An Individual Differences Analysis of Witness Response in Simulated Forensic Interviews: An Investigation of Younger and Older Adults

Hazel McMurtrie

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

School of Psychological Sciences and Health
University of Strathclyde
Glasgow
2011

Author's Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:

Date:

Acknowledgements

Dr James S. Baxter, Dr Marc C. Obonsawin and Dr Simon C. Hunter, University of Strathclyde, Glasgow

Dr Mary Thomson and Sinan Gonul of Glasgow Caledonian University for advice on certain aspects of statistical analysis.

To enable the filming of the café scene:-

The Manager and Staff of the University of Strathclyde, Lord Todd Bar for the use of the premises,

Sarah MacQuarrie, Allan McGroarty, Kate Hefferon and Trevor Agus for acting out the scene,

Norman Sharp for technical expertise.

Allan McGroarty for the use of crime event stimulus material.

Participants from the University of Strathclyde, School of Psychological Sciences and Health, and from the Centre of Lifelong Learning for their participation.

Paul for his encouragement, confidence and patience.

Table of Contents

List of Tables

List of Figures

General Abstract

The two studies presented here used an experimental procedure, and a mock witness paradigm to investigate individual differences in witness response within simulated forensic interviews. Experiment 1 adapted the general theoretical model developed by Koriat and Goldsmith (1996) from answering general knowledge questions to answering episodic memory queries for details of an event. The framework proposes that people strategically regulate the accuracy of their memory reports by using confidence judgments to guide what they report. Experiment 2 adapted the Gudjonsson Suggestibility Scale procedure (GSS; Gudjonsson, 1984, 1987a) to investigate the personality and situational determinants of participant response to interrogative pressure. The results of Experiment 1 revealed that older adults made greater gains in memory accuracy from forced to free report, such that age-related deficits were not observed. This result was contrary to the experimental hypothesis and previous results. Older adult accuracy was obtained at the expense of greater losses in quantity correct. A more conservative response criterion by older adults led to the volunteering of fewer correct (and incorrect) responses. As predicted, overall accuracy was significantly greater in response to cued-recall questions and when participants reported descriptive compared with action details, although an age effect resulted in observed interactions. The results of Experiment 2 revealed that irrespective of the type of feedback, beliefs about one's own memory were related to compliance and response change scores although contrary to expectations, this did not vary with age. Confirming expectations and previous results

negative feedback was associated with higher response change scores compared with neutral feedback. An age effect was not observed following negative feedback. An individual differences analysis was applied to both studies. The results are discussed from a theoretical and applied perspective.

Chapter 1: An Introduction to Eyewitness Testimony: Age, Interview Protocols and Empirical Research

1.1 Eyewitness Testimony

Witness information is very important to the forensic process. The interviewing of victims, witnesses and suspects forms an integral part of the police investigation into criminal activities (Williamson, 2007). During the initial investigation, detailed descriptions of an event i.e. what happened, who the perpetrators were and leads to further assist police with their investigation may be obtained. As the investigation proceeds an initial interview is often followed by additional interviews whereby witnesses can be asked to identify perpetrators, objects or places. In the final stages of bringing an offender to justice, witness evidence is central to most court cases (Kebbell & Milne, 1998; Sanders, 1996; Zander & Henderson, 1993). Witness testimony increases the likelihood that a perpetrator will not only be apprehended but also prosecuted (Lieppe, 1980; Visher, 1987). Furthermore, jurors rely heavily on witness accounts of what they have experienced (e.g., Cutler, Penrod, & Dexter, 1990). Consequently, obtaining as full and accurate an account of what a witness or victim has experienced at the outset, is vital (Dando & Milne, 2009). (Hereinafter the term witness will be used to describe both a bystander and a victim of crime).

