting interviewees with the evidence against them, in an otherwise non-intimidatory way, may be necessary (Baxter, 2004).

Forensic interviewing which avoids personal criticism is advocated by a number of authors, including Baldwin (1992), Kebbell, Hurren and Mazerolle (2006), McKenzie (2001), Shepherd (1991), Shepherd and Milne (1999), Tousignant (1991), and Williamson (1993, 1994). Their position is supported by studies showing that receiving personal criticism may entrench the attitude of a hostile witness or suspect. For example, Holmberg and Christianson (2002) found that convicted murderers and sexual offenders were more likely to have confessed to their offences if they had been interviewed courteously. In contrast, a dominant, condemning interviewing style, characterised by impatience and aggression, was associated with denials of guilt. With regard to the interviewing of suspects, these authors point to the importance of achieving *admission*, as opposed to confession, and they suggest that admission is more likely to be facilitated

when suspects feel respected and acknowledged and when they are given sufficient time for reflection and comment.

Where negative feedback is used in real forensic interviews, the resulting statements may be challenged by defence counsel on the grounds that inappropriate interview techniques were used. Bull and Milne (2004), reviewing a number of studies, reported that confessions in interviews following initial denials are as likely to occur in non-confrontational interviews as in confrontational ones. Negative feedback may, therefore, be unnecessary at best, may be a distorting influence on recall, and may be counterproductive in some cases.

# Chapter 4: Study 1a

## 4.1 Abstract

In Study 1, the interviewer varied the type of verbal feedback while aiming to hold constant his general demeanour during the interview. It is possible, however, that the communication of negative feedback affected the interviewer's manner in the requestioning phase such that he inadvertently communicated a negative attitude in the post-feedback part of the interview. The present study aimed to test this possibility. Video recordings of one interview containing negative feedback and one containing neutral feedback were shown to two groups of participants. These observers were shown only the portion of the interview following feedback, but not communication of feedback itself, and were asked to rate the interviewer's manner. Between-group behavioural ratings were comparable suggesting that the interviewer maintained a consistent manner across feedback conditions. However, the observers in the present study rated the interviewer more negatively than the interviewees in Study 1. This finding is discussed with relation to the mere exposure effect, social desirability bias and implicit personality theories, and possible implications for forensic interviewing are raised.

## 4.2 Introduction

The key manipulation in Study 1 was the variation of feedback. The interviewer communicated two distinct types of verbal feedback – neutral or negative – to the participant interviewees midway through the questioning procedure, but he otherwise attempted to maintain a consistently formal demeanour. Despite these efforts, a possible difficulty with the procedure is that the act of communicating negative verbal feedback itself may have influenced the interviewer's post-feedback behaviour making it differ from that exhibited in the neutral feedback condition. In other words, the interviewer may have been prone to a type of "halo effect" or "halo error" (Thorndike, 1920) with regard to his own behaviour such that any attitude critical of the interviewee implicit in delivering negative feedback contaminated the interviewer's post-feedback demeanour.

A halo error is a cognitive bias by which people generalise from one prominent characteristic or behaviour to the personality as a whole (Feeley, 2002; Nisbett & Wilson, 1977). Often, an initial understanding that a person has negative (or positive) traits is used to infer other uniformly negative (or positive) characteristics. Describing the same phenomenon, other authors have used the terms constant error (Wells, 1907), logical error (Newcomb, 1931), correlational bias (Kenny & Berman, 1980), and illusory or ubiquitous halo (Cooper, 1981). Such effects have generally been reported in the person perception literature and have related to erroneous inferences made about others. However, to the extent that individuals may also make similar inferences about their

own behaviour in certain situations, this may have introduced bias into the questioning procedure in Study 1.

An illustration of the halo error was reported in a classic study by Asch (1946). He presented a group of judges with a stimulus list of descriptors relating to a target person: intelligent, skilful, industrious, cold, determined, practical and cautious. Another group received the same stimulus list except in place of the word "cold" was the word "warm". Overall impressions of the target person turned out to be influenced significantly by this central trait or dimension. The "warm" group perceived the target as happy, humorous, social and popular, and 91% of this group inferred that the person was generous. The "cold" group chose the descriptors persistent, serious and restrained, and only 8% of this group inferred that the target was generous.

Kelley (1950) repeated Asch's study with a focus on real rather than hypothetical individuals. Students told in advance that a visiting lecturer was "rather cold" evaluated him subsequently in a negative manner. In contrast, other students informed that the lecturer was "rather warm" made a quite different evaluation despite witnessing the same behaviour. Also, while only 32% of the "cold" group interacted with the lecturer during class, 56% of the "warm group" did so.

It seems likely that the communication of negative verbal feedback in Study 1 was

perceived by the interviewer as a prominent or central feature of the experimental procedure and, as such, it had the potential to influence the interviewer's overall manner. It is possible that the experience of communicating negative feedback facilitated a cognitive bias by which the interviewer overestimated the covariance between the central or outstanding behaviour, i.e. negative feedback, and his own subsequent behaviour during requestioning. If such a bias occurred, it may have affected the post-feedback behaviour of the interviewer making it more abrupt and unfriendly than that portrayed in the neutral feedback conditions. He may, despite his best efforts, have remained "in character" and continued to convey a negative attitude towards the interviewee after delivery of feedback. Another mechanism which may have affected the interviewer's post-feedback behaviour is cognitive dissonance reduction (Festinger, 1957). In delivering negative feedback dishonestly, the interviewer may have felt some psychological discomfort and may have attempted to justify this deception and unfair criticism by thinking less favourably of his interviewees in the negative feedback condition and behaving more negatively toward them. Any such unplanned variations in interviewer manner could confound analysis based on variations in verbal feedback of the effects on response change observed in Study 1. (For evidence of the influence of negative interviewer manner on suggestibility, see Bain & Baxter (2000); Bain et al. (2004); Baxter & Boon (2000); Baxter et al. (2006); and Baxter et al. (2003)).

It is also possible that interviewees reacted negatively to the interviewer as a result of the negative feedback and communicated this negativity to the interviewer. If the

interviewer reacted in kind then this too may have increased the psychological distance (Baxter & Boon, 2000; Gudjonsson & Lister, 1984) between the interviewee and interviewer beyond the level intended by the design of the experiment.

These effects, if they occurred in Study 1, may have operated singly or additively. The present study was undertaken to test whether the effects of negative feedback may not be confined to a single verbal message. The importance of this point is that Study 1, lacking leading questions, offers the possibility that the role of verbal negative feedback can be isolated and quantified. It is necessary, however, to ensure that its effects were indeed confined to what was said and were not also influenced by the nonverbal and attitudinal effects postulated above.

To test for such effects, two new groups of participants were shown excerpts from either a negative or neutral feedback interview from Study 1, after the feedback had been delivered. It was assumed that, in the absence of the communication of feedback, these raters would be able to make an unbiased assessment of the interviewer's manner. If no group differences in observer ratings were found, we might assume that the interviewer's manner remained constant across conditions, and this would allow greater confidence in the assumption that results in Study 1 were due more to the intended variations in feedback and less to confounds involving unintended variations in interviewer manner.

The purpose of this study then was to investigate the possibility that interviewee responses in Study 1, particularly as they related to negative feedback, were affected not only by verbal feedback, but by perseveration of a negative attitude on the part of the interviewer. As well as serving as a manipulation check, this study offered the opportunity to examine the impact that negative feedback has on interviewee perceptions. Previous work has examined behavioural and affective responses to negative feedback, while interviewee perceptions of interviewing manner have received less attention. Consideration of these perceptions may be important in understanding aspects of interviewee behaviour such as the likelihood of denials and admissions by suspects (e.g. see Holmberg & Christianson, 2002).

## 4.3 Method

### 4.3.1 Design

The study had a two-factor, between participants design. The independent variables were type of feedback (neutral or negative), and category of rater (observer or interviewee).

### 4.3.2 Participants

Forty seven students (8 males, 39 females) were recruited as interview observers, including Arts students and mature students attending evening programmes. Ages ranged from 17 to 63 years (Mean = 21.43, SD = 10.75). Twenty three and twenty four participants were randomly assigned to the neutral and negative feedback conditions respectively.

### 4.3.3 Materials

#### 4.3.3.1 Videotaped interviews

Video recordings of two interviews, originally conducted for Study 1, were selected for use in this study. The recordings show a 48 year old male ask a set of questions to an 18 year old male participant. In both recordings, the interviewer's manner was intended to be consistently formal; neither overtly friendly, nor too abrupt. In one of the recordings, questioning by the interviewer follows the communication of negative feedback to the interviewee; the other recording is the same except that the questioning follows the communication of neutral feedback. Importantly, only the questioning phase subsequent to feedback was shown to participant observers in this study, i.e. observers were blind to the experimental manipulation of the type of feedback communicated.

#### 4.3.3.2 Interviewer Rater Form

This form used a 5-point Likert scale to assess 18 aspects of the interviewer's manner: nervous, severe, friendly, understanding, assertive, confident, professional, firm, respectful, positive, formal, warm, stern, organised, effective, authoritative, competent, and negative, with a high score being more nervous, etc. The form has been shown in previous studies (Bain & Baxter, 2000; Bain et al., 2004; Baxter et al., 2003) to reliably measure differences in perceptions of interviewer behaviour. The Interviewer Rater Form is shown in Appendix J.

### 4.3.4 Procedure

Participants were asked to watch either the neutral feedback or the negative feedback interview (although they were not shown the feedback itself) and to focus their attention

on the interviewer, rather than the interviewee. They were then asked to complete the Interviewer Rater form to record their perceptions of the interviewer's behaviour.

## 4.4 Results

Ratings by the observers were compared to the ratings made by those interviewees in Study 1 who had been interviewed by a single interviewer. Table 4-1 shows the mean interviewer behaviour ratings.

A two-way between-subjects ANOVA on these ratings found a main effect of category of rater on the following adjectives: severe ($F(1,87) = 14.80$, $p<0.001$); friendly ($F(1,87) = 66.86$, $p<0.001$); understanding ($F(1,87) = 58.65$, $p<0.001$); confident ($F(1,87) = 7.86$, $p<0.01$); professional ($F(1,87) = 14.28$, $p<0.001$); respectful ($F(1,87) = 51.76$, $p<0.001$); positive ($F(1,87) = 43.84$, $p<0.001$); warm ($F(1,87) = 85.26$, $p<0.001$); stern ($F(1,87) = 11.56$, $p=0.001$); organised ($F(1,87) = 14.12$, $p<0.001$); effective ($F(1,87) = 24.70$, $p<0.001$); competent ($F(1,87) = 16.48$, $p<0.001$); and negative ($F(1,87) = 32.54$, $p<0.001$). There were no significant differences between interviewee and observer ratings for the descriptions nervous, assertive, firm, formal, and authoritative. There was no effect of type of feedback and no interaction between feedback and category of rater.

Regardless of type of feedback, interviewees tended to rate the interviewer as more friendly, understanding, confident, professional, respectful, positive, warm, organised,

effective, and competent. Observers rated the interviewer higher on the descriptions severe, stern and negative.

**Table 4-1.** Mean (standard deviation) interviewer behaviour ratings by condition.

| | Feedback Condition | | | |
|---|---|---|---|---|
| | Neutral | | Negative | |
| Category of rater | Interviewee | Observer | Interviewee | Observer |
| | (N = 22) | (N = 23) | (N = 22) | (N = 24) |
| Nervous | 1.05 (0.21) | 1.43 (0.59) | 1.23 (0.69) | 1.25 (0.53) |
| Severe** | 2.00 (1.02) | 3.04 (1.22) | 2.59 (1.18) | 3.25 (0.74) |
| Friendly** | 3.50 (1.22) | 1.78 (0.67) | 2.95 (1.09) | 1.50 (0.59) |
| Understanding** | 3.50 (1.10) | 1.96 (0.64) | 3.05 (0.79) | 1.75 (0.94) |
| Assertive | 3.95 (0.84) | 3.70 (1.36) | 3.91 (0.75) | 3.54 (1.14) |
| Confident* | 4.59 (0.50) | 3.91 (1.35) | 4.36 (0.58) | 4.00 (0.83) |
| Professional** | 4.64 (0.49) | 3.78 (1.38) | 4.50 (0.74) | 3.75 (1.15) |
| Firm | 4.14 (0.83) | 4.17 (1.27) | 3.91 (0.75) | 4.46 (0.59) |
| Respectful** | 4.23 (0.69) | 2.91 (1.00) | 4.18 (0.73) | 2.88 (0.99) |
| Positive** | 3.68 (1.09) | 2.35 (0.83) | 3.41 (1.10) | 2.04 (0.86) |
| Formal | 3.86 (0.99) | 3.78 (1.48) | 4.18 (.080) | 3.67 (1.17) |
| Warm** | 3.05 (1.00) | 1.26 (0.69) | 2.50 (0.96) | 1.21 (0.41) |
| Stern** | 2.59 (1.44) | 3.57 (1.08) | 3.18 (1.14) | 3.88 (0.99) |
| Organised** | 4.59 (0.59) | 4.00 (1.09) | 4.55 (0.67) | 3.75 (1.03) |
| Effective** | 4.36 (0.79) | 3.18 (1.05) | 4.27 (0.70) | 3.50 (1.10) |
| Authoritative | 3.77 (1.07) | 4.00 (1.21) | 4.00 (0.93) | 4.08 (0.97) |
| Competent** | 4.36 (0.90) | 3.70 (1.18) | 4.50 (0.67) | 3.58 (0.88) |
| Negative** | 1.50 (0.74) | 2.87 (0.92) | 1.82 (1.01) | 2.67 (1.01) |

*p<0.01, **p≤0.001

## 4.5 Discussion

No significant differences in interviewer ratings were present between the neutral and negative feedback conditions. This applied to both interviewee and observer ratings. Hence, it would seem that the interviewer was successful in maintaining a consistent demeanour across these two experimental conditions despite the variation in the nature of verbal feedback communicated. This finding also suggests that the global impression formed of the interviewer, both by interviewees and observers, was not influenced significantly by the type of feedback. At least in the context of this type of simulated forensic interview, the majority of participants in Study 1 responded affectively and behaviourally to a critical message contained in the negative verbal feedback, but did not appear to generalise beyond the negative feedback to inferences about the person delivering the message. Given the wide ranging dynamics inherent in real forensic interviews such as the propensity of interviewers to apply differing degrees of IP, further work might assess the generalisability of this finding.

Had the interviewer allowed the communication of negative feedback to influence his manner during requestioning, significant differences in observer ratings would be expected across the two feedback conditions. That each group of observers made comparable ratings suggests that the communication of negative feedback did not bias the interviewer's behaviour subsequent to feedback.

The category of rater was a significant factor in 13 of the 18 interviewer behaviour ratings. Compared with the observers, interviewees rated the interviewer higher on the descriptions friendly, understanding, confident, professional, respectful, positive, warm, organised, effective, and competent, and lower on the descriptions severe, stern and negative. As such, it is clear that the interviewees formed a more positive global impression of the interviewer than did the observers.

An explanation for this finding may involve the relative degree of exposure interviewees and observers had to the interviewer. In contrast to the observers who witnessed only a portion of the interview on a television screen, interviewees experienced the interviewer in person and were exposed to him for the duration of the entire interview. It may be that the interviewees made more positive ratings of the interviewer than the observers due to the "mere exposure effect" (Zajonc, 1968). This has been defined as an increase in positive affect or reduction in negative affect resulting from the repeated or prolonged presentation of previously unfamiliar stimuli (Harmon-Jones & Allen, 2001; Seamon, Brody & Kauff, 1983). A robust phenomenon, the mere exposure effect has been demonstrated with a diverse range of stimuli such as musical sequences (Szpunar, Schellenberg & Plinner, 2004), polygons (Newell & Shanks, 2007), legal arguments (Wilson & Miller, 1968), photographs of faces (Mita, Dermer & Knight, 1977; Zajonc, 1968), and of most pertinence to the present study, real life individuals (Moreland & Beach, 1992). (For a review of studies illustrating the mere exposure effect, see Bornstein (1989)).

Also, it may be that not only the degree of exposure, but the type of exposure influenced the observer ratings. The observers did not observe the interviewer in person and they had no interaction with him. Compared with the interviewees then, there was a lack of physical immediacy and, hence, much greater social distance between the observers and the interviewer. Because they had no direct contact with the interviewer, the observers were possibly less prone than the interviewees to a social desirability response set in their ratings. (Note that the interviewer was present when interviewees were completing the rating form). Hence, the observers may have felt more able to make more negative ratings of the interviewer than the interviewees.

The finding that the interviewees perceived the interviewer differently to the observers may also be understood in terms of the observers' implicit personality theory, i.e. a set of assumptions made regarding the correlations between personality traits (Bruner & Tagiuri, 1954; Cronbach, 1955; Leyens, 1983). It is assumed that people employ their implicit personality theory in making inferences about the characteristics of a target person to the extent that these characteristics are unknown. For example, knowing only that a target person is assertive, people will rely on their implicit personality theory to infer other characteristics of the target such as level of intelligence. If assertiveness and intelligence are closely related in a given implicit personality theory, then holders of the theory will assume, often wrongly, that an assertive target person is also high in intelligence. A number of studies have indeed demonstrated that implicit personality

theories are most likely to operate in situations in which there is low familiarity of rater with the target person (e.g. Koltuv, 1962; Vonk & Heiser, 1991). In the present study, the observers were exposed only to the requestioning phase of the experiment in which the interviewer's manner was businesslike and formal. The interviewees in Study 1 experienced the questioning, but were also greeted informally by the interviewer before the experiment began. Because the observers had less information about the interviewer than the interviewees, they may have made a different set of judgements, perhaps based on an implicit personality theory that associates formality with negative traits.

The present study served primarily as a manipulation check on Study 1 of this thesis and implications of the findings for professional interviewing would seem to be limited. Nevertheless, it may be worth noting that the global impression of an interviewer formed by witnesses and suspects may be significantly influenced by degrees of exposure and familiarity. It is not certain that lack of familiarity with an interviewer can have significant consequences for the conduct of interviews. However, there is a common agreement that the task of increasing familiarity through a process of "rapport" building in the early stages of an interview is important in facilitating communication between the individuals involved (Eyewitness Evidence: A Guide for Law Enforcement, 1999; Minichiello, Aroni, Timewell & Alexander, 1990; Newberry & Stubbs, 1990; Olsen & Wells, 1991; St-Yves, 2006; Tousignant, 1991; Zulawski & Wicklander, 1993) and, in some cases, may have consequences for the outcome of an interview (Baldwin, 1993; Ord & Shaw, 1999). Indeed, Collins, Lincoln and Frank (2002) found that participants

interviewed in a rapport building manner recalled more correct information about a videotaped event than participants interviewed in neutral or abrupt styles. In another study, Walkley (1987) surveyed 100 British police officers and found that they reported a close relationship between the establishment of rapport with suspects and the likelihood that the suspects would later confess. As noted previously, Holmberg and Christianson (2002) found that convicted murderers and sexual offenders were more likely to have confessed to their offences if they had been interviewed courteously.

Certain witnesses are eligible for "special measures" such as the use of a live videolink to help them give evidence in criminal proceedings (Cooke & Davies, 2001; Davies & Noon, 1991). Videolink technology is also used occasionally to enable witnesses to give evidence from abroad. To the extent that some witnesses may react differently to a physically remote or absent questioner than to a physically present one, the present findings may also have implications for such instances. It is acknowledged, however, that the observers in this study were not questioned by the interviewer and were unable to interact with him in any way. Nevertheless, future experimental work might assess the extent to which lack of familiarity with an interviewer and/or increased social distance between interviewee and interviewer has the potential to affect not only person perception, but also interviewee responding. Following the present study, one suggestion is to compare interviewer behaviour ratings and patterns of responding by experimental participants interviewed face-to-face with those of participants interviewed via videolink. If, as might be suggested by the present findings, participants rate a remote

interviewer more negatively than a physically present one, this might indicate a need to alter interviewing practice where videolink technology is used.

# Chapter 5: Study 2

## 5.1 Abstract

This study examined the effect of self-esteem and the presence of an interviewee ally on interviewee responses in a simulated forensic interview. It was predicted that individuals high in self-esteem and those supported throughout questioning by an ally would be less susceptible to interrogative pressure and would change significantly fewer responses. A further aim was to assess the direction of response change following negative feedback. The experiment had a two-factor between participants design. Low and high self-esteem participants (N = 88) watched a videotape of a simulated crime, answered questions about it, and were requestioned after receiving negative feedback. Half of the participants were accompanied by a supportive "ally" who intervened immediately following feedback. Level of self-esteem was found to be associated with self-reported anxiety following questioning and rated difficulty of the interview, but not with response change. The presence of the ally had no effect on the dependent measures. Negative feedback increased the tendency for answers to be changed towards inaccuracy.

## 5.2 Introduction

In Study 1 of this thesis, response change was shown to be significantly affected by a specific situational factor in the interrogative context; the type of feedback communicated to participants during questioning. Gudjonsson and Clark's (1986) model of interrogative suggestibility predicts that personality variables in interviewees are also likely to mediate suggestibility. One such variable that has received attention in the research literature is self-esteem. This hypothetical construct has been defined as, "a favourable or unfavourable attitude toward oneself" (Rosenberg, 1965, p 18) and "the overall affective evaluation of one's own worth, value, or importance" (Blascovich & Tomaka, 1991, p 115). Within the psychological literature, it is generally assumed that a person's self-esteem is a reflection of attitudes toward, or evaluations of, one's self.

Gudjonsson and Clark (1986) argue that the self-esteem of subjects in an interrogative context may be temporarily affected by signals communicated by the interrogator. Specifically, they suggest that acceptance (as opposed to rejection) of negative feedback during questioning often leads to a "debilitating" cognitive evaluation and lowering of self-esteem. (Acceptance of negative feedback appears to be associated with the tendency of some individuals to rely on external social cues as a means of regulating their behaviour, a tendency similar to the concept of self-monitoring (Snyder, 1974, 1987). For evidence of the relationship between self-monitoring and interrogative suggestibility, see Bain et al. (2007)). From Gudjonsson and Clark's model, it follows that participants already low in self-esteem may be more prone than high self-esteem

individuals to the suggestive influence of negative feedback.

Research using the GSS has shown how self-esteem appears to be related to suggestibility. Gudjonsson and Singh (1984b) assessed delinquent adolescent boys on the GSS 1 and had their teachers rate the boys' self-esteem on the Coopersmith Behaviour Form (Coopersmith, 1967). A negative relationship was found between the self-esteem ratings and measures of Yield 1 and Shift. Low self-esteem ratings were associated with higher Yield 1 and Shift scores. Singh and Gudjonsson (1984) tested a sample of nurses in a psychiatric hospital on the GSS 1. They used Osgood's Semantic Differential Scales (Osgood, Suci & Tannebaum, 1957) to assess the nurses' perceptions of themselves and the experimenter. Singh and Gudjonsson found that Yield 1 and Shift correlated negatively with this measure of self-esteem. Gudjonsson and Lister (1984) also used the Semantic Differential Scales to assess the perceived distance between a sample of participants and the experimenter. The semantic differential method is intended to measure attitudes towards concepts and objects. Respondents rate target concepts on a number of 7-point scales, each consisting of a pair of bipolar adjectives (e.g. strong-weak, incompetent-competent, soft-hard). In the Gudjonsson and Lister study, participants rated the concepts "Myself during the experiment" and "The Experimenter". Where participants perceived a large semantic distance with regard to competence and control of the interrogative context, this was related to higher suggestibility scores on the GSS 1.

Baxter et al. (2003) tested university students on the Battle Self-Esteem Inventory (Battle, 1981) and the GSS 1. Students with low self-esteem scored significantly higher on all of the scale's suggestibility measures: Yield 1, Yield 2, Shift and Total Suggestibility. Baxter et al. also found an interaction between self-esteem and interviewer behaviour on Yield 2 and Shift. Low self-esteem participants' scores on these two measures increased from Friendly to Abrupt interviewer conditions, while high self-esteem participants' scores decreased from Friendly to Abrupt. Baxter et al. argue that this reflects the tendency of low self-esteem individuals to be sensitive to IP and their inclination to act to reduce psychological distance (Gudjonsson & Lister, 1984) between themselves and an unfriendly interviewer. They also note that, in contrast, high self-esteem individuals are likely to have increased resistance to IP (Gudjonsson & Lister, 1984; Singh & Gudjonsson, 1984) and that increases in psychological distance between an interviewer and interviewees with high self-esteem will heighten their suspicion and attentiveness and lead to reduced suggestibility (Gudjonsson, 2003; Gudjonsson & Clark, 1986). Testing students on the same instruments, Bain et al. (2004) also assessed the effect of self-esteem on suggestibility. They found that low self-esteem participants scored higher on the Shift measure than high self-esteem participants.

Together, the five studies cited above suggest that low self-esteem is associated with heightened suggestibility and response change in particular. One possible reason for these findings is that low self-esteem individuals may have poorer recall of the GSS narrative than high self-esteem individuals, and uncertainty concerning the content of

the narrative accounts for the tendency to more frequently change previous responses. This possibility has limited support in the literature, however. In the Bain et al. (2004) study, the low self-esteem group obtained significantly lower memory recall scores than the high self-esteem group. In contrast, Baxter et al. (2003) found no difference in memory recall between low and high self-esteem groups. Another reason for the relationship between self-esteem and response change may relate to the IP applied during the GSS procedure of which Shift is intended to measure (Gudjonsson, 1984). IP increases the "psychological distance" between the interviewer and interviewee (Baxter & Boon, 2000; Gudjonsson & Lister, 1984) and may create a social demand characteristic to change previous responses. Under these circumstances, low self-esteem individuals may be more likely to attend to interviewer demand at the expense of accurate recall in the attempt to reduce the apparent disapproval of the interviewer to which they are particularly sensitive (Gudjonsson & Clark, 1986) and/or they may experience higher anxiety, assuming perhaps that the interviewer's behaviour reflects badly on them, such that accurate recall is disrupted.

As noted, studies examining the relationship between self-esteem and interrogative suggestibility have generally yielded positive results. An exception to this trend is a study by Smith and Gudjonsson (1995) which found no relationship between self-esteem and any of the GSS 2 outcome measures. These authors assessed forensic psychiatric inpatients using the Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1965). Noting the high face validity of the RSE, Smith and Gudjonsson questioned whether the scale

yielded a reliable measure of self-esteem in their study. While not a problem with "normal" populations, they suggest that social desirability faking on the RSE may be more likely with forensic psychiatric inpatients whose responses may represent their ideal rather than actual self, i.e. concerned with "self presentation" rather than honest reporting, forensic psychiatric inpatients may choose response categories which indicate a higher level of self-esteem than they actually possess.

The RSE is the most widely used measure of global self-esteem and the standard against which the convergent validity of other measures is established (Blascovich & Tomaka, 1991; Emler, 2001). It is perhaps surprising then that, of the studies examining self-esteem and suggestibility, only one has used the RSE. The present study aimed to further examine the relationship between self-esteem and response change using the RSE with a non-forensic sample. In line with Gudjonsson and Singh (1984b), Singh and Gudjonsson (1984), Gudjonsson and Lister (1984), Baxter et al. (2003), and Bain et al. (2004), it was predicted that a low self-esteem group, compared with a high self-esteem group, would score higher on response change following negative feedback. Where this study differs from its predecessors is in examining effects of interviewee self-esteem in the absence of overtly leading questions.

The second aim of this study was to examine the effects of the presence and intervention of an interviewee ally. Adult witnesses and suspects considered to be psychologically vulnerable and juvenile suspects undergoing forensic interviewing are entitled or

required in many jurisdictions to have an appropriate adult present during questioning, e.g. a parent, someone with parental responsibility, a social worker, or another adult not employed by the police. The role of the appropriate adult is to ensure that the interviewee's interests are protected during questioning and to facilitate communication between the interviewer and interviewee (PACE, Home Office, 1985a, 1985b; Williams, 2000). This role is distinct from that of the legal representative whose purpose is to safeguard the legal rights of their client. In practice, however, legal representatives may play a more generally supportive role where there is no appropriate adult present (Gudjonsson, 2003; Medford et al., 2003). Medford et al. reported that although appropriate adults tend to speak and intervene little in interviews, their presence appeared to be associated with less IP than in interviews with no appropriate adult present. It may be that, in the presence of an interviewee ally (an appropriate adult or legal representative), forensic interviewers are less likely to apply high levels of IP in the knowledge that the ally may intervene. Alternatively, in cases where IP *is* applied, an intervention by an interviewee ally may facilitate a more critical cognitive appraisal by the interviewee leading to a more resistant coping strategy (Gudjonsson & Clark, 1986). This study tested the latter possibility.

A further issue is whether participants change their answers toward or away from accuracy during requestioning, following feedback. In Study 1, neutral feedback and negative feedback was found to elicit response change in both directions. Following neutral feedback, more responses were changed toward accuracy than away from it: this

pattern was reversed in the negative feedback condition. Compared with neutral feedback, negative feedback appeared to increase the tendency for answers to be changed toward inaccuracy. A final aim of this study was to assess the direction of response change following negative feedback, thereby testing the reliability of the findings of Study 1.

The present study had three hypotheses, therefore. Firstly, participants with low self-esteem would change more responses, rate the interview as more difficult, and report higher levels of state anxiety than those with high self-esteem. Secondly, the presence and intervention of an interviewee ally would reduce IP, such that interviewees with an ally would be less likely to change their responses following negative feedback than interviewees with no ally. Thirdly, negative feedback, in comparison with neutral feedback, would increase the tendency for answers to be changed toward inaccuracy during requestioning.

## 5.3 Method

### 5.3.1 Design

The study had a two-factor, between participants design. The independent variables were self-esteem (low or high), and presence or absence of an interviewee ally (ally or no ally).

### 5.3.2 Participants

88 undergraduate students (29 males, 59 females) took part in return for course credits.

Ages ranged from 17 to 43 years (Mean = 19.14, SD = 3.87). Forty four participants were each randomly assigned to the Ally and No Ally conditions. Participants were further divided approximately equally to the High and Low Self-Esteem groups according to their score on the RSE (details below).

The interviewer was the same 48 year old male who conducted the interviewing in Study 1; the interviewee ally was a 35 year old male. Both dressed formally for testing.

### 5.3.3 Materials

#### 5.3.3.1 Videotaped event

The stimulus was the same 77 second videotaped clip used in Study 1 of this thesis. A description of the event is provided in Section 3.3.3.1.

#### 5.3.3.2 Rosenberg Self-Esteem Scale (RSE)

This is a 10-item self-rating scale designed to assess individual differences in self-esteem. Each item is a general statement about self, e.g. "On the whole, I am satisfied with myself" and "I certainly feel useless at times". Items are answered on a 4-point scale ranging from strongly agree to strongly disagree. Scores range from 10 to 40 with high scores indicating high self-esteem. The scale has shown good reliability and validity (Blascovich & Tomaka, 1991; Greenberger, Chen, Dmitrieva & Farruggia, 2003; Rosenberg, 1965). It is shown in Appendix K.

#### 5.3.3.3 Interview Rating Form

A description of this form is provided in Section 3.3.3.3. Previous use of this form in

Study 1 showed it to discriminate between different interviewee conditions and to possess high internal consistency.

#### 5.3.3.4 State and Trait Anxiety Inventory (STAI)

A description of this instrument is provided in Section 3.3.3.2.

### 5.3.4 Procedure

A random sample of 322 students completed the RSE. Of these, 308 were returned completed with an e-mail address or telephone number and confirmation that the respondent was willing to participate in the full study. The mean, standard deviation, range and percentile scores for this initial sample are shown in Appendices L and M and these serve as normative data. The 44 highest and lowest scorers on the scale were invited for further participation. The mean self-esteem scores for the High and Low groups were 35.44 (SD = 2.15, range = 33-40) and 23.21 (SD = 2.46, range = 16-27) respectively. Comparison with the undergraduate norms shows that the High group scores were equal to and above the 75$^{th}$ percentile, and the Low Group scores were equal to and below the 25$^{th}$ percentile. The difference between means was significant, $t(86) = -24.86$, $p < 0.001$. Coefficient alpha for the scale was .93.

Participants sat at a desk, facing the interviewer across from them. In the Ally condition, the ally sat next to the interviewee facing the interviewer and explained that his role was "to ensure that the interviewer followed the correct procedure and to check that the questions asked were reasonable."

After initial questioning, all participants received negative feedback. In the Ally condition, immediately following the communication of feedback, the ally stated, "Just answer as well as you can based on what you recall from the video clip. It's what you remember that counts."

Following requestioning, participants completed the State Anxiety Inventory, the Interview Rating Form, and the Trait Anxiety Inventory. In all other respects, the procedure used was identical to that used in Study 1.

#### 5.3.4.1 Scoring

Scoring of memory recall, response change and number of responses changed toward inaccuracy were calculated according to the methods detailed in Study 1 (see Section 3.3.4.1).

## 5.4 Results

### 5.4.1 Memory recall

Table 5-1 shows mean recall scores for both closed and general/open-ended items. Separate analyses of these measures, using two-way between-subjects ANOVAs, showed no significant differences or interactions between groups.

### 5.4.2 Anxiety

Mean state and trait anxiety scores are shown in Table 5-1. A two-way between-subjects ANOVA performed on the state anxiety scores found a main effect of self-esteem ($F(1,84) = 15.65$, $p < 0.001$). Low self-esteem participants reported significantly higher levels of state anxiety (mean = 55.65) than high self-esteem participants (mean = 46.22). The presence or absence of an ally did not significantly affect this measure (ally - mean = 50.43: no ally - mean = 51.23). There was no interaction between self-esteem and presence of an ally. In analysis of the trait anxiety scores, Levene's test for homogeneity of variance was significant so equal variances of the samples could not be assumed. A Kruskal-Wallis test showed a significant difference between the low and high self-esteem groups ($\chi^2 = 52.83$, df = 1, $p < 0.001$), i.e. low self-esteem participants reported significantly higher levels of trait anxiety (mean = 52.30) than high self-esteem participants (mean = 33.09). No significant difference in trait anxiety was found between the ally (mean = 42.02) and no ally (mean = 42.93) groups. State and trait anxiety scores were positively correlated ($r = 0.51$, $n = 88$, $p < 0.001$). Participant self-esteem scores correlated negatively with both state anxiety ($r = -0.42$, $n = 88$, $p < 0.001$) and trait anxiety ($r = -0.85$, $n = 88$, $p < 0.001$).

**Table 5-1.** Mean (standard deviation) memory recall, anxiety scores, and interview ratings by condition.

| | Condition | | | |
|---|---|---|---|---|
| | No Ally | | Ally | |
| Self esteem | Low | High | Low | High |
| | (N = 21) | (N = 23) | (N = 22) | (N = 22) |
| Recall (closed items) | 10.48 (4.68) | 9.26 (4.56) | 8.27 (3.44) | 8.73 (4.79) |
| Recall (open-ended items) | 17.19 (9.02) | 16.39 (8.74) | 13.50 (4.59) | 14.64 (9.33) |
| Total recall | 27.67 (12.61) | 25.65 (12.23) | 21.77 (7.00) | 23.36 (13.02) |
| State anxiety | 56.29 (12.28) | 46.61 (11.49) | 55.05 (9.63) | 45.82 (11.29) |
| Trait anxiety | 52.10 (10.37) | 34.57 (5.68) | 52.50 (9.40) | 31.55 (5.65) |
| Rated difficulty | 29.68 (4.27) | 27.77 (6.47) | 30.95 (5.46) | 27.55 (6.63) |

### 5.4.3 Interview ratings

A reliability analysis showed the Interview Ratings form to have high internal consistency (Cronbach's alpha = .78). The item ratings were summed to give an overall measure of interview difficulty. The range of scores was 10 to 50 with a higher score indicating more perceived difficulty. Table 5-1 shows the mean total interview difficulty scores. A two-way between-subjects ANOVA showed a main effect of self-esteem ($F(1,84) = 4.65$, $p<0.05$) on the interview difficulty scores. Participants in the low self-esteem group reported higher levels of interview difficulty (mean = 30.32) than those in the high self-esteem group (mean = 27.66). Presence or absence of an ally did not affect perceived interview difficulty or interact with self-esteem.

### 5.4.4 Response change

Because the number of questions asked varied between participants, response change scores, shown in Table 5-2, were computed as a percentage of the number of questions answered. Analysis of percentage response change, using a two-way between-subjects ANOVA, showed no significant differences or interactions between groups. No correlation was found between self-esteem and percentage response change.

**Table 5-2**. Mean (standard deviation) percentage response change by condition.

| | Condition | | | |
|---|---|---|---|---|
| | No Ally | | Ally | |
| Self esteem | Low (N = 21) | High (N = 23) | Low (N = 22) | High (N = 22) |
| Response change | 11.71 (8.23) | 10.83 (12.62) | 8.05 (9.17) | 12.71 (10.72) |

### 5.4.5 Response change toward inaccuracy

Across the entire sample, a total of 207 responses were changed. Of these, 88 were changed towards accuracy and 83 away from accuracy. The remaining 36 responses were changed neither towards, nor away from accuracy. Examples of changes are given in Section 3.3.4.1 and also in Appendices E and G. Response change towards inaccuracy scores were computed as a percentage of the number of questions asked. A two-way between-subjects ANOVA found no significant differences or interactions between groups. Also, response change towards inaccuracy was calculated as a percentage of the number of response changes. Again, no significant differences or interactions were

found between groups. Descriptive statistics for theses measures are shown in Table 5-3.

**Table 5-3.** Mean (standard deviation) response change (RC) toward inaccuracy by condition.

| | Condition | | | |
|---|---|---|---|---|
| | No Ally | | Ally | |
| Self esteem | Low (N = 21) | High (N = 23) | Low (N = 22) | High (N = 22) |
| RC as % of no. of questions | 4.55 (3.54) | 4.20 (7.59) | 3.65 (4.40) | 4.45 (5.35) |
| RC as % of total changes | 41.44 (36.11) | 26.60 (36.47) | 32.88 (39.36) | 26.46 (31.17) |

To allow comparison of the means with those of a neutral feedback group, data from Study 1 of this thesis are included. That study was highly similar to the present study employing the same questioning procedure, the same interviewer in the same room, and the same video clip. The main differences were that the sample in that study was drawn from a wider age range of participants - 18 to 72 years (Mean = 31.49, SD = 17.75). Comparisons are shown in Table 5-4. (Levene's test for homogeneity of variance was significant so a Kruskal-Wallis test was used). Participants receiving negative feedback in the present study changed significantly more responses towards inaccuracy than did participants receiving neutral feedback in Study 1.

**Table 5-4.** Mean response change (RC) towards inaccuracy by feedback condition.

| | Study 2 * (N = 88) | | Study 1 ** (N = 22) | |
|---|---|---|---|---|
| | M | SD | M | SD |
| RC as % of no. of questions | 4.21[a] | 5.40 | 1.10[a] | 3.08 |
| RC as % of no. of total changes | 31.68[b] | 35.80 | 9.55[b] | 24.39 |

Participants received: negative feedback* or neutral feedback**.

Means sharing a suffix are significantly different: a. ($\chi^2 = 25.31$, df = 1, $p < 0.01$), b. ($\chi^2 = 8.65$, df = 1, $p < 0.01$).

## 5.5 Discussion

The main objectives of this study were to examine the effects of self-esteem and the intervention of an interviewee ally on response change, anxiety and ratings of interview difficulty in a simulated forensic interview. It was hypothesised that dependent measures (with the exception of memory recall) would be highest for low self-esteem participants and participants not accompanied throughout the interview by an ally.

While measures of self-esteem have been shown to correlate with constructs such as depression (Harter, 1993; Shaver & Brennan, 1990), locus of control (Biondo & MacDonald, 1971; Brockner, 1979), self-efficacy (Lane, Lane & Kyprianou, 2004), and neuroticism (Judge, Erez, Bono & Thoresen, 2002), the relationship between global self-esteem and memory recall for events is less clear. (For recent research on autobiographical memory bias associated with self-esteem, see Christensen, Wood and Barrett (2003), and Tafarodi, Marshall and Milne (2003)). As anticipated, the experimental groups in this study scored similarly on recall, i.e. level of self-esteem did not influence participants' ability to recall details of the video taped event (see Table 5-1). This finding is broadly consistent with that of Baxter et al. (2003) who found no differences in memory recall for the GSS 1 spoken narrative between low and high self-esteem groups. However, differences between the type of recall task used on the GSS (free recall) and that used in this study (cued recall) must be borne in mind.

Recall of the videotaped event was calculated on the basis of answers given prior to the intervention of the interviewee ally. It is possible, however, that the mere presence of the ally of itself could have impacted upon recall performance. Results show that the presence of the interviewee ally did not affect recall scores (see Table 5-1). A similar result was obtained in Study 1. The presence of a second interviewer did not influence recall of the videotaped event. Together, these findings suggest that the mere presence of a third person in the interview room may be neither socially facilitating nor inhibiting to the task of recall.

According to the Gudjonsson and Clark (1986) theoretical model of interrogative suggestibility, negative feedback communicated during questioning is subject to a cognitive appraisal by interviewees. It is argued that, following this appraisal, a proportion of individuals will reject the feedback and it will have no significant effect upon subsequent responses. Applied to the present procedure and to that of the GSS, the prediction is that, following feedback, some individuals will exhibit a resistant behavioural response and will not change their initial answers during requestioning. Another possibility advanced by the Gudjonsson and Clark model is that some individuals may accept the negative feedback, but nevertheless not alter their behavioural response pattern, believing that to do so would not improve their performance. Consistent with these predictions and with the findings of Study 1 of this thesis, a proportion of the sample in the present study (20.45%) did not change any responses following feedback. In Study 1, 4 of the 44 participants receiving negative

feedback (9.09%) did not change any responses.

Consistent with Smith and Gudjonsson (1995), level of self-esteem as measured by the RSE, was not associated with degree of response change in this study (see Table 5-2). Smith and Gudjonsson's sample comprised forensic psychiatric inpatients and, given the propensity of such a group for social desirability faking, they suggest that the high face validity of the RSE scale may have been problematic with such a specialised population. The present study assessed the responses of a "normal" adult sample and so this difficulty was not anticipated. Importantly, the failure to find a relationship between self-esteem and response change is not consistent with the findings of studies using the GSSs which have shown low self-esteem to be associated with an increase in Shift (Bain et al. 2004; Baxter et al., 2003; Gudjonsson & Lister, 1984; Gudjonsson & Singh, 1984b; Singh & Gudjonsson, 1984).

Using the current methodology, group differences based on level of self-esteem may not be sufficient to elicit differences in response change. As noted previously, Baxter et al. (2003) found low self-esteem to be associated with increased Shift scores on the GSS 1. Importantly, they also found an interaction between self-esteem and interviewer behaviour: low self-esteem participants' scores increased from Friendly to Abrupt conditions, whereas high self-esteem participants' scores decreased between Friendly to Abrupt (cf. Bain et al., 2004). The apparent lack of effect of self-esteem on response change in the present study may be obscuring a difference in performance which would

otherwise be observed in conjunction with differing interviewing styles. This possibility, in part, informs the rationale for Study 3 of this thesis.

The studies under focus here have each used different instruments and methods to measure self-esteem. For instance, while the present study, Smith and Gudjonsson (1995), Baxter et al. (2003) and Bain et al. (2004) used self-ratings (the RSE and the Battle Self-Esteem Inventory), Gudjonsson and Singh (1984b) employed an "other-report" measure of self-esteem (the Coopersmith Behaviour Form) based on ratings made by observers. Singh and Gudjonsson (1984) and Gudjonsson and Lister (1984) did not use a self-esteem scale, but assessed attitudes towards self with the semantic differential technique. Given this diversity of procedure, it is not certain that these varying methods are measuring the same construct (Demo, 1985; Wells & Marwell, 1976). This may limit the extent to which studies can be compared and may be a reason for the present discrepant finding.

Also, it should be noted that previous work included overtly leading questions which may have applied additional IP beyond that of negative feedback alone. Higher levels of IP may be a factor in producing a self-esteem effect on response change.