Psychological research has shown that due to the constructive nature of memories, memory for an event is not perfect. The malleable nature of eyewitness memory during the interviewing process has been extensively demonstrated (cf. Loftus, 1975, 1989;

Malpass, 1996). A person who is asked repeatedly to recall a witnessed event tends to gain, lose, and change details over time (Turtle & Yuille, 1994). Consequently, incomplete and inaccurate witness information can result in serious negative outcomes (Savage & Milne, 2006). This is graphically illustrated by work on wrongful convictions (Cutler & Penrod, 1995; Innocence Project, 2010; Wells et al., 1998). Wells et al., (1998) provides an account of the false convictions established in American courts on the basis of genetic fingerprinting (DNA) testing; forty trials that led to wrongful conviction were available for review. In each of these the testing of genetic material with new techniques established beyond any doubt that the convictions were incorrect. These were serious miscarriages of justice. All of the men who were convicted served prison sentences. Five of them spent time on death row awaiting execution. Ninety per cent of these proven cases of wrongful conviction depended upon the testimony of at least one eyewitness. In one instance, as many as five eyewitnesses were involved (Wright & Davies, 1999). Given the potentially seriousness of eyewitness mistakes, one of the most challenging tasks facing the criminal justice system is the evaluation and determination of the credibility of victims, witnesses and suspects (Brewer & Weber, 2008; Gudjonsson, 2010).

1.1.1 Estimator and System Variables

A useful model to examine the factors affecting the accuracy of witness reports is the two category system of research proposed by Wells (1978), consisting of *estimator* and *system* variables. Estimator variables include factors inherent in the nature of the crime as well as the age, nature and abilities of the witness. The criminal justice system cannot control estimator variables however such knowledge could be used to estimate, *post hoc* the likely accuracy of a witness. Examples of estimator research include: own race and own group bias (Adams-Price, 1992; Brigham & Malpas, 1985; Brigham & Ready, 1985; Shaw & Skolnick, 1994; Stroud & Wright ,1997; Yarmey, 1993), own age bias (Havard & Memon, 2009; Wilcock, Bull, & Vrij, 2007; Wright & Stroud, 2002), the effects of post-event information (PEI: Axmacher, Gossen, Elger, & Fell, 2010; Hyman, Husband, & Billings, 1995; Loftus & Palmer, 1974; Loftus, Miller, & Burns, 1978; Loftus & Pickrell, 1995; Skagerberg & Wright, 2009; Wright & Stroud, 1998; Wright, Memon, Skagerberg, & Gabbert, 2009), and of particular interest to the present research, the relationship between witness confidence and accuracy (Brewer & Weber, 2008; Brewer, Weber, & Semmler, 2007; Odinot, Wolters, & van Koppen, 2009; Penrod & Cutler, 1995; Sporer, Penrod, Read, & Cutler, 1995; Wells & Murray, 1984).

System variables, in contrast, are under the control of the criminal justice system and include how a witness is questioned as well as the structure of and instructions given during a line-up or identification parade (Howitt, 2002). One interview technique used is the Cognitive Interview (CI) (Fisher & Geiselman, 1992; Fisher, Geiselman, Raymond, Jurkevich, & Warhaftig, 1987; Geiselman et al., 1984; Geiselman, Fisher, MacKinnon, & Holland, 1985) and its revision the Enhanced Cognitive Interview (ECI) (Fisher, Geiselman, & Raymond, 1987; Fisher & Geiselman, 1992). The CI and ECI are among the most rigorously tested and widely accepted methods for improving the accuracy and completeness of eyewitness reports (Dando & Milne, 2009). The psychological principles which underpin the CI and ECI, together with a review of their efficacy are described in full later in this opening Chapter.

Examples of both estimator and system variables were investigated within this thesis. Experiment 1 investigated participant use of confidence judgements to decide whether information is reported or withheld. The experiment is reported in Chapter 3. Experiment 2 investigated the personality and situational determinants of participant response to investigative pressure. The experiment is reported in Chapter 4. The witness's age and participant individual differences were also considered. These are next considered in this opening chapter.

1.2 The Aging Eyewitness

One factor that reliably affects eyewitness performance is the age of the witness (Wilcock, Bull, & Milne, 2008). Research has primarily focused on comparing the performance of children with young adults (Hayes & Delamothe, 1997; Holliday, 2003a; 2003b; McCauley & Fisher, 1995, 1996; Memon, Holley, Wark, Bull, & Köhnken, 1996; Milne & Bull, 2002; 2003). As the proportion of the UK population aged 65 and over is projected to rise to 23% by 2034 (National Statistics, 2009), the number of elderly eyewitnesses to crime may also grow. Consequently, older adults may be involved more frequently in the Criminal Justice System (Rothman, Dunlop, & Entzel, 2000). A potential consequence of age-related declines in memory is older adults' greater susceptibility to financial scams (Jacoby, 1999), and distraction burglaries (Home Office, 2008). Therefore, the analysis of older adult witness behaviour and of interviewing strategies which may best facilitate the recall of the elderly becomes more important. As a result, research interest in this area has grown (Wilcock, 2010).