A further complication arises in the scoring of the RSE scale itself. The present study used the typical scoring scheme of a four-point response format (strongly agree, agree,

disagree, strongly disagree) giving a possible scale range of 10-40 (Blascovich & Tomaka, 1991). The broader range of self-esteem scores obtained by Smith and Gudjonsson (1995) (their scores ranged from 16-50) suggest that these authors used a 5- or 7- point Likert scale. Such differences in scoring of the RSE can make comparison of studies using this scale difficult.

In this and similar studies, participants scoring at the extreme ends of the measurement scale were allocated to low and high self-esteem groups and, as such, were deemed to be representative of those in the wider population who are "low" or "high" on this construct. The efficacy of this procedure requires a large sample of initial respondents and a normal or platykurtic distribution of scores to ensure sufficient separation between low and high scorers. A normality check conducted on the initial sample of respondents in this study showed the distribution to be normal based on kurtosis. Nonetheless, although the groups in this study were distinct in a numerical sense, it may be that, in terms of self-esteem, they were not sufficiently different to affect behavioural responding during requestioning (see Section 5.3.4). A related issue is, in undergraduate samples, scores tend to be skewed towards the high end of measurement scales. Referring to the RSE, Blascovich and Tomaka (1991, p. 123) note that, "even tripartite splits of the distribution produce "low" self-esteem groups that have relatively high self-esteem in an absolute sense."

With reference to the standard GSS procedure, Gudjonsson (1997) notes that when Shift

occurs, it is usually in the direction of increased suggestibility. That is, after negative feedback and repeated questioning, subjects tend to yield more to the leading questions than they did before. It seems then that Shift occurs often in response to both the cognitive and the social aspects of IP. Upon requestioning, interviewees may accept the message of disapproval communicated in the negative feedback and, adopting a suggestible cognitive set, may change their initial answers. At the same time, feedback may compound difficulties in the cognitive processing of leading questions and, faced with a seemingly credible "alternative" answer, interviewees may opt to shift (Gudjonsson & Clark, 1986). It is assumed that leading questions create uncertainty in interviewees, or add to existing uncertainty, as to the correct answer to questions asked. As noted previously, cognitive processes such as discrepancy detection may be important here in interfering with the ability to discriminate accurately between what was observed and what has been suggested (Schooler & Loftus, 1986; Tousignant et al., 1986). Gudjonsson and Clark (1986) suggest that the social pressures of repeated questioning and negative feedback heighten the expectations of interviewees that they should be able to provide the answer to questions rather than saying "I don't know". These pressures also absorb attention as interviewees attend to their own feelings and attempt to understand and manage what may be a tense interpersonal situation at the expense of devoting attention to accurate recall.

The measure of response change employed in the present studies primarily assesses reactions to the social aspects of IP. In other words, the absence of overtly leading

questions diminishes the cognitive aspects of susceptibility to IP. It may be that the effects of low self-esteem on response accuracy are more likely to be evident when application of IP includes specific cues to the "correct" answer. Leading questions are worded in such a way as to permit interviewees to infer some knowledge about a past event, and/or to infer the interviewer's knowledge, beliefs and motivation. Neutral questions do not cue such inferences (Fisher & Geiselman, 1992). IP applied solely by negative feedback, because it provides no specific cues as to which answers were "wrong", may generally disrupt the accuracy of participant recall, irrespective of their self-esteem. If so, then this may be an additional factor contributing to the absence of a self-esteem effect (besides the clustering around the mean of the self-esteem scores of the participants in this study) on response change in this study.

That level of self-esteem had no significant effect on response change in this study means that the first hypothesis was not confirmed. Level of self-esteem did, however, have a significant impact on the other dependent measures. The relationship between self-esteem and anxiety has been well documented (e.g. Greenberg et al., 1992; Heatherton & Ambady, 1993). Hence, it was anticipated that low self-esteem participants would report higher levels of both trait and state anxiety than high self-esteem participants. This prediction was confirmed. In addition, the low self-esteem group rated the interview as significantly more difficult than the high self-esteem group (see Table 5-1). To the extent that research findings indicate that low self-esteem individuals are more likely to perform less effectively under stress and failure

(DeLongis, Folkman & Lazarus, 1988; Schalon, 1968; Shrauger & Rosenberg, 1970), this finding was expected.

The presence and intervention of the interviewee ally did not impact on the rate of response change or on the other dependent measures following negative feedback. As such, the second hypothesis was not supported. Previous studies using the GSSs have shown that warnings to be vigilant can decrease suggestibility scores (Bain et al., 2004; Baxter et al., 2006; Boon & Baxter, 2000). The warnings in these studies were communicated by the interviewer. Nevertheless, it was anticipated that the intervention by the interviewee ally in the present study would be socially supportive and similarly, or more effectively, facilitate the resistance of IP and lead to a decrease in response change. Social support is recognised as helpful in dealing with stressful events (Cohen & Wills, 1985; Suls & Wallston, 2003) and has been shown to assist nonconformity in the face of social pressure (Asch, 1955, 1956). A question not answered in this study, however, is the extent to which participants perceived the ally's intervention and presence during questioning to be socially supportive.

It may be that the wording of the ally's intervention was not clear enough to provoke a resistant behavioural response in the participants, or that the ally's sole intervention was not sufficient to reinforce participants' confidence in their recall. Possibly, some participants perceived the intervention actually to support the negative feedback. Further work might attempt to establish that the involvement of an ally is perceived as socially

supportive and might examine the effectiveness of a more assertive ally in influencing IP in interviews.

Consistent with the findings of Study 1 of this thesis, responses following negative feedback were changed both toward and away from accuracy. In Study 1, more answers were changed away from accuracy than towards it after negative feedback. This trend was slightly reversed in the present study, but what emerges from both studies is that, using the present methods, when negative feedback changes interviewees' answers, the change is towards inaccuracy at least half of the time. This suggests that negative feedback is a significant influence for distortion in interviewees' recall, and unreliable as a means to improve recall.

At the same time, it must be acknowledged that negative feedback has the potential to seemingly *improve* the accuracy of some responses during requestioning. In the present study, 42% of all answers were changed towards accuracy following negative feedback. The reasons for this are not obvious, but it is possible that participants whose cognitive set was resistant processed the feedback as a challenge to improve their recall performance during requestioning and that this resulted in increased accuracy. Another possibility is that accuracy increased not as a consequence of conscious efforts to improve, but rather as a result of the simple choices involved in responding and/or perhaps a failure to adequately recall details of the videotaped event.

Participants in this study were subject to negative feedback only. To assess the relative impact of negative feedback on response change direction, the scores for response change towards inaccuracy were compared with related scores in the neutral feedback condition of Study 1 (see Table 5-4). Compared with neutral feedback, negative feedback increased the tendency for answers to be changed toward inaccuracy. The third hypothesis appears to be supported. However, although the present study and Study 1 were similar in procedure, sampling differences between the studies mean that the findings relating to this comparison should be regarded with some caution.

# Chapter 6: Study 3

## 6.1 Abstract

The effects of interviewer behaviour and interviewee self-esteem on response change were investigated in a simulated forensic interview. In line with Gudjonsson and Clark's (1986) model of interrogative suggestibility, it was hypothesised that increased rates of response change would be associated with an unsupportive, disapproving interviewer manner and low levels of self-esteem. Following presentation of a video taped event, low and high self-esteem participants (N = 83) were interviewed by a male interviewer portraying either a Friendly or Abrupt manner. Participants in the Abrupt conditions made significantly more response changes in response to negative feedback than those in the Friendly conditions. Contrary to previous studies, level of self-esteem did not influence this measure nor did it interact with interviewer behaviour. However, low self-esteem participants rated the interview as significantly more difficult than did high self-esteem participants. The results support some previous work and indicate that negative feedback affects response change even when questions are not overtly leading.

## 6.2 Introduction

As previously noted in this thesis, experimental work using the Gudjonsson Suggestibility Scales has shown that variations in interviewer behaviour can affect tendencies to yield to leading questions and to change responses subsequent to interpersonal pressure. Baxter and Boon (2000) varied the negative feedback component of the GSS 2 such that feedback was delivered in a "friendly", "firm" and "stern" manner. Yield 2 and Shift scores increased as the interviewer manner changed from "friendly" to "stern". Baxter and Boon argued that a "stern" as opposed to a "friendly" demeanour increased the "psychological distance" between the experimenter and participants, giving the experimenter control over the interrogative context and creating interpersonal pressure (Gudjonsson & Lister, 1984). In a similar study, Bain and Baxter (2000) found that an experimenter adopting an "abrupt" style throughout the questioning phase of the GSS procedure elicited significantly higher Shift and Total Suggestibility scores on the GSS 1 than the experimenter portraying a "friendly" style.

Bain et al. (2004) also tested participants on the GSS 1 with a "friendly" and "abrupt" interviewer behaviour manipulation. They reported significantly higher Yield 1 and Total Suggestibility scores in the "abrupt" condition. Baxter et al. (2003) found an interaction between interviewer behaviour and the self-esteem of participants on Yield 2 and Shift. Low self-esteem participants scored lower on these measures in the "friendly" condition and higher in the "abrupt" condition. The pattern was reversed for high self-esteem participants.

The authors of the studies cited above point to the clinical and forensic implications of their findings. They argue that variations in interviewer demeanour whilst administering the GSS can bias test results by producing scores significantly higher or lower than the scales` population norms. A friendly demeanour may artificially reduce scores such that people who are particularly susceptible to suggestions are not identified (a Type II error). In contrast, a stern or abrupt manner may elevate scores leading to the false identification of vulnerable witnesses (a Type I error).

The findings of the above studies have implications which extend beyond the administration of the GSS and which apply to forensic interviewing more generally. According to Baxter (2004):

"... professional interviewers who do not carefully monitor and control their demeanour risk applying inappropriate levels of IP *unintentionally*. Interviewers may bias witnesses' recall as a result, particularly by causing them to change answers which they would not have changed in the absence of interrogative pressure" (p. 305).

This assertion is supported by research findings showing that a negative interviewing demeanour may entrench the attitude of a hostile witness or suspect. For example, Holmberg and Christianson (2002) found that convicted murderers and sexual offenders

were more likely to have confessed to their offences if they had been interviewed courteously. In contrast, a dominant, condemning interviewing style, characterised by impatience and aggression, was associated with denials of guilt.

Most of the studies examining the effects of interviewer demeanour on interviewee responding have used the standard GSS procedure in which IP is applied both through negative feedback and suggestive questioning. As previously noted, these two aspects of IP seem to operate somewhat independently: the GSS measures, Shift and Yield, correlate poorly and, when factor analysed, they load on separate factors (Gudjonsson, 1984). Reactions to negative feedback appear to depend primarily on social factors (Bain & Baxter, 2000; Bain et al., 2004; Baxter & Boon, 2000; Baxter et al., 2003; Gudjonsson, 2003; Gudjonsson & Lister, 1984) and indicate how individuals cope with IP. In contrast, responses to leading questions seem to be mediated more by cognitive processes such as discrepancy detection, i.e. the ability to discriminate satisfactorily between what they observed and what is suggested to them (Schooler & Loftus, 1986; Tousignant et al., 1986).

It was argued in Study 1 of this thesis that the findings of studies using the standard GSS procedure may have limited relevance for interviews in which a narrower range of pressures, or primarily social pressures, is applied. That is, professional interviewers may pressurise a witness or suspect through explicit or implicit forms of feedback and without asking overtly leading questions. Similarly, interviewers may apply IP through a

negative manner during questioning while avoiding leading questions.

In a study based on the GSS procedure, Baxter et al. (2006) assessed the impact of two interviewing styles on recall of a spoken narrative. It was found that, in the absence of overtly leading questions, negative feedback was still associated with response change, although lower than the norm for the GSS, and a firm demeanour produced more response change than a friendly demeanour. The present study was conducted to replicate this finding with a couple of methodological changes aimed at further improving reliability and validity. As discussed elsewhere in this thesis, a spoken narrative stimulus may be of limited use in studying the effects of IP on recall of an event. Hence, in line with Studies 1 and 2, a videotaped event was used as stimulus material.

Also, Baxter et al. used six individual undergraduate students (one male, five female, age range 20-21) to portray the friendly and firm interviewing styles in their study. While analysis of participant ratings confirmed that the interviewers were successful in portraying each style, the extent to which interviewing manner may have varied across the six interviewers is not clear. In addition, the use of young, primarily female, adults as interviewers may limit the generalisability of Baxter et al.'s findings to real forensic interviews in which interviewing is typically conducted by males (Baldwin, 1993; Moston et al., 1992) in an older age group. (In a recent large scale survey in North America, the proportion of male and female interviewers was reported as 86% and 14%

respectively, and the mean age of police investigators was found to be 40.68 years (Kassin et al., 2007)). The present study sought to address these issues by using one mature male interviewer.

In Study 2, no relationship was found between level of self-esteem and response change. In light of previous work showing that such a relationship appears to exist (Bain et al., 2004; Baxter et al., 2003; Gudjonsson & Singh, 1984b; Gudjonsson & Lister, 1984; Singh & Gudjonsson, 1984), this was an unexpected finding. It was noted, however, that previous studies have varied in the methods used to assess the self-esteem of participants and that the present study differed from previous work in its use of the RSE (Rosenberg, 1965) with a "normal" population (cf. Smith & Gudjonsson, 1995). Additionally, the design of Study 2 did not take into account variations in interviewer behaviour to which interviewees low in self-esteem may be particularly sensitive. It may be that the effects of low levels of self-esteem are more likely to be revealed where social pressures are brought to bear upon interviewee responding (see Bain et al., 2004; Baxter et al., 2003). In other words, self-esteem may interact with interviewer behaviour to affect rates of response change. The second aim of the present study was to test this possibility.

The final aim of this study relates to the measurement of self-reported anxiety in simulated interrogative contexts. The STAI (Spielberger, 1983; Spielberger et al., 1970) was administered in Study 1 to participants at the end of the procedure following the second phase of questioning. The key aim was to investigate the influence of negative

feedback on levels of state anxiety. In line with the Gudjonsson and Clark (1986) model and, as predicted, participants receiving negative feedback reported higher levels of state anxiety than those receiving neutral feedback. In contrast, it was not anticipated that trait anxiety would differ between experimental groups. However, participants interviewed by two interviewers reported significantly higher levels of trait anxiety than those interviewed by a single interviewer. Although possible, it seems unlikely that this group difference occurred either as a result of the experimental manipulation or by chance alone and it was argued that this finding may raise a question about the construct validity of the trait scale of the STAI, at least as it relates to simulated forensic interviews.

Given the role of anxiety in mediating interviewee responses during interrogations (Gudjonsson, 2003; Hansdottir, Thorsteinsson, Kristinsdottir & Ragnarsson, 1990), and the potential use of the STAI in future studies of this kind, it was considered important to test the validity of the instrument, in particular the trait scale. Hence, the present study included both pre- and post-interview measures of state and trait anxiety in order to investigate this issue further. It was expected that IP applied to participants in this study would result in an increase in state anxiety, but no increase in trait anxiety.

This study, therefore, had three hypotheses. It was predicted that participants interviewed in an abrupt manner would change significantly more responses and rate the interview as more difficult overall than participants interviewed in a friendly style. It was also predicted that high and low self-esteem individuals would respond

differentially to IP such that response change and interview difficulty ratings would be highest for the low self-esteem participants. Finally, in comparing pre- and post-interview anxiety scores, it was anticipated that there would be a significant increase in state anxiety and no significant increase in trait anxiety.

## 6.3 Method

### 6.3.1 Design

The study had a two-factor, between participants design. The independent variables were interviewer behaviour (friendly or abrupt), and self-esteem (low or high).

### 6.3.2 Participants

The interviewer was the same 49 year old male used previously in Studies 1 and 2. Before the study commenced, his ability to consistently adopt two distinct interviewing styles – Friendly and Abrupt – was assessed. Video recordings were made of the interviewer adopting each manner while interviewing one participant. In portraying the Friendly and Abrupt styles, the interviewer followed the description of these styles by Bain and Baxter (2000) (and replicated by Bain et al. (2004), and Baxter et al. (2003)).

The video recordings were rated by two independent groups of psychology undergraduates. Ratings of 18 aspects of the interviewer's manner were made on the Interviewer Rater form used in Study 1a. Table 6-1 shows the mean ratings made by each group.

One-way between-subjects ANOVAs showed that raters viewing the Friendly interview gave significantly higher ratings on the demeanours friendly, understanding, assertive, respectful, positive, warm, effective, and competent. Those viewing the Abrupt interview gave significantly higher ratings on severe, formal, stern, and negative.

Ratings for nervous, confident, professional, firm, organised, and authoritative were not significantly different between the two conditions.

**Table 6-1.** Pre-study interviewer behaviour ratings by condition.

| | Condition | | | |
|---|---|---|---|---|
| | Friendly (N = 36) | | Abrupt (N = 30) | |
| | M | SD | M | SD |
| Nervous | 1.69 | 0.95 | 1.90 | 0.99 |
| Severe** | 2.64 | 0.93 | 3.40 | 0.89 |
| Friendly*** | 2.83 | 0.85 | 1.43 | 0.50 |
| Understanding*** | 3.17 | 0.74 | 2.13 | 0.63 |
| Assertive* | 4.03 | 0.70 | 3.53 | 0.90 |
| Confident | 4.00 | 0.76 | 3.70 | 0.65 |
| Professional | 3.75 | 0.77 | 3.53 | 0.73 |
| Firm | 3.94 | 0.79 | 4.17 | 0.53 |
| Respectful*** | 3.42 | 0.69 | 2.77 | 0.63 |
| Positive** | 2.92 | 0.81 | 2.43 | 0.63 |
| Formal** | 3.44 | 1.11 | 4.13 | 0.73 |
| Warm*** | 2.08 | 0.91 | 1.40 | 0.50 |
| Stern*** | 3.14 | 0.96 | 4.03 | 1.00 |
| Organised | 4.17 | 0.74 | 3.87 | 0.90 |
| Effective** | 3.64 | 0.76 | 3.00 | 0.83 |
| Authoritative | 4.03 | 0.74 | 4.03 | 0.81 |
| Competent** | 4.00 | 0.68 | 3.33 | 0.92 |
| Negative** | 2.17 | 0.81 | 2.87 | 0.90 |

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

For the study itself, participant interviewees were recruited from first-year undergraduate psychology and engineering classes, and evening classes for mature students. The sample had a mean age of 25 years (SD = 12.65, range = 17-69) and

comprised 23 males and 60 females. Forty two and forty one participants were randomly assigned to the Friendly and Abrupt conditions respectively. Participants were further divided approximately equally to the High and Low Self-Esteem groups according to their score on the RSE (details below).

### 6.3.3 Materials

#### 6.3.3.1 Videotaped event

The stimulus was the same 77 second videotaped clip used in Studies 1 and 2 of this thesis. A description of the event is provided in Section 3.3.3.1.

#### 6.3.3.2 Rosenberg Self-Esteem Scale (RSE)

A description of this instrument is provided in Section 5.3.3.2. It is shown in Appendix K.

#### 6.3.3.3 Interview Rating Form

A description of this form is provided in Section 3.3.3.3 and it is shown in Appendix I. Previous use of this form in Studies 1 and 2 showed it to discriminate between different interviewee experiences and to possess high internal consistency.

#### 6.3.3.4 Interviewer Rater Form

A description of this form is provided in Section 4.3.3.2 and is shown in Appendix J.

#### 6.3.3.5 State and Trait Anxiety Inventory (STAI)

A description of this instrument is provided in Section 3.3.3.2.

### 6.3.4 Procedure

A random sample of 627 students completed the RSE. Of these, 534 were returned completed with an e-mail address or telephone number indicating that the respondent was willing to participate in the full study. The 42 highest and lowest scorers on the scale were invited for further participation. The mean self-esteem scores for the High and Low groups were 34.73 (SD = 2.64, range = 32-40) and 26.23 (SD = 1.95, range = 20-28) respectively. Comparison with the norms (see Appendices L and M) shows that a proportion of the High group scores fell below the 75$^{th}$ percentile, and that some of Low Group scores were slightly above the 25$^{th}$ percentile. Coefficient alpha for the scale was .87.

After reading and signing a consent form, half of the participants completed the SAI and the other half completed the TAI to provide a pre-interview anxiety measure. The participants were then asked to watch the videotaped event, after which they sat at the desk, facing the interviewer across from them. Each interview was video recorded.

The interviewer portrayed either a Friendly or Abrupt demeanour as described by Bain and Baxter (2000). In the Friendly condition, the interviewer smiled when participants entered the room, thanking them for taking part. He attempted to respond in a friendly manner to any conversation initiated by the participant prior to testing, and to maintain this manner in explaining the procedure to participants. The interviewer smiled frequently and always smiled back if participants smiled. A body position of leaning

back, away from the table and the participant, was adopted for the entire period of the experiment. Eye contact was maintained throughout testing, except when consulting the scoring sheet.

In the Abrupt condition, no attempt was made to build rapport or to be friendly when the participant entered the room. The interviewer gave minimal responses to any attempts at conversation by the participant and limited speaking to issuing instructions in an abrupt manner. The interviewer did not smile or make any facial response to anything the participant said. Instead, an expression intended to convey mild annoyance was maintained throughout the experiment. The interviewer adopted a body position of leaning forward across the table towards the participant. Again, eye contact was maintained by the interviewer throughout the period of testing, except when it was necessary to consult the scoring sheet.

After initial questioning, the interviewer delivered negative feedback in a firm, formal fashion then resumed the same interviewing manner, either friendly or abrupt, during requestioning. Participants then completed the Interview Rating Form and the Interviewer Rater form. Finally, to provide a post-interview anxiety measure, participants completed the same anxiety scale as before (either the State Anxiety Inventory or Trait Anxiety Inventory). In all other respects, the procedure used was the same as that used in Study 1.

#### 6.3.4.1 Scoring

Scoring of memory recall, response change and number of responses changed toward inaccuracy were calculated according to the methods detailed in Study 1 (see Section 3.3.4.1).

## 6.4 Results

### 6.4.1 Memory recall

Analysis of the overall mean recall scores found no significant differences in recall between experimental groups. The mean recall scores for both closed and general/open-ended items, shown in Table 6-2, were also analysed separately. A two-way between-subjects ANOVA on the closed item scores found a main effect of interviewer behaviour ($F(1,79) = 6.26$, $p<0.05$). Participants interviewed in an Abrupt style showed significantly higher scores (mean = 12.61) than those interviewed in a Friendly style (mean = 10.57). Level of self-esteem was not related to this measure and there was no interaction between interviewer behaviour and self-esteem. Analysis of memory recall on the open-ended items found no main effect of interviewer behaviour and no interaction. However, there was a main effect of self-esteem ($F(1,79) = 4.21$, $p<0.05$). Participants in the high self-esteem group obtained significantly higher scores (mean = 20.40) than those in the low self-esteem group (mean = 16.98).

### 6.4.2 Interview ratings

A reliability analysis showed the Interview Ratings form to have high internal consistency (Cronbach's alpha = .84). The item ratings were summed to give an overall measure of interview difficulty. The range of scores was 10 to 50 with a higher score indicating more perceived difficulty. Table 6-2 shows the mean total interview difficulty scores. A two-way between-subjects ANOVA showed a main effect of self-esteem ($F(1,78) = 14.32$, $p<0.001$) on the interview difficulty scores. Participants in the low self-esteem group reported higher levels of interview difficulty (mean = 30.40) than

those in the high self-esteem group (mean = 25.40). Interviewer behaviour did not affect perceived interview difficulty or interact with self-esteem.

**Table 6-2.** Mean (standard deviation) memory recall and interview ratings by condition.

| Self esteem | Interviewer behaviour | | | |
|---|---|---|---|---|
| | Friendly | | Abrupt | |
| | Low (N = 22) | High (N = 20) | Low (N = 21) | High (N = 20) |
| Recall (closed items) | 11.00 (4.62) | 10.10 (3.73) | 11.95 (3.58) | 13.30 (2.89) |
| Recall (open-ended items) | 18.14 (7.72) | 20.05 (7.84) | 15.76 (7.01) | 20.75 (8.05) |
| Total recall | 29.14 (10.76) | 30.15 (10.46) | 27.71 (9.21) | 34.05 (9.85) |
| Rated difficulty | 29.95 (4.94) | 26.80 (6.31) | 30.86 (5.69) | 24.00 (6.89) |

### 6.4.3 Response change

As noted above, memory recall performance on both closed and open-ended items varied as a function of experimental condition. To determine the effect this may have had on response change, separate response change scores were calculated for the closed and open-ended items. Again, because the number of questions asked varied between participants, these were computed as a percentage of the number of closed and open-ended items answered. A main effect of interviewer behaviour was found on the closed items ($F(1,79) = 4.31$, $p<0.05$). Abrupt interviewer groups (mean = 12.76) changed significantly more closed item responses than Friendly interviewer groups (mean = 8.44). There was no effect of level of self-esteem and no interaction. Analysis of response change on the open-ended items showed no significant differences between

experimental groups and no interaction.

**Table 6-3.** Mean (standard deviation) percentage response change by condition and question type.

| | Interviewer behaviour | | | |
|---|---|---|---|---|
| | Friendly | | Abrupt | |
| Self esteem | Low (N = 22) | High (N = 20) | Low (N = 21) | High (N = 20) |
| % Response change (closed items) | 8.57 (9.48) | 8.29 (8.07) | 14.11 (10.17) | 11.35 (9.81) |
| % Response change (open-ended items) | 20.29 (20.81) | 17.36 (16.44) | 25.00 (18.29) | 23.71 (23.18) |
| % Total response change | 11.55 (10.89) | 10.41 (7.42) | 16.68 (10.77) | 14.32 (11.54) |

To assess the extent to which negative feedback with an abrupt interviewing manner may have increased response change scores above those obtained with negative feedback alone, a between-groups comparison using data from Study 1 was conducted. Using a one-way ANOVA, levels of overall response change in the Abrupt conditions in the present study (N = 41, mean = 15.53) were compared with response change scores in the negative feedback/one-interviewer condition in Study 1 (N = 22, mean = 12.11). The difference, although in the expected direction, was not significant.

### 6.4.4 Response change toward inaccuracy

Across the four experimental groups, a total of 274 responses were changed. Of these, 111 (40.51%) were changed towards accuracy and 125 (45.62%) away from accuracy. The remaining 38 (13.87%) responses were changed neither towards nor away from accuracy. Examples of changes are given in Section 3.3.4.1 and also in Appendices E and G.

**Table 6-4.** Mean (standard deviation) response change (RC) toward inaccuracy by condition and question type.

| | Interviewer behaviour | | | | |
|---|---|---|---|---|---|
| | Friendly | | Abrupt | | |
| Self esteem | Low | High | Low | High | Sig. |
| *Closed items* | | | | | |
| RC as % of no. of questions | 3.93 (6.32) | 4.20 (5.04) | 6.79 (7.45) | 6.24 (7.15) | n.s. |
| RC as % of no. of changes | 36.67 (41.04) | 52.56 (39.00) | 45.99 (38.81) | 47.82 (39.37) | n.s. |
| *Open-ended items* | | | | | |
| RC as % of no. of questions | 6.17 (9.87) | 7.11 (9.27) | 10.34 (12.70) | 10.14 (12.99) | n.s. |
| RC as % of no. of changes | 34.62 (38.16) | 40.38 (40.23) | 40.20 (41.27) | 42.44 (42.99) | n.s. |
| *All items* | | | | | |
| RC as % of no. of questions | 4.47 (5.90) | 4.89 (4.83) | 7.61 (6.47) | 7.16 (6.61) | $p<0.05$ |
| RC as % of total changes | 33.30 (30.27) | 42.99 (30.88) | 44.21 (29.65) | 42.89 (31.87) | n.s. |

Table 6-4 shows response change toward inaccuracy scores separately for closed and open-ended questions. These were computed as a percentage of the number of closed or open-ended questions answered, and also as a percentage of the number of closed or open-ended response changes. Two-way between-subjects ANOVAs were conducted on these measures, and no significant group differences or interactions were found.

Analysis was also conducted on the closed and open-ended questions combined. On the first measure, response change toward inaccuracy as a percentage of the number of questions, a main effect of interviewer behaviour was found ($F(1,79) = 4.23$, $p<0.05$). Abrupt interviewer groups (mean = 7.39) were more likely to change away from accuracy than Friendly interviewer groups (mean = 4.67). Level of self-esteem did not affect this measure and no interaction was found. On the second measure, response change toward inaccuracy score as a percentage of the number of response changes, no significant differences or interactions were found between groups.

### 6.4.5 Interviewer behaviour ratings

Table 6-5 shows the mean interviewer behaviour ratings in the Friendly and Abrupt conditions. A one-way between-subjects ANOVA on these ratings revealed significant differences between conditions on the following adjectives: severe ($F(1,79) = 12.20$, $p=0.001$); friendly ($F(1,79) = 7.08$, $p<0.01$); professional ($F(1,79) = 4.64$, $p<0.05$); firm ($F(1,79) = 7.15$, $p<0.01$); warm ($F(1,79) = 8.15$, $p<0.01$); stern ($F(1,79) = 6.96$, $p<0.05$); authoritative ($F(1,79) = 10.78$, $p<0.01$); and negative ($F(1,79) = 4.81$, $p<0.05$). Ratings for nervous, understanding, assertive, confident, respectful, positive, formal, organised, effective and competent were not significantly different between the two conditions. Participants in the Friendly condition rated the interviewer higher on the descriptions friendly and warm. Those in the Abrupt condition rated the interviewer higher on the descriptions severe, professional, firm, stern, authoritative and negative.

**Table 6-5.** Mean interviewer behaviour ratings by condition.

| | Condition | | | |
|---|---|---|---|---|
| | Friendly (N = 42) | | Abrupt (N = 41) | |
| | M | SD | M | SD |
| Nervous | 1.07 | 0.26 | 1.02 | 0.16 |
| Severe*** | 1.95 | 1.01 | 2.71 | 0.93 |
| Friendly** | 3.55 | 1.06 | 2.95 | 1.05 |
| Understanding | 3.36 | 1.06 | 3.10 | 1.14 |
| Assertive | 3.95 | 0.70 | 4.12 | 0.81 |
| Confident | 4.50 | 0.55 | 4.66 | 0.53 |
| Professional* | 4.45 | 0.55 | 4.71 | 0.51 |
| Firm* | 3.88 | 0.77 | 4.32 | 0.69 |
| Respectful | 4.07 | 0.68 | 4.02 | 0.88 |
| Positive | 3.64 | 0.79 | 3.51 | 1.12 |
| Formal | 3.86 | 0.84 | 4.20 | 0.85 |
| Warm** | 3.05 | 0.88 | 2.44 | 1.07 |
| Stern* | 2.55 | 1.13 | 3.20 | 1.10 |
| Organised | 4.36 | 0.66 | 4.56 | 0.55 |
| Effective | 4.10 | 0.73 | 4.29 | 0.68 |
| Authoritative** | 3.71 | 1.09 | 4.37 | 0.66 |
| Competent | 4.45 | 0.55 | 4.46 | 0.67 |
| Negative* | 1.52 | 0.67 | 1.88 | 0.81 |

* $p<0.05$, ** $p<0.01$, *** $p\leq0.001$

### 6.4.6 Anxiety

Pre-interview and post-interview anxiety scores are shown in Tables 6-6 and 6-7. A three-way mixed ANOVA on the state anxiety scores revealed a within-subjects main effect. There was a significant increase from pre- to post-interview state anxiety across all conditions ($F(1,40) = 68.32$, $p<0.001$). No other significant differences or interactions were found. A three-way mixed ANOVA on the trait anxiety scores showed a between-

subjects main effect of self-esteem. The low self-esteem groups (mean = 40.20) were significantly higher on trait anxiety than the high self-esteem groups (mean = 30.74). Pre- and post-interview trait anxiety scores did not differ significantly and no other main effects or interactions were found.

**Table 6-6.** Mean (standard deviation) pre- and post-interview state anxiety scores.

| | Interviewer behaviour | | | |
|---|---|---|---|---|
| | Friendly | | Abrupt | |
| Self esteem | Low | High | Low | High |
| Pre-interview state anxiety | 37.08 (7.56) | 33.60 (9.38) | 39.09 (4.11) | 35.36 (7.57) |
| Post-interview state anxiety | 49.83 (7.13) | 48.80 (12.69) | 52.00 (7.78) | 44.27 (13.30) |

**Table 6-7.** Mean (standard deviation) pre- and post-interview trait anxiety scores.

| | Interviewer behaviour | | | |
|---|---|---|---|---|
| | Friendly | | Abrupt | |
| Self esteem | Low | High | Low | High |
| Pre-interview trait anxiety | 41.20 (7.74) | 30.50 (2.84) | 39.20 (7.30) | 32.44 (5.20) |
| Post-interview trait anxiety | 41.40 (6.70) | 28.50 (2.92) | 39.00 (7.72) | 31.78 (4.79) |

## 6.5 Discussion

The interviewer behaviour ratings made by participants in the Friendly and Abrupt conditions were significantly different on eight of the 18 adjectives used to describe the interviewer's manner (see Table 6-5). Participants in the Friendly conditions tended to rate the interviewer as more friendly and warm. Those in the Abrupt conditions rated the interviewer as more severe, professional, firm, stern, authoritative and negative. These ratings suggest that the experimental manipulation of interviewer demeanour was successful.

Overall memory recall scores did not differ significantly between experimental conditions, i.e. each group appeared to have comparable memories for the videotaped event (see Table 6-2). It is worth noting that studies using the standard GSS procedure employ a free recall task prior to questioning and, because variations in interviewing manner are not operative at this stage, interviewer behaviour does not influence free recall performance in these studies (see Bain & Baxter, 2000; Bain et al., 2004; Baxter & Boon, 2000; Baxter et al., 2003). In contrast, recall scoring in the present study was based on answers given during the pre-feedback questioning phase. As such, there was some possibility that variations in interviewer behaviour could have influenced cued recall performance. With regard to overall memory scoring, this was not the case. However, consistent with Studies 1 and 2 of this thesis, separate memory recall scores were calculated for both closed and general/open-ended items and, here, between-group differences were found (see Table 6-2). Interviewer behaviour had a significant effect on

recall of the closed items: participants interviewed in an Abrupt style gained higher scores on these questions than those interviewed in a Friendly manner. Interviewer behaviour had no effect on recall of the general/open-ended items.

By their nature, closed questions tend to lead respondents in a particular direction (Oppenheim, 1992; Schuman & Presser, 1979). Rather than replying in their own words, respondents to closed questions often must select from a series of pre-assigned categories the answer that best fits their own viewpoint (Payne, 1980; Vinten, 1995). Hence, in contrast to open "free response" questions which require respondents to structure their own answer, it is assumed that closed questions provide more retrieval cues. The following closed questions used in the present study illustrate this point:

Q2. Was he wearing a jacket?

Q8. Could he have taken the case?

Q9. Did he leave the scene to the left or to the right?

These questions may have provided specific cues to interviewees, namely what the man in question may have been wearing, that he may have stolen the case, and that he left the scene. Also, the restricted, dichotomous nature of closed questions means that, compared with open-ended questions, there is a greater likelihood of achieving a correct answer irrespective of cognitive effort employed. That closed questions afford these "advantages" to respondents is closely related to the finding that people tend to perform

better on formal tests of recognition than they do on recall tests (Baddeley, 1990, 1999). It may be then that the abrupt questioning style increased the vigilance of participants leading them to attend more closely to internal rather than external cues and, under these circumstances, performance on the closed questions was influenced differentially.

In Study 2 of this thesis, low and high self-esteem participants did not differ in their recall of the video taped event. This finding was partially replicated in the present study. Level of self-esteem had no effect on recall of the closed items, but it was related to recall of the open-ended items: high self-esteem participants obtained higher scores on the open-ended questions than low self-esteem participants. Given the similarities between the present study and Study 2, it is not entirely clear why this occurred. Both studies used the RSE (Rosenberg, 1965) to measure self-esteem and the stimulus, question set, and the interviewer were the same. The studies differed to a degree only in the composition of their samples. The Study 2 sample comprised undergraduate students while the present study used undergraduates and a proportion of mature students.

According to Tafarodi and Milne (2002), self-esteem comprises two interdependent dimensions, self-liking and self competence. The latter is described as "the valuative experience of oneself as a causal agent, an intentional being with efficacy and power" (p. 444), and in this conceptualisation, self-esteem would seem to be related to the construct of self-efficacy (Bandura, 1977, 1986, 1989, 1992). Indeed, a number of investigations of self-esteem and self-efficacy have failed to fully discriminate between

the two constructs (e.g. Bernard, Hutchison, Lavin & Pennington, 1996; Stanley & Murphy, 1997). As such, individuals with high self-esteem may be more likely than low self-esteem individuals to possess a sense of personal competence and to expend effort on difficult tasks. As previously discussed, open-ended questions impose a greater cognitive load than tasks based on closed questions and there would seem to be more scope for error in answering open-ended questions compared with the more limited "forced-choice" options presented in closed questions. Faced with the difficulties of answering the open-ended questions, the high self-esteem participants in the present study may have had more confidence in their ability to provide a correct answer than the low self-esteem participants and hence may have engaged more constructively in the task to achieve a successful outcome. A related explanation is that the high self-esteem participants may have been distracted less by concerns about their self-efficacy and, as a result, were able to perform better on the recall task than low self-esteem participants.

As stated in Study 1, recall is a secondary measure intended only to allow a basic comparison between the event memory available to each group. Recall scoring does, however, serve as a baseline measure against which group comparisons of response change are made. In this respect, it is important that a variation in recall between groups is at a minimum if valid conclusions are to be drawn. In the present study, overall memory for the videotaped event did not vary as a function of experimental condition, i.e. differences in recall were apparent only when a distinction was made between question types. To the extent that there may exist some underlying variation in recall

between groups, however, interpretation of response change data must bear this in mind.

Participants interviewed in an abrupt manner changed significantly more responses overall than participants interviewed in a friendly style (see Table 6-3). This result supports the findings of Bain and Baxter (2000), Bain et al. (2004), Baxter and Boon (2000), and Baxter et al. (2003). These studies used the standard GSS procedure which employs both negative feedback and leading questions to apply IP. The present study shows that, even in the absence of overtly leading questions, negative feedback is still associated with response change and that response change can be further increased through a negative interviewing manner. In this respect, the present findings are also consistent with those of Baxter et al. (2006), and the first hypothesis appears to be supported.

Comparing the findings of Study 1 with the present study, it can be argued that negative feedback together with an abrupt interviewing manner increased levels of IP to a small degree above those obtained with negative feedback alone: in Study 1, half of the participants received negative feedback alone. Of these, 65.91% changed two or more responses; in the present study, half of the participants received negative feedback and were interviewed abruptly. Of these, 82.93% changed two or more responses. A formal analysis of the response change data showed this difference in levels of IP not to be significant, however.

Because memory recall scores varied across experimental conditions according to the type of question asked, rates of response change for closed and open-ended questions were examined separately. Following negative feedback, the experimental groups did not differ significantly in their tendency to change open-ended responses. However, participants interviewed in an abrupt manner changed significantly more closed responses than participants interviewed in a friendly manner (see Table 6-3). This is perhaps a surprising finding to the extent that the abrupt questioning manner appeared to improve the pre-feedback recall performance on the closed questions. Because participants in the Abrupt conditions appeared to have better memories for the videotaped event and, presumably, higher certainty in their recall, it can be argued that they should have been less likely than participants in the Friendly conditions to change their initial responses (Gudjonsson & Clark, 1986). The fact that the Abrupt groups actually changed significantly more closed responses raises the possibility that they acted in a compliant, rather than a suggestible manner, opting to change answers during requestioning without any accompanying belief that their initial answers were incorrect. Further, it may be that response change in this instance was aided by the dichotomous or "forced-choice" nature of the closed items, i.e. less cognitive effort is required to change a closed response – such as "left" to "right" – than is necessary to generate an alternative answer to an open-ended question.

With regard to response change, it should be noted that 11 of the 83 participants in this

study did not change any answers during requestioning. Across the three main studies of this thesis, the percentage of participants changing no responses following negative feedback was 9.09% (Study 1), 20.45% (Study 2), and 13.25% (Study 3). It may be that, following cognitive appraisal of the negative feedback, these participants chose to reject the feedback and adopt a resistant attitude (Gudjonsson & Clark, 1986). Another possibility, however, is that the participants did not adequately process the negative feedback and, despite the implicit negative feedback communicated during requestioning, opted simply to repeat their initial answers.

Interestingly, although variation in interviewer manner affected participant behaviour with regard to response change, it did not appear to influence perceptions of how difficult the interview was. Interview difficulty ratings were, in fact, closely comparable across Friendly and Abrupt conditions (see Table 6-2).

As discussed in Study 2 of this thesis, previous studies using the GSS 1 have shown that low levels of self-esteem appear to be related to the tendency to shift responses following negative feedback (Bain et al., 2004; Baxter et al., 2003; Gudjonsson & Lister, 1984; Gudjonsson & Singh, 1984b; Singh & Gudjonsson, 1984). In Study 2, however, no association was found between interviewee self-esteem and response change. It was noted that Study 2 differed from previous work with regard to stimulus material, self-esteem measure and degree of IP applied to participants. Regarding the last of these, negative feedback was communicated to participants in Study 2 in the absence of

leading questions. It is assumed that IP was lessened as a result. In the present study, it was anticipated that negative feedback together with an unfriendly interviewing manner would increase uncertainty and anxiety in low self-esteem participants motivating them to reduce psychological distance between themselves and the interviewer (Gudjonsson & Lister, 1984). This in turn would elevate levels of response change above those expected with negative feedback alone. Hence, it was hypothesised that level of self-esteem would affect response change with low self-esteem participants scoring highest on this measure.

This hypothesis was not supported. Low and high self-esteem groups scored similarly on response change and interviewee self-esteem did not interact with interviewer manner to affect response change (see Table 6-3). This result appears to contradict the findings of the studies cited above that have shown that individuals with lower levels of self-esteem are particularly vulnerable to IP. One explanation for this finding may relate to the degree of IP applied in this study relative to previous work. As noted above, in the studies conducted by Bain et al. (2004), Baxter et al. (2003), Gudjonsson and Lister (1984), Gudjonsson and Singh (1984b), and Singh and Gudjonsson (1984), IP was applied through negative feedback and overtly leading questions. By dispensing with the latter in this study (and in Study 2 of this thesis), it must be assumed that the level of IP was reduced. Indeed, Baxter et al. (2006) noted that mean Shift scores on an amended GSS 2 (overtly leading questions were removed) were significantly lower than the population norm for the standard GSS (Gudjonsson, 1997).

It was suggested in discussing Study 2 that the absence of overtly leading questions minimises the cognitive aspects of susceptibility to IP. Neutral questions, as used in the present studies, as opposed to the overtly leading questions used in GSS studies, do not provide specific cues to the "correct" answer. The results of the present study and that of Study 2 suggest that the social aspects of IP, whether applied solely by negative feedback, or by negative feedback *and* a negative interviewing manner, may not be sufficient on their own to elicit an effect of low self-esteem on response change.

Another important factor when considering the absence of a self-esteem effect on response change in this study (and Study 2) relates to the range of self-esteem scores. Firstly, scores for the original sample tested were somewhat negatively skewed, i.e. the distribution of scores was weighted towards the higher end of the scale. As previously noted, this is commonly observed in undergraduate samples assessed on the RSE (Blascovich & Tomaka, 1991) and means that, in an absolute sense, fewer participants could be considered as having relatively low self-esteem. A second, and related, consideration is perhaps an insufficient split between the low and high self-esteem groups. Using the standard deviation of the original sample as a measure of split, the difference between the overall mean score and the mean for the high self-esteem groups exceeded one standard deviation. However, the difference between the overall mean score and the mean for the low self-esteem groups was slightly less than one standard deviation. In future studies of this kind using the RSE, undergraduate samples may be

best avoided to perhaps render more likely a wider distribution of scores on which the lower end of the scale is better represented. Where an undergraduate population *is* employed, it may be necessary to increase considerably the size of the general sample from which low and high scorers are drawn. Note that although no difference in rate of response change was found between the two self-esteem groups, the group split in the present study did successfully discriminate between the self-esteem groups with regard to interview difficulty ratings and self-reported anxiety, as reported below.