1.2.1 Age and Memory

Cognitive functioning in many different areas gradually deteriorates with age across normal healthy adults. It is acknowledged that there are very likely large differences in cognitive deficits between healthy adults and those who experience dementia. However, this thesis is only concerned with the way in which normal adults differ from one another, and not how they differ from patient populations. Deficits to episodic memory are more pronounced (Park & Minear, 2004). Episodic memory refers to the system of long-term memory concerned with to-be-remembered events. Age-related deficits have been found across different materials such as word lists, sentences, fragments of prose, faces, drawings, and photographs (see Bäckman, Small, & Larsson, 2000, for a review), and with measures of recall (Botwinick, 1978; Burke & Light, 1981; Craik, 1977; Craik & MacDowd, 1987; Whiting & Smith, 1997) and recognition (Erber, 1974; Gordon & Clark, 1974; Isingrini, Hauer, & Fontaine, 1996; White & Cunningham, 1982). The dominant explanation for this decline is age-related changes in the brain (Prull, Gabrieli, & Bunge, 2000; Raz, 2000; West, 1996; Woodruff-Pak, 1997) and these changes compromise the effectiveness of neuronal functioning and hence cognitive processing (for a review see Zacks, Hasher, & Li, 2000). These changes include general shrinkage of neurons, reduced cerebral blood flow, and decreased availability of certain neurotransmitters (McDaniel, Einstein, & Jacoby, 2008)

A number of hypotheses deriving from the information-processing framework (encoding, storage, retrieval) have been proposed to account for such cognitive deficits (for a review see Luo & Craik, 2008). Specifically, there are noted decreases in the speed at which mental operations are performed (e.g., Salthouse, 1996), long-term memory functions less effectively (Kausler, 1991; Park, Smith, et al., 1996), and the amount of working memory capacity available declines (Craik, Morris, & Gick, 1990; Hasher & Zacks, 1979, 1984, 1988; Park, Smith, et al., 1996; Salthouse & Babcock, 1991). Hasher and Zacks (1988) suggest that age-related decline in working memory results from a decreased ability to inhibit irrelevant information (Hasher & Zacks, 1988; Lustig, Hasher, & Tonev, 2001; Zacks & Hasher, 1988). Other researchers argue that limited resources are available to process cognitive tasks (e.g., Hashtroudi, Johnson, & Chrosniak, 1990), with older adults having fewer available resources than younger adults. This makes it difficult for older adults to engage in cognitively demanding tasks such as elaborate encoding strategies (Hashtroudi et al., 1990). Similarly, the reduced cognitive control approach (e.g., Park & Hedden, 2001) suggests that decrements in both processing efficiency and working memory capacity cause age-related memory deficits (Park, Smith, Dudley, & Lafronza, 1989). According to these theories, older adults' recall is optimal when they have sufficient time to process cognitive tasks, when their attention is undivided and when activities are uncomplicated (Anderson, 1999). Such processing theories may also explain the observation that older adults are less proficient than young adults at remembering the temporal sequence of to-be-remembered events (e.g., Vakil, Weise, & Enbar, 1997). According to the associate deficit hypothesis

(Naveh-Benjamin, 2000) memory declines because older adults are less able to form connections between different items or features of an event (e.g., Chalfonte & Johnson, 1996; Henkel, Johnson, & De Leonardis, 1998; Koutstaal, Schacter, & Brenner, 2001; Kroll, Knight, Metcalfe, Wolf, & Tulving, 1996; Light, Patterson, Chung, & Healy, 2004; Mitchell, Johnson, Raye, & D'Esposito, 2000, Naveh-Benjamin, Hussain, Guez, & Bar-on, 2003; Schacter, Norman, & Koutstaal, 1998). Older adults are impaired at remembering the source of recently acquired information (cf. Balota, Dolan, & Duchek, 2000; Dodson & Schacter, 2002; Johnson, Hashtroudi, & Lindsay, 1993; Memon, Bartlett, Rose, & Gray, 2003; Schacter et al., 1998). Events are encoded in a less elaborate, more general way by older adults (Craik & Simon, 1980; Hasher & Zacks, 1979; Hashtroudi et al., 1990; Koutstaal, Schacter, Galluccio, & Stofer, 1999; Koutstaal et al., 2003; Puglisi, Park, Smith, & Dudley, 1988; Rabinowitz & Ackerman, 1982), but see also Park, Puglisi, Smith and Dudley (1987).