An alternative and perhaps less likely explanation for the lack of a self-esteem effect on response change involves the measure of self-esteem employed. It was noted in Study 2 that previous work has used a range of methods to measure self-esteem making it difficult to draw comparisons between studies. Gudjonsson and Singh (1984b) used the Coopersmith Behaviour Form (Coopersmith, 1967), Gudjonsson and Lister (1984) and Singh and Gudjonsson (1984) used Osgood's Semantic Differential Scales (Osgood et al., 1957), while Baxter et al. (2003) and Bain et al. (2004) used the Battle Self-Esteem Inventory (Battle, 1981). The only previous study to use the RSE (Rosenberg, 1965) was Smith and Gudjonsson (1995). Although they found no relationship between self-esteem and measures of interrogative suggestibility and this appears to be in line with the findings of Study 2 and the present study, a direct comparison with the Smith and Gudjonsson study is problematic due to differences in the samples used. While the present study used an undergraduate student sample, Smith and Gudjonsson assessed a group of forensic psychiatric inpatients and they suggest that such a sample may be

more prone than a "normal" population to presentational biases in using the RSE, an instrument noted for its high face validity (Blascovich & Tomaka, 1991).

In Studies 1 and 2 of this thesis, it was noted that IP in the form of negative feedback was associated, to a degree, with the tendency to shift answers toward inaccuracy. The present study used two measures to examine whether an abrupt, compared to a friendly interviewing manner might also be related to this tendency. On one measure (response change towards inaccuracy as a percentage of questions answered), this was observed (see Table 6-4). However, no group differences were found in response change toward inaccuracy when responses to closed and open-ended items were analysed separately. With regard to real forensic interviews then, there is some limited support for the idea that not only negative feedback, but negative feedback together with an unfriendly manner has the potential to be a distorting influence on recall. Again, this result must be considered together with the aforementioned finding in this study that an abrupt questioning style on its own, prior to feedback, appeared to improve recall on closed questions.

The apparently positive impact of negative feedback on response change direction must also be taken into account. Consistent with Studies 1 and 2, a proportion of answers during requestioning were changed towards accuracy in the present study. Across the three studies, the percentage of answers changed towards accuracy varied between 37.14% and 40.51% while the percentage changed towards inaccuracy ranged from 40%

to 45.62%. (The remaining responses (13.87% to 22.86%) were judged to have been changed in no specific direction). It can be seen then that negative feedback is associated with response changes in both directions. The overall tendency, albeit marginal, is for answers to shift towards inaccuracy following negative feedback. Hence, professional interviewers who communicate negative feedback to interviewees may indeed prompt increases in the accuracy of testimony in some cases. The findings in this thesis, however, suggest that such a strategy is high risk at best and indeed may be more likely to result in the opposite outcome.

While level of interviewee self-esteem had no effect on levels of response change in this study, it did appear to influence ratings of interview difficulty. Consistent with the findings in Study 2, individuals lower in self-esteem rated the interview higher in difficulty than the high self-esteem groups (see Table 6-2). This result may reflect the tendency of low self-esteem individuals to have more severe affective and cognitive reactions to failure feedback than high self-esteem individuals (Brown & Dutton, 1995; Kernis, Frankel & Brockner, 1989). It may be that the low self-esteem groups engaged in "overgeneralisation following failure" (Carver & Ganellen, 1983) in response to the negative verbal feedback and requestioning which, by their nature, imply that participants have failed in the task of accurate recall. The above term describes a psychological process in which persons with low self-esteem have feelings and thoughts of personal inadequacy following a negative outcome. The elevated difficulty ratings made by the low self-esteem groups in the present study may have occurred as a

function of such global self-attributions. Another reason for differences in ratings of interview difficulty may relate to the relative attentional resources available to low and high self-esteem participants during questioning. It is possible that low self-esteem individuals are more likely than those with high self-esteem to be concerned with their affective state and so have an extra demand on their attention.

A question left unanswered in Study 1 related to the measurement of self-reported anxiety in simulated interrogative contexts. In Study 1, participants interviewed by two interviewers reported significantly higher levels of trait anxiety than those interviewed by a single interviewer. This was an unexpected finding. It was anticipated instead that the experimental groups would not differ significantly on reported trait anxiety. It seems unlikely that the two-interviewer group was generally more anxious than the one-interviewer group and so a question was raised about the use of the STAI, particularly the trait scale, in simulated forensic interviews.

In Study 1, state and trait anxiety were measured once at the end of the procedure, immediately following questioning. The absence of both pre- and post-interview measures of anxiety precluded a proper assessment of the STAI as it relates to an interrogative context. In the present study, the STAI was administered to participants before and after questioning to allow comparison of the effects of IP on state and trait anxiety. It was anticipated that IP applied to participants in this study would result in an increase in state anxiety, but no increase in trait anxiety. The findings upheld this

prediction and so the third hypothesis was supported (see Tables 6-6 and 6-7). At least on the basis of the present findings, it can be argued that the two scales of the STAI appear to be sensitive to the respective constructs that they purport to measure.

Clearly, the absence of pre-interview measures of anxiety limits the conclusions that can be drawn about the effects of various stressors present in simulated interviews, and precludes assumptions about the levels of trait anxiety held by experimental groups prior to participation. It was suggested in Study 1 of this thesis, however, that pre-interview measures of anxiety in studies of this kind may detract from attempts to closely simulate the dynamics of a forensic interview. Nevertheless, in instances where levels of self-reported anxiety are central to a particular study's hypothesis, pre- and post-interview anxiety comparisons may be useful and the present findings suggest that the STAI may be appropriate for such measurements.

# Chapter 7: General Discussion

## 7.1 Key results

The three main studies presented above used an experimental procedure adapted from the GSSs (Gudjonsson, 1984, 1987b). While the repeated questions and feedback aspects of the standard GSS procedure were retained in the present work, questions asked of participants were minimally rather than overtly leading and were applied to a video presentation of a simulated crime as opposed to a narrative passage. These amendments were made in the interests of improving ecological validity.

The first aim of this thesis was to investigate situational determinants of interrogative suggestibility. In Study 1, manipulation of the type of verbal feedback communicated to participants showed that negative feedback has the tendency to increase levels of response change above those associated with neutral feedback. In the same study, the use of negative feedback was also found to raise self-reported state anxiety and ratings of interview difficulty. These results support the stated hypothesis and are closely in accordance with the Gudjonsson and Clark (1986) model of interrogative suggestibility. Another finding of Study 1 was that, compared with neutral feedback, negative feedback increased the likelihood that answers changed would be changed in the direction of inaccuracy. Study 3 varied the manner of the interviewer during questioning and this too influenced the behaviour of the interviewees. An abrupt, as opposed to a friendly, questioning manner increased the accuracy of recall on closed questions while also

increasing the level of response change on these. It was argued that the application of IP increased interviewees' uncertainty and expectations, and diverted their attentional resources away from the task of accurate recall towards demand characteristics of the interview. Response change may have occurred as a consequence of attempts to reduce perceived psychological distance (Bain & Baxter, 2000; Baxter & Boon, 2000; Gudjonsson & Clark, 1986; Gudjonsson & Lister, 1984). These findings point to a need for forensic interviewers to closely monitor and attempt to control both their verbal communication and general demeanour during questioning. Failure to do so runs the risk of disproportionate levels of IP being applied unintentionally and the introduction of bias into what might otherwise be a straightforward exchange of information (Baxter, 2004).

The second aim was to assess the extent to which the presence and involvement of others in an interrogative setting might influence the pattern of interviewee responses. In Study 1, the presence and involvement of a second interviewer did not influence recall of the videotaped event and, contrary to the hypothesis, had no effect on response change, level of state anxiety or ratings of interview difficulty. In Study 2, the presence of an interviewee ally did not affect recall scores and, again contrary to expectation, the ally's intervention following the communication of negative feedback did not affect the rate of response change. Possible reasons for these findings are considered below and suggestions made for retesting these hypotheses.

As a third aim, the present work also examined the relationship between interrogative suggestibility and levels of interviewee self-esteem. In Study 2, the low self-esteem and high self-esteem groups gained closely comparable scores on the recall task and, while level of self-esteem influenced levels of state anxiety and ratings of interview difficulty, it was not associated with the tendency to change responses. These findings were replicated in Study 3 except that high self-esteem participants obtained higher recall scores on the open-ended questions than low self-esteem participants.

The failure of these studies to find a relationship between the susceptibility of interviewees to yield to IP and their level of self-esteem requires clarification. It may be that the relationship between these variables is not as robust as suggested by previous work on the Gudjonsson and Clark theoretical model and as anticipated at the outset of this thesis. It is worth reiterating that previous studies have been based on the GSS procedure in which IP is applied both by means of negative feedback and suggestive questioning. The latter was absent from the present studies and it can be assumed therefore that the level of IP was lower as a consequence. In comparing the present findings with those of previous studies, it is possible to conclude that the influence of self-esteem on suggestibility is likely to be lessened in instances where a narrower range of pressures is applied to interviewees. It may be informative in future work to test this interpretation by varying systematically not only the level of self-esteem, but also the level of IP applied to participants.

The final aim of this thesis was to further investigate the relationship between IP, interrogative suggestibility and anxiety. As previously noted, the application of IP by negative feedback in Study 1 was associated with increases in self-reported state anxiety. The comparison in that study was between the effects of neutral and negative feedback and only post-feedback measures of anxiety were obtained. Study 3 was designed to assess change in levels of anxiety across the question-feedback-requestion procedure. Measures of both state and trait anxiety were taken at pre-interview and post-interview. As predicted, there was a pre to post increase in state anxiety and no concomitant increase in trait anxiety. In Study 1, a positive correlation was found between state anxiety and interviewees' tendencies to change responses. Together, these results confirm the mediating effects of anxiety on interrogative suggestibility and support previous work showing that suggestibility is more closely associated with state than trait anxiety (e.g. Gudjonsson, 1988b; Haraldsson, 1985; Smith & Gudjonsson, 1995).

### 7.1.1 Practical and theoretical implications

Clearly, the experimental studies conducted for this thesis were focused on an area of application – forensic interviewing – and, to some extent, were addressed to practical questions of special interest to professional forensic interviewers and others working in the field of criminal justice. Research studies of this kind which are directed towards practical issues might be criticised on the grounds that they are somewhat atheoretical. (Small (1993, cited in Ogloff, 1999) has asserted that legal psychologists sometimes fail to employ or develop psychological theories to explain the phenomena they study in the legal system). At the same time, work motivated by purely theoretical concerns may

seem to have no obvious relevance to real-world problems. It can be argued, however, that the extremities of these two viewpoints are unhelpful and neglect the broader and truer intentions of applied experimental psychology. As Weiner (2003) states: "It is not essential that practical work be atheoretical or that theoretical work be divorced from applications" (p. 655).

The aims of this thesis were to contribute to both practical *and* theoretical interests. The practical implications include an improved understanding of the effects of IP in instances where overtly leading questions are not used. The results here suggest that many interviewees will change aspects of their original testimony following explicit negative feedback in the form of critical verbal statements and/or an unfriendly manner of interviewing as well as in response to implicit types of feedback such as the repetition of questions. These findings have clear relevance for investigative practice in instances where witnesses or suspects, particularly those vulnerable to suggestion, have provided an accurate and "true" representation of events during initial questioning; in such instances, shifts in answers may lead to a decrease in accuracy. The most useful application of the present findings might simply be to increase the awareness of professional forensic interviewers to the possibility that they may *inadvertently* pressurise witnesses or suspects through a communication of disapproval, whether this is verbal, non-verbal, or both. A general recommendation following from the work in this thesis for police interviewing is that, not only the content of questions, but the manner of questioning should be subject to close monitoring and control.

The present work found that levels of suggestible responding were not influenced by the presence and involvement of an additional interviewer nor were they impacted by the intervention of an interviewee ally. Hence, no obvious practical implications follow. As noted below, further experimental work is needed to determine whether these factors, which are common to many police interviews, have the potential to affect the suggestibility of interviewees. If this proves to be the case, such a finding would be of clear interest to those in the criminal justice system who seek to better understand the various influences that witnesses and suspects face during police interviewing.

As detailed in Chapter 2, the theoretical background for the present work is the Gudjonsson and Clark (1986) model of interrogative suggestibility. The strength of their theoretical propositions is reflected in the number of hypotheses that can be predicted from the model and the extent to which these are support by empirical findings (see Gudjonsson, 1991, 2003 and Chapter 2 of this thesis). In relation to Gudjonsson and Clark's model, the present work must be considered as confirmatory, i.e. the findings do not in any sense undermine the theoretical assumptions on which the model is based, nor do they implicate significant amendments to the model. Rather, the findings are corroborative in nature as they largely validate the results of previous studies.

The key theoretical statement arising from this thesis is that interrogative suggestibility

is dynamic in nature and situation bound. It was demonstrated that suggestibility, as measured by the number of responses changed during requestioning, varied as a function of type of verbal feedback and interviewer manner. Hence, although there is some acceptance in the literature that interrogative suggestibility can be regarded as a "trait", i.e. a stable tendency within individuals resulting in consistent patterns of behaviour (e.g. see Stukat, 1958), the present work confirms that levels of suggestibility can be significantly affected by situational factors.

A further theoretical contribution of the thesis relates to the composition of interrogative suggestibility. Gudjonsson (1984, 1991, 2003) argued that interrogative suggestibility comprises two components – leading questions (measured by Yield) and negative feedback (measured by Shift) – and that these are relatively independent on the basis that they are poorly correlated and tend to load on separate factors (Gudjonsson, 1984). As previously discussed, the former is assumed to be mediated by cognitive processes while the latter depends on social factors (Bain & Baxter, 2000; Bain et al., 2004; Baxter & Boon, 2000; Baxter et al., 2003; Gudjonsson, 2003; Gudjonsson & Lister, 1984). Because the present studies did not use overtly leading questions, it was possible to test the assumption that negative feedback would operate independently of suggestive questioning. That negative feedback increased response change in the absence of leading questions supports Gudjonsson's argument.

The Gudjonsson and Clark theoretical model predicts that negative feedback, assuming

it is accepted by the individuals it is communicated to, can affect their mood. Specifically, it is predicted that acceptance of negative feedback commonly causes a temporary increase in anxiety. Again, the results here support the model: comparisons of pre and post feedback measures of state anxiety showed a significant increase. It was also of theoretical interest to further examine the relationship between self-esteem and interrogative suggestibility. As discussed throughout this thesis, the bulk of research evidence, based on use of the GSSs, has indicated a strong association between these two variables (Bain et al. 2004; Baxter et al., 2003; Gudjonsson & Lister, 1984; Gudjonsson & Singh, 1984b; Singh & Gudjonsson, 1984). The present findings suggest that, when levels of IP are lowered beyond those applied in the GSS procedure, the self-esteem of interviewees may be less important as a mediating influence on their responses during questioning.

## 7.2 Methodological issues

There are aspects of experimental design in the present studies which may limit the generalisability of the findings to forensic settings. First, witnesses and suspects in real interviews are subject to a wider range of pressures than can be applied to experimental participants. They often face not only IP during questioning, but also custodial pressures surrounding the interview (or interviews), i.e. detention and confinement (Irving, 1986). Second, it should not be overlooked that while the participants in these studies were drawn from a "non-forensic", educated and presumably healthy demographic, where they exist, issues relating to the physical and mental health of witnesses and suspects

have the potential to influence their behaviour during questioning (Gudjonsson, 2006). Third, it is common for there to be a delay between the reporting of an incident and the interviewing of witnesses and suspects. Some memories during this time may be prone to decay and to interference by post-event information (Loftus & Loftus, 1980; Tuckey & Brewer, 2003). Participants in the present research were invited to recall details of the videotaped event almost immediately after witnessing it. It is possible, therefore, that their memories for the event were artificially strong to the extent that they would be less likely than real interviewees to rely on external cues present in the interview. Given these differences between laboratory conditions and what may occur in the field, it can reasonably be argued that the present results offer only conservative estimates of response change following the application of IP.

Two potentially important variables in the questioning procedure are the prior knowledge of the subject matter held by the interviewer and the degree to which such knowledge may be evident to the interviewee. Importantly, it is usual for police interviewers not to have witnessed the incident in question and to the extent that there is little or no physical evidence or other reliable witness accounts, the interviewer may be totally reliant on the account of the interviewee to further the criminal investigation. (It is clearly Inbau et al.'s (2001) belief that this places interviewers at a significant disadvantage when interrogating suspects. Where no evidence against a suspect exists, they recommend confronting the suspect with false evidence of their guilt). In real police interviews, the extent of an interviewer's knowledge about an incident can vary

considerably and, in some cases, it may be evident to interviewees that the interviewer has much less or perhaps more knowledge of an incident than they do.

From a psychological standpoint, it seems pertinent to ask whether such perceptions might influence the response pattern of interviewees. It may be that interviewees are generally more susceptible to IP when they perceive the interviewer to possess sure knowledge of the event in question. In the present studies, the videotaped event was presented to participants by the interviewer himself. Hence, it was clear to participants that the interviewer was familiar with its contents and this may have increased artificially the tendency of participants to shift their previous answers following feedback. A suggested methodological improvement for future studies to bring them closer to the dynamics of real interviews is to ensure that the interviewer is blind to the contents of the video presentation and that this is clearly communicated to participants.

## 7.3 Future work

As discussed throughout this thesis, a large body of research using the GSSs has shown that, in response to IP, many individuals will react in a suggestible or compliant manner when questioned about the content of a spoken narrative. The work presented here confirms that similar responses can be elicited when questioning is based on a videotaped event. The use of a video presentation rather than a narrative passage means that details are encoded in a manner similar to that of an event personally experienced by

a witness, i.e. both auditory and visual information is encoded. In studies using the GSSs, participants are essentially tasked with the retrieval of information from semantic memory. By contrast, the present studies required participants to retrieve details of an event from autobiographical episodic memory. While these two types of memory are conceptually and, it seems, neurophysiologically distinct (Tulving, 1989; Warrington, 1986), and processes of encoding and retrieval between them differ, such differences may have limited importance for the study of interrogative suggestibility. On the basis of the present findings, it can be argued that suggestibility effects, at least as they pertain to IP in the form of negative feedback and abrupt questioning style, are not restricted to spoken narrative stimulus material and, as argued by O'Connor (2001), may be fairly insensitive to mode of presentation. Further work might usefully extend and assess this viewpoint by testing the effects of IP upon recall of a real-life event. If it can be demonstrated that recall of a real event is similarly prone to the influences of IP, then it can be more confidently assumed that effects observed in GSS studies are not associated solely to the scale's stimulus material and that they have relevance beyond the laboratory.

As previously noted, research on social facilitation and social inhibition has demonstrated a relationship between social presence and task performance, and a number of psychological explanations have been proposed, e.g. drive theory (Zajonc, 1965), evaluation apprehension (Cottrell, 1972), distraction-conflict theory (Sanders, Baron & Moore, 1978), and self-awareness theory (Carver & Scheier, 1981). Hence,

there are sound theoretical reasons to support the prediction that interviewees facing two interviewers may perform differently during questioning than interviewees facing a single interviewer. Research of this kind is also of clear forensic interest: a large number of police interviews with suspects are conducted by two interviewers (Baldwin, 1993; Leo, 1996c; Moston et al., 1992) - a recent large study in London reported a figure of 79% (Medford et al., 2003). The present work did not find a relationship between the number of interviewers and the tendency to change responses following feedback. This is perhaps because the second interviewer played a relatively limited role in the questioning procedure. Future studies could assess the possibility that two more equally active interviewers might raise levels of IP above those associated with a single interviewer.

Despite the finding in this thesis that support from an interviewee ally did not reduce suggestible responding, it can be argued on both practical and theoretical grounds that further work on this would be of value. In many jurisdictions, suspects do not face the pressures of questioning alone. They are often accompanied by a legal representative or an appropriate adult whose role it is to uphold the suspect's legal rights and/or assist suspects identified as vulnerable. This was confirmed in a recent study by Medford et al. (2003). Examining a large number of interviews by the Metropolitan Police, they found that appropriate adults were present during 58% of adult suspect interviews and 99% of juvenile suspect interviews. It seems appropriate then that attention is devoted to better understanding the influence of such allies upon interrogative suggestibility.

Social support is widely accepted as an important environmental resource (Suls & Wallston, 2003) and, in its absence, the likelihood that individuals will yield to social pressure is increased (Asch, 1955, 1956). It appears that Binet (1905), an early investigator of suggestibility, was sensitive to this. He warned, "Woe betide him who is alone" (Malheur a celui qui est seul). A number of studies have shown that warnings to be vigilant during questioning, as communicated by an interviewer, can decrease suggestibility scores on the GSSs (Bain et al., 2004; Baxter et al., 2006; Boon & Baxter, 2000). These authors argued that by drawing interviewees' attention to the possible presence of misinformation, it is possible to alter the conditions of uncertainty, expectation and interpersonal trust to the extent that the likelihood of suggestible responding is reduced. It was in light of these findings that a similar prediction was formulated in this thesis with regard to warnings made by an interviewee ally. In the present work, the ally made a single, short intervention immediately following the delivery of negative feedback which may not have been explicit enough to heighten the vigilance of the participants and ultimately to influence their behaviour. It may be useful in future studies to vary not only the content of the warning, but also its timing and frequency and to check that the ally's intervention and presence during questioning is indeed perceived to be socially supportive by participants.

## References


Ainsworth, P. (1998). *Eyewitness Testimony*. Chichester: Wiley.

Ajzen, I. (1985). From intentions to actions: A theory of planned behaviour. In: J. Kuhl & J. Beckmann (Eds) *Action Control: From Cognition to Behaviour*. Heidelberg: Springer.

Ajzen, I. & Madden, T.J. (1986). Prediction of goal-directed behaviour: Attitudes, intentions, and perceived behavioural control. *Journal of Experimental Social Psychology*, 22, 453-474.

Arther, R.O. & Caputo, R.R. (1959). *Interrogation for Investigators*. New York: Copp.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258-290.

Asch, S.E. (1951). Effects of group pressure upon the modification and distortion of judgement. In: H. Guetzkow (Ed) *Groups, Leadership, and Men*. Pittsburgh: Carnegie.

Asch, S.E. (1955). Opinions and social pressure. *Scientific American*, 193, 31-35.

Asch, S.E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs*, 70 (9, Whole No. 416).

Baddeley, A.D. (1990). *Human Memory: Theory and Practice*. Needham Heights, MA: Allyn and Bacon.