Finally, elderly participants are less efficient at encoding information in relation to scripts and schematic prior knowledge. This is especially the case where such information has little association with the script or schema and is therefore dependent upon further processing for integration in memory (Bäckman, 1991; Hess, Donley, & Vandermaas, 1989). Older adults may respond on the basis of familiarity, plausibility, or an easily generated alternative. As a result, memory distortion may occur (cf. Jacoby & Hollingshead, 1990; Reder, 1987; Reder, Wible, & Martin, 1986). The environmental support hypothesis (Craik, 1986, 1994) proposes that age-related reductions in attentional capacity make older adults less able to engage in self-initiated remembering processes (e.g., retrieval searches). According to this theory, older adults' memories are greatly influenced by the amount of contextual cues available from the environment (Hasher, Tonev, Lustig, & Zacks, 2001). When extensive contextual cues from the original event are present at retrieval, age-related recall deficits are reduced (Fernandez & Alonso, 2001; Sharps & Antonelli, 1997; Yarmey & Yarmey, 1997).

The distinction between recognition, and recall memory tasks is an important one. Recognition memory is an example of environmental support (Craik & Jennings, 1992), because a target item is re-presented. This is in contrast to free recall which involves more self-initiated activity and which is a more explicit source identification task. Age-related decrements have been found to be greater in recall compared with recognition (Craik & McDowd, 1987; Nyberg et al., 2003; Rabinowitz 1984, 1986) and greater in recall compared with cued-recall (Ceci & Tabor 1981; Craik, Byrd, & Swanson, 1987).

1.2.2 Individual Differences in Cognitive Aging

Researchers have begun to move beyond conceptualizing cognitive aging merely as a population-level phenomenon. There is a growing appreciation for person-to-person individual differences in the cognitive aging process. The shift in research interest is illustrated by the following two quotes.

"In some people cognition declines precipitously, but in many others cognition declines only slightly or not at all, or improves slightly. Determining the factors that contribute to this variabity is likely to require detailed knowledge about individual differences in patterns of change in different cognitive abilities in old age" (Wilson et al., 2002).

"Researchers are recognizing increasingly that the study of mean change with age does not give a full account of cognitive change across the life span. Although the average performance on most tasks may decline with age, studies have suggested that many older individuals may change very little whereas others deteriorate dramatically" (Christensen et al., 1999).

Within the current thesis an individual differences analysis across younger and older adults was used to assess the relationship between participant subjective memory beliefs, years of education, crystallised intelligence, health, sensory deficits, positive and negative affect, and objective memory performance. Work on the role of lifestyle or pharmaceuticals is not reviewed. The potential effects of education, measures of crystallised intelligence, health, sensory deficits, and affective states on memory performance will be addressed in subsequent paragraphs. Subjective memory beliefs fall within the construct of metamemory and will be introduced in the next section.

1.3 Moderators of Cognitive Aging

1.3.1 Metacognition and Metamemory

Metacognition is broadly defined as the knowledge or awareness of one's cognitive processes (Brown, 1978; Niemi, 2002; Shimamura, 2000). Metamemory is an aspect of metacognition and is a person's beliefs about his or her own memory and the memory of others. It also includes the person's ability to use such beliefs to regulate his or her own memory processes (Cavanaugh & Green, 1990; Flavell, 1979; Gilweski & Zelinski, 1986; Hertzog, Dixon, & Hultsch, 1990; Hultsch, Hertzog, Dixon, & Davidson, 1988;

Lovelace, 1990; Nelson & Narens, 1990). Three general categories of metamemory constructs include:-

1. *Declarative knowledge* about memory tasks and memory processes; i.e. knowledge about both how memory functions and the viability of strategic behaviours for tasks requiring memory processes,
2. *Memory monitoring*; i.e. awareness of the current state of one's memory system, and
3. *Self-referent beliefs* about memory; particularly memory self-efficacy; i.e. an individual's ability to use memory effectively in memory demanding situations.

Hultsch et al., (1988) identified a fourth aspect of metamemory; memory-related affect. This was defined as a variety of emotional states relating to or generated by memory-demanding situations, including anxiety, depression, and fatigue.

1.3.2 Measurement of Memory Beliefs

The measurement of subjective memory beliefs has mainly relied on self-report questionnaires. A literature review identified five different research methodologies:

(a) *metamemory questionnaires,* such as the Metamemory in Adulthood (MIA; Dixon, Hultch, & Hertzog, 1988) and the Memory Functioning Questionnaire (MFQ; Gilewski, Zelinski, Schaie, & Thompson, 1983),

(b) *memory complaints questionnaires,* such as the Inventory of Memory Experiences (IME; Herrmann & Neisser, 1978),

(c) *self-efficacy measures,* (Balcerak & Rebok, 1986; Berry, West, & Scogin, 1983; Berry, West, & Dennehy, 1989),

(d) *single, task-specific predictions,* whereby individuals predict the number of items that they expect to recall (Bruce, Coyne, & Botwinick, 1982; Coyne, 1985; Murphy, Sanders, Gabriesheski, & Schmitt, 1981), and

(e) *feeling of knowing or confidence ratings* administered after encoding or retrieval (Lachman & Lachman, 1980; Perlmutter, 1978).

These methods vary in terms of the timing and the type of prediction obtained: prospective (prior to encoding) or retrospective (e.g., a confidence judgement about a previous recall response), (Nelson & Narens, 1990). Three aspects of metamemory are of particular interest to the present research. Experiments 1 and 2 obtained participant self-reports of memory prior to memory encoding. Individual differences in memory beliefs were measured. Experiment 1 also investigated retrospective monitoring in the form of participant confidence judgments. The accuracy-confidence relation will be addressed in subsequent paragraphs, but next the chapter will turn to the measurement and indirect effect of subjective memory beliefs on objective memory performance, specifically self-referent beliefs of control and self-efficacy measures.

1.4 Prospective Monitoring

1.4.1 Social Context and Beliefs about Aging

Within cultures, social norms exist about the aging process, including ideas about individual ability at different points in the life span. Such norms may affect the manner in which society responds to individuals on the basis of their age, both in terms of interpersonal interactions and public policies relating to work and retirement. Consequently, development may be influenced by individuals internalizing age-relevant beliefs. This in turn may affect memory performance as perceptions of one's own ability and behaviours are influenced (Hess, 2005).

Research on memory-related beliefs in relation to aging has been conducted from various perspectives including the examination of *stereotypes*, *control beliefs*, and *memory self-efficacy*. Hertzog and Hultsch (2000) distinguish between two general categories of i*mplicit and self-referent beliefs*. Implicit beliefs are the informal ideas that individuals have about the nature of memory and its development course, as applied to most people. In contrast, self-referent beliefs reflect expectations about change in one's own ability, and the factors that influence performance and change (Hertzog & Hultsch, 2000).

1.4.2 Stereotypes

Within Western culture stereotypically, aging is associated with negative cognitive attributes, such as slow thinking, senile behaviour, and forgetfulness (e.g., Hummert, Garstka, Shaner, & Strahm, 1994). Negative aging stereotypes are present in adults of all ages (e.g., Heckhausen, Dixon, & Baltes, 1989; Hummert et al., 1994). However, older women are assigned more negative stereotypes at an earlier age than older men (Hummert, Garstka, & Shaner, 1997; Mueller-Johnson, Toglia, Sweeney, & Ceci, 2007). These negative stereotypes affect how older individuals are viewed and how others respond to them. For example, memory failures in older adults are more likely to be viewed as reflections of mental difficulty than are the same failures in younger adults (e.g., Erber & Rothberg, 1991; Erber, Szuchman, & Rothberg, 1990), resulting in more sympathetic reactions toward older than younger adults (Erber, Szuchman, & Prager, 1997). Additionally, cues associated with older adults' physical appearance and behaviour may activate aging stereotypes in others, which in turn influence behaviour toward these same individuals (e.g., patronising talk Kemper, 1994; Ryan & Cole, 1990; Ryan, Meredith, & Shantz, 1994). Older adults can respond to conditions that prime age stereotypes by performing more poorly on memory tasks (e.g., Cavanaugh, Feldman, & Hertzog, 1998; Chasteen, Bhattacharyya, Horhota, Tam, & Hasher, 2005; Hess, Auman, Colcombe, & Rahhal, 2003; Hess, Hinson, & Statham, 2004; Hess & Hinson, 2006; Levy, 1996; Rahhal, Hasher, & Colcombe, 2001; Stein, Blanchard-Fields, & Hertzog, 2002 – but see Horton, Baker, Pearce, & Deakin, 2010). For example Rahhal et al., (2001) observed significant age differences when participants were *repeatedly* told that the purpose of the study was to test their *memory,* when compared to memory-neutral instructions that focussed on the learning of facts and the acquisition of knowledge.