Baddeley, A.D. (1999). *Essentials of Human Memory*. Hove: Psychology Press.

Bain, S.A. & Baxter, J.S. (2000) Interrogative suggestibility: the role of interviewer behaviour. *Legal and Criminological Psychology*, 5, 123-133.

Bain, S.A., Baxter, J.S. & Ballantyne, K. (2007). Self-monitoring style and levels of interrogative suggestibility. *Personality and Individual Differences*, 42, 623-630.

Bain, S.A., Baxter, J.S. & Fellowes, V. (2004). Interacting influences on interrogative suggestibility. *Legal and Criminological Psychology*, 9, 239-252.

Baldwin, J. (1992). *Video taping police interviews with suspects: A national evaluation. Police Research Series Paper 1*. London: Home Office Police Department.

Baldwin, J. (1993). Police interview techniques: establishing truth or proof? *British Journal of Criminology*, 33, 325-352.

Baldwin, J. (1994). Police interrogation: what are the rules of the game? In: D. Morgan & G.M. Stephenson (Eds) *Suspicion and Silence: The Right to Silence in Criminal Investigations*. London: Blackstone.

Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology, Special Issue: Self-Efficacy Theory in Contemporary Psychology*, 4, 359-373.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, 44, 1175-1184.

Bandura, A. (1992). Exercise of personal agency through the self-efficacy mechanism. In: R. Schwarzer (Ed) *Self-Efficacy: Thought Control of Action*. Washington, DC: Hemisphere Publishing Corporation.

Barnes, L., Harp, D. & Sik Jung, W. (2002). Reliability generalization of scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement*, 62, 603-618.

Bartlett, F.C. (1932). *Remembering*. Cambridge: Cambridge University Press.

Battle, J. (1981). *Culture-free SEI: Self-Esteem Inventories for Children and Adults*. Seattle, WA: Special Child.

Baxter, J.C. & Rozelle, R.M. (1975). Nonverbal expression as a function of crowding during a simulated police-citizen encounter. *Journal of Personality and Social Psychology*, 32, 40-54.

Baxter, J.S. (1990). The suggestibility of child witnesses: a review. *Applied Cognitive Psychology*, 3, 393-407.

Baxter, J.S. (2004). Managing interrogative pressure in police interviews. *The Police Journal*, 77, 303-308.

Baxter, J.S. & Bain, S.A. (2002). Faking interrogative suggestibility: the truth machine. *Legal and Criminological Psychology*, 7, 219-225.

Baxter, J.S., Bain, S.A. & McAusland, J.H. (2007). "Are you thick?": A case study of confrontational aspects of a police interview with a suspect. *The Police Journal*, 80, 99-108.

Baxter, J.S. & Boon, J.C.W. (2000). Interrogative suggestibility: the importance of being earnest. *Personality and Individual Differences*, 28, 753-762.

Baxter, J.S., Boon, J.C.W. & Marley, C. (2006). Interrogative pressure and responses to minimally leading questions. *Personality and Individual Differences*, 40, 87-98.

Baxter, J.S., Jackson, M. & Bain, S.A. (2003). Interrogative suggestibility: interactions between interviewees' self-esteem and interviewer style. *Personality and Individual Differences*, 35, 1285-1292.

Beail, N. (2002). Interrogative suggestibility, memory and intellectual disability. *Journal of Applied Research in Intellectual Disabilities*, 15, 129-137.

Bernard, L.C., Hutchison, S., Lavin, A. & Pennington, P. (1966). Ego-strength, hardiness, self-esteem, self-efficacy, optimism, and maladjustment: Health-related personality constructs and the "Big Five" model of personality. *Assessment*, 3, 115-131.

Bexton, W.H., Heron, W. & Scott, T.H. (1954). Effects of decreased variation in the sensory environment. *Canadian Journal of Psychology*, 8, 70-76.

Bickman, L. (1974). The social power of a uniform. *Journal of Applied Social Psychology*, 1, 47-61.

Bieling, P.J., Antony, M.M. & Swinson, R.P. (1998). The State-Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behaviour Research and Therapy*, 36, 777-788.

Binet, A. (1900). *La Suggestibilite*. Paris: Doin.

Binet, A. (1905). La science du temoignage. *Annee Psychologique*, 11, 128-136.

Biondo, J. & MacDonald, A.P. (1971). Internal-external locus of control and response to influence attempts. *Journal of Personality*, 39, 407-419.

Blagrove, M., Cole-Morgan, D. & Lambe, H. (1994). Interrogative suggestibility: the effects of sleep deprivation and relationship with field dependence. *Applied Cognitive Psychology*, 8, 169-179.

Blascovich, J. & Tomaka, J. (1991). Measures of Self-Esteem. In: J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds) *Measures of Personality and Social Psychological Attitudes* (3rd Edition). New York: Academic Press.

Bond, C.F. & Titus, L.J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin*, 94, 265-292.

Boon, J.C.W. & Baxter, J.S. (2000). Minimizing interrogative suggestibility. *Legal and Criminological Psychology*, 5, 273-284.

Boon, J.C.W. & Baxter, J.S. (2004). Minimizing extraneous, interviewer-based interrogative suggestibility: *Legal and Criminological Psychology*, 9, 229-238.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968—1987. *Psychological Bulletin*, 106, 265-289.

Brockner, J. (1979). The effects of self-esteem: Success-failure, and self-consciousness on task performance. *Journal of Personality and Social Psychology*, 37, 1732–1741.

Brown, J.D. & Dutton, K.A. (1995). The thrill of victory, the complexity of defeat: Self-esteem and people's emotional reactions to success and failure. *Journal of Personality and Social Psychology*, 68, 712-722.

Brown v. Mississippi, 297 U.S. 278 (1936).

Bruner, J.S. & Tagiuri, R. (1954). The perception of people. In: G. Lindzey (Ed) *Handbook of Social Psychology*. Reading, MA: Addison-Wesley.

Bull, R. (1999). Police investigative interviewing. In A. Memon & R. Bull (Eds) *Handbook of the Psychology of Interviewing*. Chichester: Wiley.

Bull, R. & Cherryman, J. (1995). *Helping to Identify Skill Gaps in Specialist Investigative Interviewing*. London: Home Office.

Bull, R. & Milne, B. (2004). Attempts to improve the police interviewing of suspects. In: G.D. Lassiter (Ed) *Interrogations, Confessions, and Entrapment*. Kluver/Plenum: New York.

Burley, P.M. & McGuiness, J. (1977). Effects of social intelligence on the Milgram paradigm. *Psychological Reports*, 40, 767-770.

Bushman, B. (1984). Perceived symbols of authority and their influence on compliance. *Journal of Applied Social Psychology*, 14, 501-508.

Cardone, D. & Dent, H. (1996). Memory and interrogative suggestibility: The effects of modality of information presentation and retrieval conditions upon the suggestibility scores of people with learning disabilities. *Legal and Criminological Psychology*, 1, 165-177.

Cartwright, D. (1965). Influence, leadership, control. In: J. March (Ed) *Handbook of Organizations*. Chicago: Rand McNally.

Carver, C.S. & Ganellen, R.J. (1983). Depression and components of self-punitiveness: High standards, self-criticism, and overgeneralisation. *Journal of Abnormal Psychology*, 92, 330-337.

Carver, C.S. & Scheier, M.F. (1981). *Attention and Self-Regulation: A Control Theory Approach to Human Behaviour*. New York: Springer.

Carver, C.S., Scheier, M.F. & Weintraub, J.K. (1989). Assessing coping strategies: A theoretically based approach. *Journal of Personality and Social Psychology*, 56, 267-283.

Ceci, S.J. & Bruck, M. (1995). *Jeopardy in the Courtroom: A Scientific Analysis of Children's Testimony*. Washington, DC: American Psychological Association.

Christensen, T.C., Wood, J.V. & Barrett, L.F. (2003). Remembering everyday experience through the prism of self-esteem. *Personality and Social Psychology Bulletin*, 29, 51-62.

Clare, I.C.H., Gudjonsson, G.H., Rutter, S.C. & Cross, P. (1994). The inter-rater reliability of the Gudjonsson Suggestibility Scale (Form 2). *British Journal of Clinical Psychology*, 33, 357-365.

Clarke, C. & Milne, R. (2001). *National Evaluation of the PEACE Investigative Interviewing Course*. Police Research and Award Scheme. London: Home Office.

Clifford, B.R. & Bull, R. (1978). *The Psychology of Person Identification*. London: Routledge & Kegan Paul.

Cohen, S. & McKay, G. (1984). Social support, stress and the buffering hypothesis: A theoretical analysis. In: A. Baum, J.E. Singer, & S.E. Taylor (Eds) *Handbook of Psychology and Health, Vol. 4*. Hillsdale, NJ: Erlbaum.

Cohen, S. & Wills, T.W. (1985). Stress, social support, and buffering. *Psychological Bulletin*, 98, 310-357.

Collins, R., Lincoln, R. & Frank, M.G. (2002). The effect of rapport in forensic interviewing. *Psychiatry, Psychology & Law*, 9, 69-78.

Cooke, P. & Davies, G. (2001). Achieving best evidence from witnesses with learning disabilities: New guidance. *British Journal of Learning Disabilities*, 29, 84-87.

Cooper, W.H. (1981). Ubiquitous halo. *Psychological Bulletin*, 90, 218-244.

Coopersmith, S. (1967). *The Antecedents of Self-Esteem*. San Francisco: W.H. Freeman.

Cottrell, N.B. (1972). Social facilitation. In: C. McClintock (Ed) *Experimental Social Psychology*. New York: Holt, Rinehart & Winston.

CPTU (1992a). *The Interviewer's Rule Book*. London: Home Office, The Central Planning and Training Unit.

CPTU (1992b). *A Guide to Interviewing*. London: Home Office, The Central Planning and Training Unit.

Crombag, H. F. M., Wagenaar, W. A. & van Koppen, P. J. (1996). Crashing memories and the problem of "source monitoring". *Applied Cognitive Psychology*, 10, 95–104.

Cronbach, L. J. (1955). Processes affecting scores on understanding of others and assumed similarity. *Psychological Bulletin*, 52, 177–193.

Custaud, M.A., Belin de Chantemele, E., Larina, I.M., Nichiporuk, I.A., Grigoriev, A., Duvareille, M. et al. (2004). Hormonal changes during long-term isolation. *European Journal of Applied Physiology*, 91, 508-515.

Cutler, B.L. & Penrod, S.D. (1995). *Mistaken Identification: The Eyewitness, Psychology and the Law*. New York: Cambridge University Press.

Davies, G. & Noon, E. (1991). *An Evaluation of the Live Link For Child Witnesses*. Home Office Report.

Delongis, A., Folkman, S. & Lazarus, R.S. (1988). The impact of daily stress on health and mood: psychological and social resources as mediators. *Journal of Personality and Social Psychology*, 54, 486-495.

Demo, D.H. (1985). The measurement of self-esteem: Refining our methods. *Journal of Personality and Social Psychology*, 48, 1490-1502.

DePaulo, B.M. & Pfeifer, R.L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology*, 16, 249-267.

Deutsch, M. & Gerard, H.B. (1955). A study of normative and informational social influences upon individual judgement. *Journal of Abnormal and Social Psychology*, 51, 629-636.

Driver, E.D. (1968). Confessions and the social psychology of coercion. *Harvard Law Review*, 82, 42-61.

Emler, J. (2001). *Self-Esteem: The Costs and Causes of Low Self-Worth*. York: York Publishing Services.

Evans, F.J. (1967). Suggestibility in the normal waking state. *Psychological Bulletin*, 67, 114-129.

Eysenck, H. (1947). *Dimensions of Personality*. London: Routledge and Kegan Paul.

Eysenck, H.J. & Furneaux, W.D. (1945). Primary and secondary suggestibility: An experimental and statistical study. *Journal of Experimental Psychology*, 35, 485-523.

Fazio, R.H. & Roskos-Ewoldsen, D.R. (1994). Acting as we feel: When and how attitudes guide behaviour. In: S. Shavitt & T.C. Brock (Eds) *Persuasion*. Boston: Allyn & Bacon.

Feeley, T.H. (2002). Comment on halo effects in rating and evaluation research. *Human Communication Research*, 28, 578-586.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Fiaschetti, M. (1930). *You Gotta Be Rough. The Adventures of Detective Fiaschetti of the Italian Squad*. New York: Doubleday, Doran.

Fisher, R.P. (1995). Interviewing victims and witnesses of crime. *Psychology, Public Policy & Law*, 1, 732-764.

Fisher, R.P. & Geiselman, R.E. (1988). Enhancing eyewitness memory with the cognitive interview. In: M.M. Gruneberg, P.E. Morris & R.N. Sykes (Eds) *Practical Aspects of Memory: Current Research and Issues. Volume 1: Memory in Everyday Life*. Chichester: John Wiley & Sons.

Fisher, R.P. & Geiselman, R.E. (1992). *Memory Enhancing Techniques For Investigative Interviewing: The Cognitive Interview*. Springfield, IL: Thomas.

Fisher, R.P., Geiselman, R.E. & Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of actual victims and witnesses of crime. *Journal of Applied Psychology*, 74, 722-727.

Fisher, R.P., Geiselman, R.E. & Raymond, D.S. (1987). Critical analysis of police interview techniques. *Journal of Police Science and Administration*, 15, 291-297.

Fisher, R., Geiselman, R.E., Raymond, D., Jurkevich, L. & Warhaftig, M. (1987). Enhancing enhanced eyewitness memory: Refining the cognitive interview. *Journal of Police Science and Administration*, 15, 291-296.

Folkman, S. & Lazarus, R.S. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behaviour*, 21, 219-239.

Geen, R.G. (1989). Alternative conceptions of social facilitation. In: P.B. Paulus (Ed.) *Psychology of Group Influence* (2nd edition). Hillsdale, NJ: Erlbaum.

Geiselman, R.E., Fisher, R.P., Cohen, G., Holland, H. & Surtes, L. (1986). Eyewitness responses to leading and misleading questions under the cognitive interview. *Journal of Police Science and Administration*, 14, 31-39.

Geiselman, R.E., Fisher, R.P., Firstenberg, I., Hutton, L.A., Sullivan, S., Avetissa, I. et al. (1984). Enhancement of eyewitness memory: An empirical evaluation of the cognitive interview. *Journal of Police Science and Administration*, 12, 74-80.

Geiselman, R.E., Fisher, R.P., Mackinnon, D.P. & Holland, H.L. (1985). Eyewitness memory enhancement in the police interview: Cognitive retrieval, mnemonics versus hypnosis. *Journal of Applied Psychology*, 70, 401-412.

George, R. & Clifford, B.R. (1992). Making the most of witnesses. *Policing*, 8, 185-198.

Gilbert, S.J. (1981). Another look at the Milgram obedience studies: The role of the graduated series of shocks. *Personality and Social Psychology Bulletin*, 7, 690-695.

Godden, D. & Baddeley, A.D. (1975). Context dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, 66, 325-331.

Gorassini, D.R., Harris, J.A., Diamond, A. & Flynn-Dastoor, E. (2006). Computer assessment of interrogative suggestibility. *Personality and Individual Differences*, 40, 569-577.

Gordon, N.J., Fleisher, W.L. & Weinberg, C.D. (2002). *Effective Interviewing and Interrogation Techniques*. San Diego, CA: Academic Press.

Grassian, S. & Friedman, N. (1986). Effects of sensory deprivation in psychiatric seclusion and solitary confinement. *International Journal of Law and Psychiatry*, 8, 49-65.

Greenberg, J., Solomon, S., Pyszczynski, T., Rosenblatt, A., Burling, J., Lyon, D. et al. (1992). Why do people need self-esteem? Converging evidence that self-esteem serves an anxiety-buffering function. *Journal of Personality and Social Psychology*, 63, 913-922.

Greenberger, E., Chen, C., Dmitrieva, J. & Farruggia, S.P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: Do they matter? *Personality and Individual Differences*, 35, 1241-1254.

Greene, E., Flynn, M.S. & Loftus, E.F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behaviour*, 21, 207-219.

Griffiths, A. & Milne, R. (2006). Will it all end in tiers? Police interviews with suspects in Britain. In: T. Williamson (Ed) *Investigative Interviewing: Rights, Research, Regulation*. Devon: Willan.

Griffiths, A. & Milne, R. (in prep). *The Griffiths Question Map (GQM). A Training Manual*.

Gudjonsson, G.H. (1983). Suggestibility, intelligence, memory recall and personality: An experimental study. *British Journal of Psychiatry*, 142, 35-37.

Gudjonsson, G.H. (1984). A new scale of interrogative suggestibility. *Personality and Individual Differences*, 5, 303-314.

Gudjonsson, G.H. (1987a). Historical background to suggestibility: How interrogative suggestibility differs from other types of suggestibility. *Personality and Individual Differences*, 8, 347-355.

Gudjonsson, G.H. (1987b). A parallel form of the Gudjonsson Suggestibility Scale. *British Journal of Clinical Psychology*, 26, 215-221.

Gudjonsson, G.H. (1988a). The relationship of intelligence and memory to interrogative suggestibility: the importance of range effects. *British Journal of Clinical Psychology*, 27, 185-187.

Gudjonsson, G.H. (1988b). Interrogative suggestibility: its relationship with assertiveness, social-evaluative anxiety, state anxiety and method of coping. *British Journal of Clinical Psychology*, 27, 159-166.

Gudjonsson, G.H. (1991). The application of interrogative suggestibility to police interviewing. In J.F. Schumaker (Ed.) *Human Suggestibility: Advances in Theory, Research, and Application*. London: Routledge.

Gudjonsson, G.H. (1997). *The Gudjonsson Suggestibility Scales Manual*. Hove: Psychology Press.

Gudjonsson, G.H. (2002). Suggestibility since Eysenck and Furneaux (1945): Theoretical and empirical aspects [Abstract]. *Proceedings of The British Psychological Society*, 10, 70.

Gudjonsson, G.H. (2003). *The Psychology of Interrogations and Confessions: A Handbook*. Chichester: Wiley.

Gudjonsson, G.H. (2006). The psychology of interrogations and confessions. In: T. Williamson (Ed) *Investigative Interviewing: Rights, Research, Regulation*. Devon: Willan.

Gudjonsson, G.H. & Bownes, I. (1992). The reasons why suspects confess during custodial interrogation: Data for Northern Ireland. *Medicine, Science and Law*, 32, 204-213.

Gudjonsson, G.H. & Clark, N.K. (1986). Suggestibility in police interrogation: a social psychological model. *Social Behaviour*, 1, 83-104.

Gudjonsson, G.H. & Gunn, J.G. (1982). The competence and reliability of a witness in a criminal court: A case report. *British Journal of Psychiatry*, 141, 624-627.

Gudjonsson, G.H. & Lister, S. (1984). Interrogative suggestibility and its relationship with perceptions of self-concept and control. *Journal of the Forensic Science Society*, 24, 99-110.

Gudjonsson, G.H. & MacKeith, J.A.C. (1988). Retracted confessions: Legal, psychological and psychiatric aspects. *Medicine, Science and the Law*, 28, 187-194.

Gudjonsson, G.H. & Petursson, H. (1991). Custodial interrogation: Why do suspects confess and how does it relate to their crime, attitude and personality. *Personality and Individual Differences*, 12, 295-306.

Gudjonsson, G.H., Rutter, S.C. & Clare, I.C.H. (1995). The relationship between suggestibility and anxiety among suspects detained at police stations. *Psychological Medicine*, 25, 875-878.

Gudjonsson, G.H. & Sigurdsson, J.F. (1999). The Gudjonsson Confession Questionnaire – Revised (GCQ-R): Factor structure and its relationship with personality. *Personality and Individual Differences*, 27, 953-968.

Gudjonsson, G.H. & Singh, K.K. (1984a). The relationship between criminal conviction and interrogative suggestibility among delinquent boys. *Journal of Adolescence*, 7, 29-34.

Gudjonsson, G.H. & Singh, K.K. (1984b). Interrogative suggestibility and delinquent boys: an empirical validation study. *Personality and Individual Differences*, 5, 425-430.

Guerin, B. (1993). *Social Facilitation*. Cambridge: Cambridge University Press.

Hall, E.T. (1966). *The Hidden Dimension*. New York: Doubleday.

Haller, M.H. (1976). Historical roots of police behaviour: Chicago, 1890-1925. *Law and Society Review*, 10, 303-323.

Haney, C., Banks, C. & Zimbardo, P.G. (1973). Interpersonal dynamics in a simulated prison. *International Journal of Criminology and Penology*, 1, 69-97.

Hansdottir, I., Thorsteinsson, H.S., Kristinsdottir, H. & Ragnarsson, R.S. (1990). The effects of instructions and anxiety on interrogative suggestibility. *Personality and Individual Differences*, 11, 85-87.

Haraldsson, E. (1985). Interrogative suggestibility and its relationship with personality, perceptual defensiveness and extraordinary beliefs. *Personality and Individual Differences*, 5, 765-767.

Harmon-Jones, E. & Allen, J.J.B. (2001). The role of affect in the mere exposure effect: Evidence from psychophysiological and individual differences approaches. *Personality and Social Psychology Bulletin*, 27, 889-898.

Harter, S. (1993). Causes and consequences of low self-esteem in children and adults. In: R. Baumeister (Ed) *Self-Esteem: The Puzzle of Low Self-Regard*. New York: Plenum.

Heatherton, T. & Ambady, N. (1993). Self-esteem, self-prediction, and living up to commitments. In: R. Baumeister (Ed) *Self-Esteem: The Puzzle of Low Self-Regard*. New York: Plenum.

Heaton-Armstrong, A. (1995). Recording and disclosing statements by witnesses: law and practice. *Medicine, Science and the Law*, 35, 136-143.

Hess, J.E. (1997). *Interviewing and Interrogation for Law Enforcement*. Cincinnati, OH: Anderson.

Holmberg, U. & Christianson, S.A. (2002). Murderers' and sexual offenders' experiences of police interviews and their inclination to admit or deny crimes. *Behavioural Sciences and the Law*, 20, 31-45.

Home Office (1985a). *Police and Criminal Evidence Act 1984*. London: HMSO.

Home Office (1985b). *Police and Criminal Evidence Act 1984 (S.66), Codes of Practice*. London: HMSO.

Hopkins, E.J. (1931). *Our Lawless Police: A Study of the Unlawful Enforcement of the Law*. New York: Viking Press.

Howard, R. & Hong, N.S. (2002). Effects of coping style on interrogative suggestibility. *Personality and Individual Differences*, 33, 479-485.

Hull, C.L. (1933). *Hypnosis and Suggestibility*. New York: Appleton-Century-Crofts.

Inbau, F.E. (1942). *Lie Detection and Criminal Interrogation*. Baltimore: Williams and Wilkins.

Inbau, F.E. (1948). *Lie Detection and Criminal Interrogation*. (2nd edition). Baltimore: Williams and Wilkins.

Inbau, F.E. & Reid, J.E. (1953). *Lie Detection and Criminal Interrogation*. (3rd edition). Baltimore: Williams and Wilkins.

Inbau, F.E. & Reid, J.E. (1962). *Criminal Interrogations and Confessions*. Baltimore: Williams and Wilkins.

Inbau, F.E. & Reid, J.E. (1967). *Criminal Interrogations and Confessions*. (2nd edition). Baltimore: Williams and Wilkins.

Inbau, F.E., Reid, J.E. & Buckley, J.P. (1986). *Criminal Interrogations and Confessions*. (3rd edition). Baltimore: Williams and Wilkins.

Inbau, F.E., Reid, J.E., Buckley, J.P. & Jayne, B.C. (2001). *Criminal Interrogations and Confessions*. (4th edition). Gaithersberg: Aspen.

Irving, B. (1980). *Police Interrogation: A Case Study of Current Practice. Research Studies No. 2*. London: HMSO.

Irving, B. (1986). The interrogation process. In: J. Benyon & C. Bourn (Eds) *The Police: Powers, Procedures and Proprieties*. Oxford: Pergamon Press.

Irving, B. & Hilgendorf, L. (1980). *Police Interrogation: The Psychological Approach. Research Studies No. 1*. London: HMSO.

Jelicic, M., Smeets, T., Peters, M.J.V., Candel, I., Horselenberg, R. & Merckelbach, H. (2006). Assassination of a controversial politician: Remembering details from another non-existent film. *Applied Cognitive Psychology*, 20, 591-596.

Johnson, M.K., Hashtroudi, S. & Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3-28.

Judge, T.A., Erez, A., Bono, J.E. & Thoresen, C.J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a

common core construct? *Journal of Personality and Social Psychology*, 83, 693-710.

Kassin, S.M., Goldstein, C.C. & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behaviour*, 27, 187-203.

Kassin, S.M. & Gudjonsson, G.G. (2004). The psychology of confessions: a review of the literature and issues. *Psychological Science in the Public Interest*, 5, 33-67.

Kassin, S.M., Leo, R.A., Meissner, C.A., Richman, K.D., Colwell, L.H., Leach, A.M. et al. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behaviour*, 31, 381-400.

Kassin, S.M. & McNall, K. (1991). Police interrogations and confessions: communicating promises and threats by pragmatic implication. *Law and Human Behaviour*, 15, 233-251.

Kebbell, M.R. & Giles, D.C. (2000). Lawyers' questions and witness confidence: Some experimental influences of complicated lawyers' questions on witness confidence and accuracy. *The Journal of Psychology*, 134, 129-139.

Kebbell, M.R., Hurren, E. & Mazerolle, P. (2006). *An investigation into the effective and ethical interviewing of suspected sex offenders*. Criminological Research Council and Crime and Misconduct Commission End of Grant Report, Brisbane, Australia: Griffith University.

Kebbell, M.R., Hurren, E.J. & Roberts, S. (2006). Mock-suspects' decisions to confess: The accuracy of eyewitness evidence is critical. *Applied Cognitive Psychology*, 20, 477-486.

Kebbell, M.R., Milne, R. & Wagstaff, G. (1999). The cognitive interview: A survey of its forensic effectiveness. *Psychology, Crime and Law*, 5, 101-116.

Keedy, E.R. (1937). The third degree and legal interrogation of suspects. *University of Pennsylvania Law Review*, 85, 761-777.

Kelley, H.H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality*, 18, 431-439.

Kelly, W.E. (2004). Examining the relationship between worry and trait anxiety. *College Student Journal*, 38, 370-373.

Kenny, D.A. & Berman, J.S. (1980). Statistical approaches to the correction of correlational bias. *Psychological Bulletin*, 88, 288-295.

Kernis, M.H., Frankel, B.S. & Brockner, J. (1989). Self-esteem and reactions to failure: The mediating role of overgeneralization. *Journal of Personality and Social Psychology*, 57, 707-714.

Kidd, W.R. (1940). *Police Interrogation*. New York: Basuino.

Kilham, W. & Mann, L. (1974). Level of destructive obedience as a function of transmitter and executant roles in the Milgram obedience paradigm. *Journal of Personality and Social Psychology*, 29, 696-702.

Kimmel, A.J. (1996). *Ethical Issues in Behavioural Research*. Cambridge, MA: Blackwell.

Koltuv, B. B. (1962). Some characteristics of intrajudge trait intercorrelations. *Psychological Monographs*, 76 (Whole no. 552).

Kubzansky, P.E. (1961). The effects of reduced environmental stimulation on human behaviour: A review. In: A.D. Biderman & H. Zimmer (Eds) *The Manipulation of Human Behaviour*. New York: Wiley.

Lane, J., Lane, A.M. & Kyprianou, A. (2004). Self-efficacy, self-esteem and their impact on academic performance. *Social Behaviour and Personality*, 32, 247-256.

Lane, S.M., Mather, M., Villa, D. & Morita, S.K. (2001). How events are reviewed matters: Effects of varied focus on eyewitness suggestibility. *Memory and Cognition*, 29, 940-947.

Leo, R.A. (1992). From coercion to deception: the changing nature of police interrogation in America. *Crime, Law and Social Change*, 18, 35-39.

Leo, R.A. (1996a). Miranda's revenge: Police interrogation as a confidence game. *Law and Society Review*, 30, 259-288.

Leo, R.A. (1996b). Questioning the relevance of Miranda in the Twenty-First Century. *Michigan Law Review*, 99, 1000-1029.

Leo, R.A. (1996c). Inside the interrogation room. *Journal of Criminal Law and Criminology*, 86, 266-303.

Leo, R.A. (2004). The third degree and the origins of psychological interrogation in the United States. In: G.D. Lassiter (Ed) *Interrogations, Confessions, and Entrapment*. Kluver/Plenum: New York.

Leyens, J-P. (1983). *Sommes-nous tous de psychologues? Approche psychosociale de theories implicites de la personalite*. Brussels: Mardaga.

Linton, C.P. & Sheehan, P.W. (1994). The relationship between interrogative suggestibility and susceptibility to hypnosis. *Australian Journal of Clinical and Experimental Hypnosis*, 22, 53-64.

Loftus, E.F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7, 560-572.

Loftus, E.F. (1979). Reactions to blatantly contradictory information. *Memory and Cognition*, 7, 368-374.

Loftus, E.F. & Loftus, G.R. (1980). On the permanence of stored information in the human brain. *American Psychologist*, 35, 409-420.

Loftus, E.F. & Palmer, J.C. (1974). Reconstruction of an automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behaviour*, 13, 585-589.

McGurk, B.J., Carr, M.J. & McGurk, D. (1993). *Investigative Interviewing Courses for Police Officers: An Evaluation. Police Research Series Paper 4*. London: Home Office.

McKenzie, I.K. (1994). Regulating custodial interviews: A comparative study. *International Journal of the Sociology of Law*, 22, 239-259.

McKenzie, I.K. (2001). Forensic investigative interviewing. In J.F. Gubrium & J.A. Holstein (Eds) *Handbook of Interview Research: Context and Method*. London: Sage.

McLean, M. (1995). Quality investigation? Police interviewing of witnesses. *Medicine, Science and the Law*, 35, 116-122.

Medford, S., Gudjonsson, G.H. & Pearse, J. (2003). The efficacy of the appropriate adult safeguard during police interviewing. *Legal and Criminological Psychology*, 8, 253-266.

Meeus, W.H.J. & Raaijmakers, Q.A.W. (1986). Administrative obedience: Carrying out orders to use psychological-administrative violence. *European Journal of Social Psychology*, 16, 311-324.

Meissner, C.A. & Kassin, S.M. (2004). "You're guilty, so just confess!" Cognitive and behavioural confirmation biases in the interrogation room. In: G.D. Lassiter (Ed) *Interrogations, Confessions, and Entrapment*. Kluver/Plenum: New York.

Meissner, C.A. & Russano, M.B. (2003). The psychology of interrogations and false confessions: Research and recommendations. *The Canadian Journal of Police and Security Services*, 1, 53-64.

Memon, A., Hope, L. & Bull, R. (2003). Exposure duration: effects on eyewitness accuracy and confidence. *British Journal of Psychology*, 94, 339-354.

Memon, A., Vrij, A. & Bull, R. (2003). *Psychology and Law: Truthfulness, Accuracy and Credibility* (2nd edition). Chichester: Wiley.

Memon, A. & Wright, D. (2000). Eyewitness testimony: theoretical and practical issues. In: J. McGuire, T. Mason & A. O'Kane (Eds) *Behaviour, Crime and Legal Process*. Chicester: Wiley.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 67, 371-378.

Milgram, S. (1974). *Obedience to Authority*. New York: Harper.

Milne, R. & Bull, R. (1999). *Investigative Interviewing: Psychology and Practice*. Chichester: Wiley.

Milne, R. & Bull, R. (2003). Interviewing by the police. In D. Carson & R. Bull (Eds) *Handbook of Psychology in Legal Contexts (2nd edition)*. Chichester: Wiley.

Milne, R. & Shaw, G. (1999). Obtaining witness statements: The psychology, best practice and proposals for innovation. *Medicine, Science and the Law*, 39, 127-138.

Minichiello, V., Aroni, R., Timewell, E. & Alexander, L. (1990). *In-Depth Interviewing: Researching People*. Melbourne: Longman Cheshire.

Miranda v. Arizona, 384 U.S. 336 (1966).

Mita, T.H., Dermer, M. & Knight, J. (1977). Reversed facial images and the mere-exposure hypotheses. *Journal of Personality and Social Psychology*, 35, 597-601.

Moreland, R.L. & Beach, S.R. (1992). Exposure effects in the classroom: The development of affinity among students. *Journal of Experimental Social Psychology*, 28, 255-276.

Mortimer, A. & Shepherd, E. (1999). Frames of mind: Schemata guiding cognition and conduct in the interviewing of suspected offenders. In: A. Memon & R. Bull (Eds) *Handbook of the Psychology of Interviewing*. Chichester: Wiley.

Moston, S. (1992). Can police officers distinguish between deceptive and truthful statements? *Policing*, 8, 26-39.

Moston, S. & Engelberg, T.M. (1993). Police questioning techniques in tape recorded interviews with criminal suspects. *Policing and Society*, 3, 223-237.

Moston, S. & Stephenson, G.M. (1993). The changing face of police interrogation. *Journal of Community and Applied Social Psychology*, 3, 101-115.

Moston, S., Stephenson, G.M. & Williamson, T.M. (1992). The effects of case characteristics on suspect behaviour during police questioning. *British Journal of Criminology*, 32, 23-40.

Mulbar, H. (1951). *Interrogation*. Springfield, Illinois: Thomas.

Munsterberg, H. (1908). *On the Witness Stand*. Garden City, NY: Doubleday. Retrieved February 23, 2007, from http://psychclassics.yorku.ca/Munster/Witness/detection.htm.

Newberry, J.J. & Stubbs, C.A. (1990). *Advanced Interviewing Techniques*. Glynco, Georgia: Bureau of Alcohol Tobacco and Firearms National Academy.

Newcomb, T. (1931). An experiment designed to test the validity of a rating technique. *Journal of Educational Psychology*, 22, 279-289.

Newell, B.R. & Shanks, D.R. (2007). Recognising what you like: Examining the relation between the mere-exposure effect and recognition. *European Journal of Cognitive Psychology*, 19, 103-118.

Nisbett, R. E. & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35, 250–256.

O'Connor, D.M. (2001). *An investigation of the robustness of the Gudjonsson Suggestibility Scales*. Unpublished PhD Thesis, University of Strathclyde.

Ofshe, R.J. & Leo, R.A. (1997). The social psychology of police interrogation: the theory and classification of true and false confessions. *Studies in Law, Politics and Society*, 16, 189-251.

Ogloff, J.R.P. (1999). Law and human behaviour: Reflecting back and looking forward. *Law and Human Behavior*, 23, 1-7.

O'Hara, C.E. (1956). *Fundamentals of Criminal Investigation*. Springfield, Illinois: Thomas.

Olsen, L. & Wells, R. (1991). Cognitive interviewing and the victim/witness in crisis. *The Police Chief*, February, 28-29.

Oppenheim, A.N. (1992). *Questionnaire Design, Interviewing, and Attitude Measurement*. New York: Pinter Publications.

Ord, B. & Shaw, G. (1999). *Investigative Interviewing Explained*. Surrey: NPB, Benson Publications.

Orne, M.T. (1959). The nature of hypnosis: Artifact and essence. *Journal of Abnormal and Social Psychology*, 58, 277-299.

Orne, M.T. & Holland, C.C. (1968). On the ecological validity of laboratory deceptions. *International Journal of Psychiatry*, 6, 282-293.

Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. (1957). *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.

Payne, S. (1980). *The Art of Asking Questions*. Princeton, NJ: Princeton University Press.

Pearse, J. & Gudjonsson, G.H. (1996). Police interviewing techniques at two south London police stations. *Psychology, Crime and Law*, 3, 63-74.

Pearse, J. & Gudjonsson, G.H. (1999). Measuring influential police interviewing tactics: a factor analytic approach. *Legal and Criminological Psychology*, 4, 221-238.

Polczyk, R. (2005). Interrogative suggestibility: cross-cultural stability of psychometric and correlational properties of the Gudjonsson Suggestibility Scales. *Personality and Individual Differences*, 38, 177-186.

Pollard, R., Trowbridge, B., Slade, P.D., Streissguth, A.P., Laktonen, A. & Townes, B.D. (2004). Interrogative suggestibility in a US context: Some preliminary data on normal subjects. *Personality and Individual Differences*, 37, 1101-1108.

Powers, P.A., Andricks, J.L. & Loftus, E.F. (1979). Eyewitness account of females and males. *Journal of Applied Psychology*, 64, 339-347.

Raven, J.C., Court, J.H. & Raven, J. (1983). *Manual for Raven's Progressive Matrices and Vocabulary Scales (Section 4, Advanced Progressive Matrices)*. London: H.K. Lewis.

Redlich, A.D. (2004). Mental illness, police interrogations, and the potential for false confession. *Psychiatric Services*, 55, 19-21.

Register, P.A. & Kihlstrom, J.F. (1988). Hypnosis and interrogative suggestibility. *Personality and Individual Differences*, 9, 549-558.

Reik, T. (1959). *The Compulsion to Confess: On the Psychoanalysis of Crime and Punishment*. New York: John Wiley and Sons.

Richardson, S.A., Dohrenwend, B.S. & Klein, D. (1965). *Interviewing: Its Forms and Functions*. London: Basic Books.

Richardson, G. & Smith, P. (1993). The inter-rater reliability of the Gudjonsson Suggestibility Scale. *Personality and Individual Differences*, 14, 251-253.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Rubenstein, C. (1982). Psychology's fruit flies. *Psychology Today*, 16, 83-84.

Russano, M.B., Meissner, C.A., Narchet, F.M. & Kassin, S.M. (2005). Investigating true and false confessions within a novel experimental paradigm. *Psychological Science*, 16, 481-486.

Sanders, G.S., Baron, R.S. & Moore, D.L. (1978). Distraction and social comparison as mediators of social facilitation. *Journal of Experimental Social Psychology*, 14, 291-303.

Schalon, C. (1968). Effect of self-esteem upon performance following failure stress. *Journal of Consulting and Clinical Psychology*, 32, 497.

Schifter, D.B. & Ajzen, I. (1985). Intention, perceived control, and weight loss: An application of the theory of planned behaviour. *Journal of Personality and Social Psychology*, 49, 843-851.

Schooler, J.W. & Loftus, E.F. (1986). Individual differences and experimentation: complementary approaches to interrogative suggestibility. *Social Behaviour*, 1, 105-112.

Schuman, H. & Presser, S. (1979). The open and closed question. *American Sociological Review*, 44, 692-712.

Scullin, M.H. & Ceci, S.J. (2001). A suggestibility scale for children. *Personality and Individual Differences*, 5, 203-209.

Seamon, J.G., Brody, N. & Kauf, D.M. (1983). Affective discrimination of stimuli that are not recognized: II. Effect of delay between study and test. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 9, 544-555.

Sells, S.B. (1973). The taxonomy of man in enclosed space. In: J.E. Rasmussen (Ed) *Man in Isolation and Confinement*. Illinois: Aldine.

Shanab, M.E. & Yahya, K.A. (1978). A cross-cultural study of obedience. *Bulletin of the Psychonomic Society*, 11, 267-269.

Shaver, P.R. & Brennan, K.A. (1990). Measures of depression and loneliness. In: J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds) *Measures of Personality and Social Psychological Attitudes* (3rd Edition). New York: Academic Press.

Shepherd, E. (1988). Developing interviewing skills: A career span perspective. In: P. Southgate (Ed) *New Directions in Police Training*. London: HMSO.

Shepherd, E. (1991). Ethical interviewing. *Policing*, 7, 42-60.

Shepherd, E. (1993). Resistance in interviews: The contribution of police perceptions and behaviour. In: E. Shepherd (Ed) *Aspects of Police Interviewing, Issues in Criminological and Legal Psychology, No. 18*. Leicester: British Psychological Society.

Shepherd, E. & Milne, R. (1999). Full and faithful: Ensuring quality practice and integrity of outcome in witness interviews. In: A. Heaton-Armstrong, E. Shepherd & D. Wolchover (Eds) *Analysing Witness Testimony: A Guide for Legal Practitioners and Other Professionals*. London: Blackstone Press.

Shrauger, J. & Rosenberg, S. (1970). Self-esteem and the effects of success and failure feedback on performance. *Journal of Personality*, 33, 404-414.

Sigurdsson, J.F. & Gudjonsson, G.H. (1994). Alcohol and drug intoxication during police interrogation and the reasons why suspects confess to the police. *Addiction*, 89, 985-997.

Singh, K.K. & Gudjonsson, G.H. (1984). Interrogative suggestibility, delayed memory and self-concept. *Personality and Individual Differences*, 5, 203-209.

Singh, K.K. & Gudjonsson, G.H. (1992). Interrogative suggestibility among adolescent boys and its relationship with intelligence, memory, and cognitive set. *Journal of Adolescence*, 15, 155-161.

Small, M.A. (1993). Legal psychology and therapeutic jurisprudence. *Saint Louis University Law Journal*, 37, 675-700.

Smeets, T., Leppink, J., Jelicic, M. & Merckelbach, H. (in press). Shortened versions of the Gudjonsson Suggestibility Scale meet the standards. *Legal and Criminological Psychology*.

Smith, E. (1930). The fallibility of eyewitness testimony. *The American Journal of Police Science*, 1, 487-495.

Smith, P. & Gudjonsson, G.H. (1995). Confabulation among forensic inpatients and its relationship with memory, suggestibility, compliance, anxiety, and self-esteem. *Personality and Individual Differences*, 19, 517-523.

Snyder, M. (1974). Self-monitoring and expressive behaviour. *Journal of Personality and Social Psychology*, 30, 526-537.

Snyder, M. (1987). *Public Appearances as Private Realities: The Psychology of Self-Monitoring*. New York: Freeman.

Softley, P. (1980). *Police Interrogation: An Observational Study in Four Police Stations. Home Office Research Study No.61*. London: HMSO.

Sommer, R. (1969). *Personal Space*. Englewood Cliffs, NJ: Prentice Hall.

Spielberger, C.D. (1969). *The State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C.D. (1972). *Anxiety: Current Trends in Research*. London: Academic Press.

Spielberger, C.D. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C.D., Gorsuch, R.L. & Lushene, R.E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Stanley, K.D. & Murphy, M.R. (1997). A comparison of general self-efficacy with self-esteem. *Genetic, Social and General Psychology Monographs*, 123, 79-99.

Stern, W. (1910). Abstracts of lectures on the psychology of testimony and on the study of individuality. *American Journal of Psychology*, 21, 273-282.

Stern, W. (1938). *General Psychology: From the Personalistic Standpoint*. New York: Macmillan.

Stern, W. (1939). The psychology of testimony. *Journal of Abnormal and Social Psychology*, 34, 3-20.

Stukat, K.G. (1958). *Suggestibility: A Factor and Experimental Analysis*. Stockholm: Almgvist & Wiksell.

St-Yves, M. (2006). The psychology of rapport: Five basic rules. In: T. Williamson (Ed) *Investigative Interviewing: Rights, Research, Regulation*. Devon: Willan.

Suls, J.M. & Wallston, K.A. (2003). *Social Psychological Foundations of Health and Illness*. New York: Blackwell.

Szpunar, K.K., Schellenberg, G. & Plinner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30, 370-381.

Tafarodi, R.W., Marshall, T.C. & Milne, A.B. (2003). Self-esteem and memory. *Journal of Personality and Social Psychology*, 84, 29-45.

Tafarodi, R.W. & Milne, A.B. (2002). Decomposing global self-esteem. *Journal of Personality*, 70, 443-483.

Tata, P.R. & Gudjonsson, G.H. (1990). The effects of mood and verbal feedback on interrogative suggestibility. *Personality and Individual Differences*, 11, 1079-1085.

Technical Working Group for Eyewitness Evidence (1999). *Eyewitness Evidence: A Guide for Law Enforcement [Booklet]* United States Department of Justice, Officer of Justice Programs: Washington DC.

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, 4, 469-477.

Tousignant, D.D. (1991). Why suspects confess. *FBI Law Enforcement Bulletin*, March, 14-18.

Tousignant, J.P., Hall, D. & Loftus, E.F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory and Cognition*, 14, 329-338.

Trippi, J.A. (1973). Special class placements and suggestibility of mentally retarded children. *American Journal of Mental Deficiency*, 79, 220-222.