In a forensic setting, negative stereotypes surrounding older adults might have far-reaching consequences. Older adults, particularly older females may erroneously be treated as less credible witnesses (Mueller-Johnson et al., 2007).

1.4.3 Memory Self-efficacy and Control Beliefs

Memory self-efficacy is a dimension of metamemory that specifically reflects the appraisals of one's capability to attain a given type or level of performance in designated settings (Bandura, 1977, 1997). Memory self-efficacy allows for the possibility that an older individual may have extensive and accurate knowledge about how his / her memory functions (metamemory) but may also believe that his or her ability to remember in a given context is poor (Jorm, Christensen et al.,1994; Jorm, Christensen, et al., 1997). Some people believe that their memory is much poorer than that of others from their own age group (Crombag, Merckelbach, & Elffers, 2000; Magnussen et al., 2006). However, subjective ideas about memory do not always correspond to objective memory performance (Ponds & Jolles, 1996; Ponds, Van Boxtel, & Jolles, 2000). Bandura (1986) suggests that memory self-efficacy depends not only on experiences of perceived performance, but also on a number of psychosocial factors such as anxiety, and as described above, stereotypes about aging.

Although efficacy beliefs are built primarily on past experiences of performance mastery (Bandura, 1977), self-efficacy appraisals predict future levels of cognitive performance even after controlling for the effects of past attainments (Cervone, Jiwani, & Wood, 1991; Cervone & Wood, 1995). Cognitive performance is influenced in several ways. First, self-efficacy beliefs can influence the construction of task strategies that is, planning for attaining performance goals. Second, high self-efficacy beliefs lead to higher levels of effort and persistence in the face of challenging tasks (e.g., Bandura & Cervone, 1983; Stock & Cervone, 1990). Lastly, low self-efficacy beliefs can cause high levels of negative affect, especially anxiety, in the performance situation, which can lead to poor performance (Bandura, 1988). In research with older populations, individuals with a higher sense of self-efficacy outperform those with a low sense of personal efficacy on memory, metamemory, and cognitive tasks (e.g., Berry, 1999;

Berry et al., 1989; Devolder, Brigham, & Pressley, 1990; Hertzog et al., 1990; Seeman, McAvay, Merril, Albert, & Rodin, 1996). Furthermore, older adults in general have lower self-efficacy beliefs about their memory than do younger or middle-aged adults (e.g., Berry & West, 1993; Berry et al., 1989; Gilewski, Zelinski, & Schaie, 1990; Hultsch, Hertzog, & Dixon, 1987; West, Dennehy-Basile, & Norris, 1996). Aging is also associated with an increase in beliefs about the lack of control over one's own cognitive and memory functioning, including the course of decline (e.g., Heckhausen & Baltes, 1991; Hertzog, McGuire, & Lineweaver, 1998; Hultsch et al., 1987; Lachman, 1986; Lachman, Bandura, Weaver, & Elliot, 1995; Lachman & McArthur, 1986).

1.4.4 Memory-related Beliefs and Performance

The pervasiveness of stereotype-based beliefs may account for age-related deficits in memory performance. Using a variety of responses (e.g., recall, recognition), materials (e.g., prose, words), and task contexts (e.g., laboratory, everyday), researchers have observed relationships between memory and self-referent beliefs of control and self-efficacy, and have demonstrated that such associations account for age differences in memory performance (e.g., Berry et al., 1989; Cavanaugh & Poon, 1989; Hertzog et al., 1998; Lachman, Steinberg, & Trotter, 1987; Luszcz & Hinton, 1995; Riggs, Lachman, & Wingfield, 1997; West et al., 1996; Zelinski, Gilewski, & Anthony-Bergstone, 1990). For example, Cavanaugh and Poon (1989) tested young and older adults on immediate and delayed recall of words and text. After controlling for general ability, 25%-53% of recall variance was accounted for by measures relating to control and self-efficacy. The strength of the relationship between memory beliefs and performance was stronger in the older adults than in the young.