Trovillo, P.V. (1940). [Review of the book Police Interrogation]. *Journal of Criminal Law and Criminology*, 31, 252-254.

Tuckey, M.R. & Brewer, N. (2003). The influence of schemas, stimulus ambiguity and interview schedule on eyewitness memory over time. *Journal of Experimental Psychology: Applied*, 9, 101-118.

Tulving, E. (1972). Episodic and semantic memory. In: E. Tulving & W. Donaldson (Eds) *Organisation of Memory*. London: Academic Press.

Tulving, E. (1989). Remembering and knowing the past. *American Scientist*, 77, 361-367.

Tulving, E. & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352-373.

Vinten, G. (1995). Open versus closed questions – an open issue. *Management Decision*, 33, 27-31.

Vonk, R. & Heiser, W.J. (1991). Implicit personality theory and social judgement: Effects of familiarity with a target person. *Multivariate Behavioral Research*, 26, 69-81.

Vrij, A. (2000). *Detecting Lies and Deceit: The Psychology of Lying and the Implications for Professional Practice*. Chichester: Wiley.

Vrij, A. & Graham, S. (1997). Individual differences between liars and the ability to detect lies. *Expert Evidence: The International Digest of Human Behaviour, Science and Law*, 5, 144-148.

Vrij, A. & Mann, S. (2001). Who killed my relative? Police officers' ability to detect real-life high stake lies. *Psychology, Crime and Law*, 7, 119-132.

Walkley, J. (1987). *Police Interrogation. A Handbook for Investigators*. London: Police Review Publication

Warrington, E.K. (1986). Memory for facts and memory for events. *British Journal of Clinical Psychology*, 25, 1-12.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised: Manual*. The Psychological Corporation, New York.

Weiner, I.B. (2003). *Handbook of Psychology*. New York: Wiley.

Weitzenhoffer, A.M. (1953). *Hypnotism: An Objective Study in Suggestibility*. New York: Wiley.

Weitzenhoffer, A.M. & Hilgard, E.R. (1962). *Stanford Hypnotic Susceptibility Scale: Form C*. Palo Alto, California: Consulting Psychologists Press.

Wells, F.L. (1907). A statistical study of literary merit. *Archives of Psychology*, 1.

Wells, L.E. & Marwell, G. (1976). *Self-Esteem: Its Conceptualization and Measurement*. Beverly Hills, CA: Sage.

Wickersham, G.W. (1931). *Report on Lawlessness in Law Enforcement: National Commission on Law Observance and Law Enforcement*. Washington, DC: United States Government Printing Office.

Williams, J. (2000). The inappropriate adult. *Journal of Social Welfare and Family Law*, 22, 43-57.

Williamson, T.W. (1993). From interrogation to investigative interviewing: strategic trends in police questioning. *Journal of Community and Applied Social Psychology*, 3, 89-99.

Williamson, T.W. (1994). Reflections on current police practice. In: D. Morgan & G. Stephenson (Eds) *Suspicion and Silence: The Right to Silence in Criminal Investigations*. London: Blackstone Press.

Wilson, W. & Miller, H. (1968). Repetition, order of presentation, and timing of arguments as determinants of opinion change. *Journal of Personality and Social Psychology*, 9, 184-188.

Woolston, R., Bain, S.A. & Baxter, J.S. (2006). Patterns of malingering and compliance in measures of interrogative suggestibility. *Personality and Individual Differences*, 40, 453-461.

Zachary, R.A. (1986). *Shipley Institute of Living Scale: Revised Manual*. Los Angeles: Western Psychological Services.

Zajonc, R.B. (1965). Social facilitation. *Science*, 149, 269-274.

Zajonc, R.B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monographs*, 9, 1-27.

Zimbardo, P.G. (1967). The psychology of police confessions. *Psychology Today*, 1, 17-27.

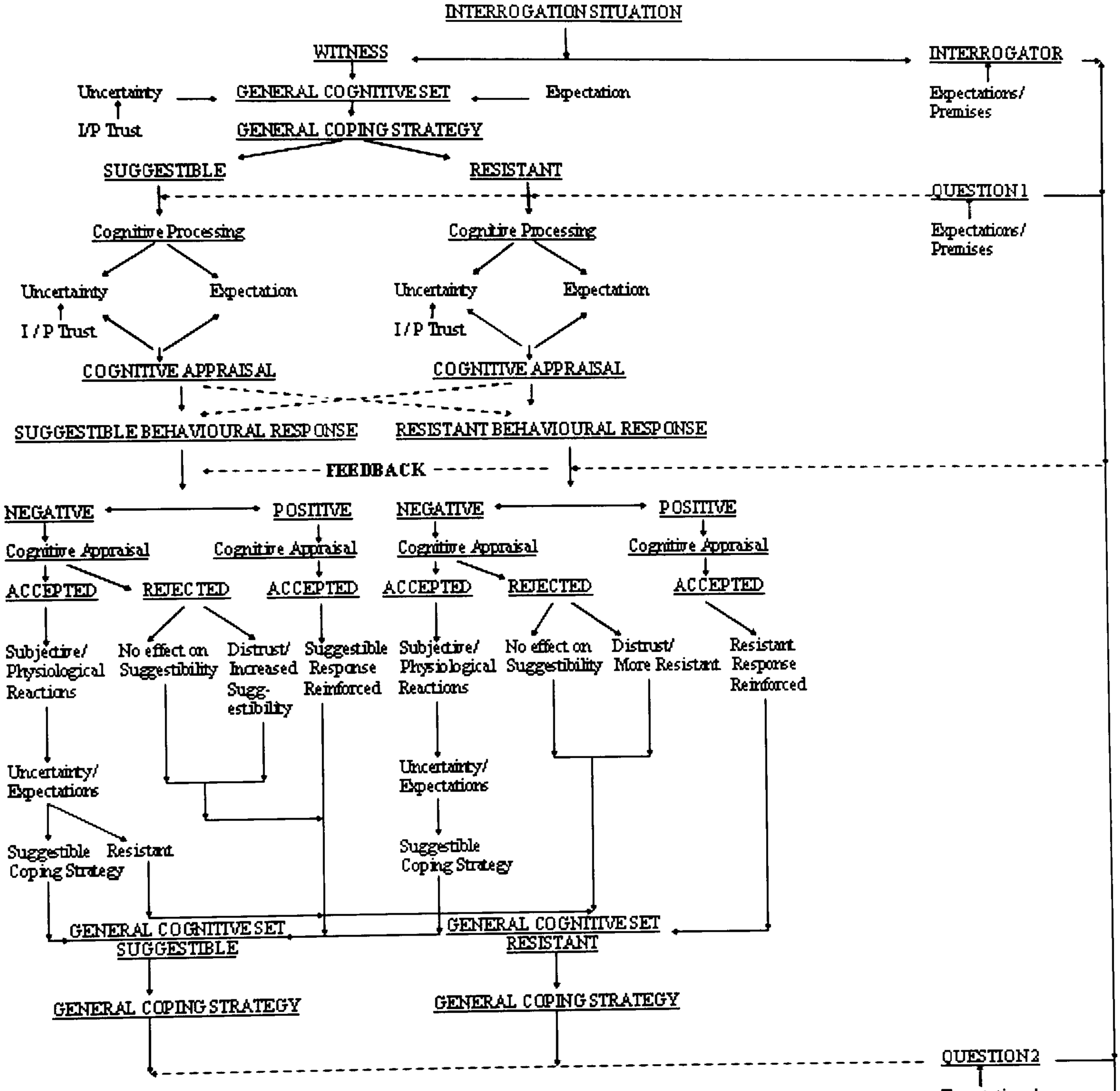
Zimbardo, P.G. (1968). Coercion and compliance: The psychology of police confessions. In: R. Perrucci & M. Pilisuk (Eds) *The Triple Revolution: Social Problems in Depth*. Boston: Little, Brown and Company.

Zubeck, J.P., Bayer, L. & Shepherd, J.M. (1969). Relative effects of prolonged social isolation and confinement: behavioural and physiological effects. *Journal of Abnormal Psychology*, 74, 625-631.

Zuckerman, M., Persky, H., Link, K.E. & Basu, G.K. (1968). Experimental and subject factors and defining responses to sensory deprivation, social isolation and confinement. *Journal of Abnormal Psychology*, 73, 183-194.

Zulawski, D.E. & Wicklander, D.E. (1993). *Practical Aspects of Interview and Interrogation*. Boca Raton: CRC Press.

# Appendix A – The Gudjonsson and Clark (1986) Model of Interrogative Suggestibility



# Appendix B – Standard GSS 1 and 2 Procedure and Measures (Gudjonsson 1984, 1987b)

| Procedure | Measure obtained |
|---|---|
| Presentation of spoken narrative | - |
| Interviewees asked to recall what they can remember about narrative | Immediate recall score<br>maximum score 40 (range 0-40)<br><br>Total confabulation 1 |
| 50 min delay | - |
| Interviewees asked again for free recall | Delayed recall score<br>maximum score 40 (range 0-40)<br><br>Total confabulation 2 |
| Interviewees asked 20 questions about the narrative (15 leading questions) | **Yield 1** – number of leading items accepted by interviewee<br>maximum score 15 (range 0-15) |
| Negative feedback administered to interviewee | - |
| Interviewees asked 20 questions again | **Shift** – number of responses which change following negative feedback; includes responses to 5 non-leading questions<br>maximum score 20 (range 0-20)<br><br>**Yield 2** - number of leading items accepted by interviewee<br>maximum score 15 (range 0-15)<br><br>Total suggestibility = Yield 1 + Shift<br>maximum score 35 (range 0-35) |

# Appendix C – Adapted GSS 2 Procedure and Measures (Baxter, Boon & Marley, 2006)

| **Procedure** | **Measure obtained** |
|---|---|
| Presentation of spoken narrative | - |
| Interviewees asked to recall what they can remember about narrative | Immediate recall score<br>maximum score 40 (range 0-40)<br><br>Total confabulation 1 |
| 50 min delay | - |
| Interviewees asked again for free recall | Delayed recall score<br>maximum score 40 (range 0-40)<br><br>Total confabulation 2 |
| Interviewees asked 20 questions about the narrative (all questions "minimally leading") | - |
| Negative feedback administered to interviewee | - |
| Interviewees asked 20 questions again | **Shift** – number of responses which change following negative feedback<br>maximum score 20 (range 0-20) |

# Appendix D – Variant of GSS Procedure and Measures

| Procedure | Measure obtained |
|---|---|
| Presentation of videotaped event | - |
| Interviewees asked to recall what they can remember about event | - |
| No delay | - |
| Interviewees asked up to 29 questions about the videotaped event (all questions "minimally leading") | Recall score<br>maximum score 80 (range 0-80) |
| Negative feedback administered to interviewee | - |
| Interviewees asked up to 29 questions again | **Response change** – number of responses which change following negative feedback<br>maximum score 29 (range 0-29)<br><br>**% Response change** – number of responses changed as percentage of no. of questions<br><br>**Response change towards inaccuracy** – number of responses which change towards inaccuracy following negative feedback<br>maximum score 29 (range 0-29)<br><br>RC toward inaccuracy as % of no. of questions<br>RC toward inaccuracy as % of total changes |

# Appendix E – Case 1: Example Answer Sheet

| | Questions | Recall | | Answers 1 | Answers 2 | RC | RC direction |
|---|---|---|---|---|---|---|---|
| | | Closed | Open | | | | |
| 1 | The second man to appear on the scene: What age was he? | | 8 | Late 20s, early 30s | 28 or so | N | |
| 2 | Was he wearing a jacket? | 1 | | Yes | No | Y | A |
| 3 | Did he have dark or light hair? | 1 | | Dark | Dark | | |
| 4 | How tall was he? | | 0 | 5'6 | 5'8 | N | |
| 5 | What did he say? | | 1 | Brother is mechanic | Brother is mechanic | | |
| 6 | What accent did he have? | 0 | | English | Scottish | Y | T |
| 7 | Could he have seen who took the case? | 1 | | No | No | | |
| 8 | Could he have taken the case? | 0 | | Yes | No | Y | T |
| 9 | Did he leave scene to the left or to the right? | 1 | | Left | Left | | |
| 10 | Did you see anyone walk past in the background? (How many?) | | 2 | Yes, 2 people | Yes, 2 people | | |
| 11 | The (first) by passer in the background: What age was he? | | 5 | 25 | 25 | | |
| 12 | Was he wearing a jacket? | 1 | | Yes | No | Y | A |
| 13 | Did he have short or long hair? | 0 | | Long | Long | | |
| 14 | Was he carrying anything? | 1 | | A rucksack | A rucksack | | |
| 15 | Did he say anything? | 1 | | No | No | | |
| 16 | Could he have taken the case? | 1 | | No | No | | |
| 17 | Did he leave scene to the left or to the right? | 1 | | Right | Left | Y | A |
| 18 | The second by passer in the background: What age was he? | | 2 | Early 50s | In his 50s | N | |
| 19 | Was he wearing a jacket? | 1 | | Yes | Yes | | |
| 20 | Did he have short or long hair? | 1 | | Short | Short | | |
| 21 | How tall was he? | | 0 | Don't know | Don't know | | |
| 22 | Did he say anything? | 1 | | No | No | | |
| 23 | Could he have taken the case? | 1 | | Yes | Don't know | Y | A |
| 24 | Did he leave scene to the left or to the right? | 1 | | Right | Right | | |
| 25 | Did you see the case? | 1 | | Yes, at first | Yes, at first | | |
| 26 | Did you see the case being removed? | 0 | | No | No | | |
| 27 | Where was the case when it was stolen? | 1 | | Side of car | Side of car | | |
| 28 | Could the woman have seen who took the case? | 1 | | No | Yes | Y | A |
| 29 | Who is most likely to have taken the case? | 0 | | 2nd man | 2nd man | | |

# Appendix F – Case 1: Scores

Scoring Key

Response change

Y = Answers 1 and 2 are sufficiently different to indicate response change

N = Answers 1 and 2 are *not* sufficiently different to indicate response change

Response change direction

A = Response changed away from accuracy

T = Response changed towards accuracy

X = Response changed in neither direction

Scores

| | | | |
|---|---|---|---|
| Recall – closed items | = | 17 | (max. 22) |
| Recall – open-ended items | = | 18 | (max. 58) |
| Response change | = | 7 | (max. 29) |
| Response change toward inaccuracy | = | 5 | (max. 29) |
| Number of questions answered | = | 29 | (max. 29) |

# Appendix G – Case 2: Example Answer Sheet

| | Questions | Recall | | Answers 1 | Answers 2 | RC | RC direction |
|---|---|---|---|---|---|---|---|
| | | Closed | Open | | | | |
| 1 | The second man to appear on the scene: What age was he? | | 5 | 20s | Late 20s | N | |
| 2 | Was he wearing a jacket? | 1 | | Yes | Yes | | |
| 3 | Did he have dark or light hair? | 1 | | Dark | Dark | | |
| 4 | How tall was he? | | 8 | 5'10 | 5'7 | Y | A |
| 5 | What did he say? | | 2 | Can I help; head gasket | Can I help; head gasket | | |
| 6 | What accent did he have? | 0 | | Scottish | Glasgow | N | |
| 7 | Could he have seen who took the case? | 1 | | No | No | | |
| 8 | Could he have taken the case? | 0 | | Yes | Yes | | |
| 9 | Did he leave scene to the left or to the right? | 0 | | Don't know | Don't know | | |
| 10 | Did you see anyone walk past in the background? (How many?) | | 1 | Yes, 1 person | Yes, 1 person | | |
| 11 | The (first) by passer in the background: What age was he? | | 3 | Early 20s | Mid 20s | Y | T |
| 12 | Was he wearing a jacket? | 1 | | Yes | Yes | | |
| 13 | Did he have short or long hair? | 1 | | Short | Short | | |
| 14 | Was he carrying anything? | 0 | | Don't know | No | Y | A |
| 15 | Did he say anything? | 0 | | Yes | No | Y | T |
| 16 | Could he have taken the case? | 1 | | No | Yes | Y | A |
| 17 | Did he leave scene to the left or to the right? | 1 | | Right | Right | | |
| 18 | The second by passer in the background: What age was he? | | | | | | |
| 19 | Was he wearing a jacket? | | | | | | |
| 20 | Did he have short or long hair? | | | | | | |
| 21 | How tall was he? | | | | | | |
| 22 | Did he say anything? | | | | | | |
| 23 | Could he have taken the case? | | | | | | |
| 24 | Did he leave scene to the left or to the right? | | | | | | |
| 25 | Did you see the case? | 0 | | No | No | | |
| 26 | Did you see the case being removed? | 0 | | No | No | | |
| 27 | Where was the case when it was stolen? | 0 | | Don't know | Don't know | | |
| 28 | Could the woman have seen who took the case? | 1 | | No | No | | |
| 29 | Who is most likely to have taken the case? | 0 | | 2nd man | Bypasser | Y | T |

# Appendix H – Case 2: Scores

Scoring Key

Response change

Y = Answers 1 and 2 are sufficiently different to indicate response change

N = Answers 1 and 2 are *not* sufficiently different to indicate response change

Response change direction

A = Response changed away from accuracy

T = Response changed towards accuracy

X = Response changed in neither direction

Scores

| | | | |
|---|---|---|---|
| Recall – closed items | = | 8 | (max. 22) |
| Recall – open-ended items | = | 19 | (max. 58) |
| Response change | = | 6 | (max. 29) |
| Response change toward inaccuracy | = | 3 | (max. 29) |
| Number of questions answered | = | 22 | (max. 29) |

# Appendix I – Interview Rating Form

Below are some questions which ask you about your experience of the interview. To answer each question, circle one of the numbers on the scale. A score of 1 means not at all, 3 is an average rating, and 5 means very. Please answer honestly and do not omit any items.

How easy did you find it to recall details of the crime scene during the questioning?

| not at all | | | | very |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

To what extent did you feel distracted during the questioning?

| not at all | | | | very |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

To what extent did you feel confident during the questioning?

| not at all | | | | very |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

To what extent did you feel influenced during the questioning?

| not at all | | | | very |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

To what extent did you feel comfortable during the questioning?

| not at all | | | | very |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

How easy was it for you to concentrate during the questioning?

| not at all | | | | very |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

Did you feel that the questions were reasonable?

| not at all | | | | very |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

To what extent did you feel sure of your answers during the questioning?

| not at all | | | | very |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

To what extent did you feel under pressure during the questioning?

| not at all | | | | very |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

Overall, how difficult did you find the interview?

| not at all | | | | very |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

# Appendix J – Interviewer Rater Form (Bain & Baxter, 2000)

Listed below are 18 words which can be used to describe a person's behaviour. Consider the behaviour of the person who interviewed you. For each word, circle somewhere on the 5-point scale which best represents your perception of the interviewer's manner. For example, if you thought the interviewer was not at all assertive, you would circle 1 opposite this word; or, if you perceived the interviewer as very assertive, you would circle 5, and so on. Please do not omit any items.

| | not at all | | | | very |
|---|---|---|---|---|---|
| Nervous | 1 | 2 | 3 | 4 | 5 |
| Severe | 1 | 2 | 3 | 4 | 5 |
| Friendly | 1 | 2 | 3 | 4 | 5 |
| Understanding | 1 | 2 | 3 | 4 | 5 |
| Assertive | 1 | 2 | 3 | 4 | 5 |
| Confident | 1 | 2 | 3 | 4 | 5 |
| Professional | 1 | 2 | 3 | 4 | 5 |
| Firm | 1 | 2 | 3 | 4 | 5 |
| Respectful | 1 | 2 | 3 | 4 | 5 |
| Positive | 1 | 2 | 3 | 4 | 5 |
| Formal | 1 | 2 | 3 | 4 | 5 |
| Warm | 1 | 2 | 3 | 4 | 5 |
| Stern | 1 | 2 | 3 | 4 | 5 |
| Organised | 1 | 2 | 3 | 4 | 5 |
| Effective | 1 | 2 | 3 | 4 | 5 |
| Authoritative | 1 | 2 | 3 | 4 | 5 |
| Competent | 1 | 2 | 3 | 4 | 5 |
| Negative | 1 | 2 | 3 | 4 | 5 |

# Appendix K – Rosenberg Self-Esteem Scale (Rosenberg, 1965)

BELOW IS A LIST OF STATEMENTS DEALING WITH YOUR GENERAL FEELINGS ABOUT YOURSELF. IF YOU **STRONGLY AGREE**, CIRCLE **SA**. IF YOU **AGREE** WITH THE STATEMENT, CIRCLE **A**. IF YOU **DISAGREE**, CIRCLE **D**. IF YOU **STRONGLY DISAGREE**, CIRCLE **SD**.

PLEASE ANSWER HONESTLY

| | | 1. STRONGLY AGREE | 2. AGREE | 3. DISAGREE | 4. STRONGLY DISAGREE |
|---|---|---|---|---|---|
| 1 | I feel that I`m a person of worth, at least on an equal plane with others. | SA | A | D | SD |
| 2 | I feel that I have a number of good qualities. | SA | A | D | SD |
| 3 | All in all, I am inclined to feel that I am a failure. | SA | A | D | SD |
| 4 | I am able to do things as well as most other people. | SA | A | D | SD |
| 5 | I feel I do not have much to be proud of. | SA | A | D | SD |
| 6 | I take a positive attitude toward myself. | SA | A | D | SD |
| 7 | On the whole, I am satisfied with myself. | SA | A | D | SD |
| 8 | I wish I could have more respect for myself. | SA | A | D | SD |
| 9 | I certainly feel useless at times. | SA | A | D | SD |
| 10 | At times I think I am no good at all. | SA | A | D | SD |

PLEASE CIRCLE AS APPROPRIATE. I AM:

Male Female

## Appendix L – Mean and standard deviation scores on the RSE for undergraduate students

| *n* | *Mean* | *S.D.* | *Range* |
|---|---|---|---|
| 308 | 29.78 | 4.63 | 14-40 |

## Appendix M – Percentile scores on the RSE for undergraduate students

| *Percentile points* | | | | | | |
|---|---|---|---|---|---|---|
| *5* | *10* | *25* | *50* | *75* | *90* | *95* |
| 22.00 | 24.00 | 27.00 | 30.00 | 33.00 | 36.00 | 38.00 |