While some researchers (e.g., Berry et al., 1989; Cavanaugh & Poon, 1989) have found that a substantial amount of performance variance (25%-53%) is accounted for by beliefs, other researchers report a more modest relationship (3%-15%), (Hertzog & Hultsch, 2000). Participant characteristics and the specific assessments of beliefs and memory may account for the variance. For example Berry et al. (1989) found that the amount of performance variance accounted for by self-efficacy beliefs was stronger for every-day type memory tasks (e.g., remembering a grocery list) than for laboratory-type tasks (e.g., remembering lists of words) suggesting that experience may be an important part of this relationship. Variability may also be related to the specificity of the relationship between beliefs and the type of memory being studied (e.g., Hertzog et al., 1990).

In a forensic setting, pessimistic ideas about one's own memory might have far-reaching consequences. Negative ideas about one's own memory are also associated with elevated *interrogative suggestibility* levels and an enhanced susceptibility to false recollections (Gudjonsson & MacKeith, 1982). The relationship between participant subjective memory beliefs and the psychological construct of interrogative suggestibility is investigated in Experiment 2. Interrogative suggestibility and related empirical research will be introduced in Chapter 2.

In the experiments reported in this thesis, memory self-referent beliefs were assessed with two measures which were administered prior to encoding. Control beliefs were measured with the Metamemory in Adulthood Questionnaire (MIA; Dixon et al., 1988). Memory self-efficacy was measured with four scales from the Memory Self-Efficacy Questionnaire (MSEQ; Berry et al., 1989). (Further information about the scales is provided in the Method Section of Experiment 1).

Adult age differences in control beliefs have been found on several of the MIA scales across multiple samples (Dixon & Hultsch, 1983b). Older adults perceive more change in their memory, report lower levels of memory capacity, and perceive they have less control over their memory than younger adults. Also, self-reports on MIA scales have been shown to predict performance on various cognitive measures. In a sample of women (age range 21-78 years), Dixon, Hertzog and Hultsch (1986) observed low to moderate correlations between MIA scales and measures of intellectual abilities (e.g., verbal comprehension, induction and memory span). Low to moderate correlations were also observed by Dixon and Hultsch (1983a) for the relationship between MIA scales and measures of text recall, showing that poorer memory beliefs were associated with poorer memory performance.

Having introduced the psychological construct of subjective memory beliefs, the chapter next turns to the effects of education, measures of crystallised intelligence, health status, sensory deficits and affective states in moderating the relationship between age and memory performance. The Cognitive Reserve Hypothesis as indexed by education and measures of crystallised intelligence will next be introduced.

1.5 The Cognitive Reserve Hypothesis

The Cognitive Reserve Hypothesis generally refers to the prediction that those who have experienced more enriched socio-economic environments during childhood and early adulthood have more resilient cognitive and / or neurobiological architectures that protect against the aging-related cognitive deficits in early adulthood.Versions of the Cognitive Reserve Hypothesis can be classified as either *passive* or *active*. Passive models are more frequently conceptualised at the neurobiological level, where reserve usually refers to the capacity to replace damaged brain areas. In contrast, active models are most often conceptualised at the cognitive level, where reserve is functional; high reserve individuals are generally viewed as better able to compensate for tissue loss or brain damage (Stern, 2009). Although evidence that educational attainment protects against normative cognitive declines is mixed, there is evidence that those with higher levels of education function have higher average levels of cognitive function throughout adulthood (Tucker-Drob & Salthouse, in press; Van Hooren et al., 2007). Both education and intelligence have been used as proxy measures of brain reserve (Reece & Cherry, 2006). Education and intelligence are highly correlated, and their association varies as a function of the type of intelligence (verbal, performance, full scale) or education measure (schooling, continuing education etc). Two principal dimensions of intelliegence have been researched: *crystallised intelligence* and *fluid intelligence*. Crystallised intelligence represents verbal intelligence, and consists of the knowledge and expertise accumulated over a life time of experience, and is measured by vocabulary and comprehension subtests of IQ tests. Fluid intelligence represents non-verbal reasoning requiring rapid understanding of novel relationships, and is measured by tests such as Raven's Progressive Matrices (Horn & Cattell, 1966). Crystallised intelligence remains fairly intact or even increases across the lifespan, whereas fluid intelligence deteriorates with advancing age (Baltes, 1993). Although education is related to both higher crystallised intelligence and fluid intelligence, crystallised intelligence appears to be particularly sensitive to education (Alley, Suthers, & Crimmins, 2007).

In the present research a demographics questionnaire recorded participant's years of education. Crystallised intelligence was measured by the Wechsler Test of Adult Reading (WTAR; Wechsler, 2001). This reading test involves asking the client to read out loud 50 words that have atypical grapheme to phoneme translations, and allows an initial estimation of pre-morbid intellectual and memory functioning for individuals aged 16 to 89 years. Further information about the test is provided in Experiment 1 – methodology.

1.6 Health

Health status may be important to memory change and the degree to which individuals experience memory problems late in life. Significant amounts of age-related variance in memory performance are accounted for by objective measures of health status (Albert et al., 1995; Nilsson et al., 1997). For example Nilsson et al., (1997) found that age accounted for 34% in performance, while objective health indicators (e.g., systolic blood pressure) accounted for 62%. However, little relations or no association have been reported between memory performance and self-rated health (e.g., Hultsch et al., 1993;

Hultsch et al., 1999; Nilsson et al., 1997; Perlmutter & Nyquist, 1990; Salthouse, Kausler, & Saults,1990). Small but significant amounts of age-related variance have been found using a self-report status indicator (e.g., perceptions of health, reported symptoms and diseases) (e.g., Hultsch et al., 1993; Perlmutter & Nyquist, 1990).The use of more subjective assessments of health (e.g. self-report of health on a 5-point scale) may have accounted for the reduced health-behaviour relationships (Hess, 2005).

Mental health is also implicated in cognitive functioning. A longitudinal study by Paterniti, Verdier-Taillefer, Dufouil, and Alpérovitch (2002) found that high levels of persistent, depressive symptoms were associated with cognitive decline in a sample of 1003 persons aged 59 to 71 years.

In the present research self-report measures of physical and mental health were obtained on a 5-point Likert-type scale and their moderating effect on the age and memory performance relationship was assessed. Objective measures of physical and mental health and the effects of pharmaceuticals were beyond the scope of this thesis.

1.7 Sensory Deficits

Age-related *memory* deficits may not reflect declines in memory *per se*, but rather the effects of aging on sensory processes (Schneider & Pichora-Fuller, 2000). An effortfulness hypothesis proposes that older adult memory deficit may be due to encoding limitations arising from shifting resources to compensate for sensory problems (McCoy et al., 2005; Wingfield, Tun, & McCoy 2005). Many older adults suffer hearing loss, particularly for high frequency sounds, important for the faithful perception of speech. Most older adults with mild to moderate hearing loss do not use hearing aids and for those who do, hearing aids tend to be less than completely effective in correcting the deficit (Schneider & Pichora-Fuller, 2000). In addition, everyday speech often occurs in noisy contexts and this exacerbates the accuracy of speech perception especially for older adults (Tun, 1998). According to the effortfulness hypothesis, people with hearing loss utilise central processing resources to successfully identify items from a degraded sensory trace. However, the extra effort devoted to perceptual processing, results in fewer resources available for rehearsing, elaborating, and / or organizing the information. Thus, this information will be less memorable (McDaniel, et al., 2008). Support for the hypothesis has come from McCoy et al., (2005), and Tun, McCoy, Cox, and Wingfield (2006).

In the present research self-report measures of vision and hearing on a 5-pint Likert-type scale were obtained and their moderating effect on the age and memory performance relationship assessed.

1.7.1 Affective States and Performance

Cognitive performance is also influenced by affective states (Oaksford, Morris, Grainger, & Williams, 1996;). For the purposes of the experiments reported in this thesis, the term affect was specifically differentiated as positive or negative affect as defined by the Positive and Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988). Thoughts produced when anxious or depressed can interfere with other thoughts and behaviours. Negative affective states can lead to poorer performance on tasks requiring cognitive attention (Ingram, Lumry, Cruet, & Sieber, 1987), particularly those measuring memory performance. For example, Dux, Woodard et al., (2008) found that several negative affect measures moderated the relation between objective memory functioning and subjective memory complaints in a sample of healthy participants aged over 65. The authors concluded that negative affect, particularly anxiety sensitivity, distorts the subjective appraisal of one's own memory, such that people high on negative factors report more episodes of forgetting, even in the absence of objective cognitive impairments. Additionally, positive affect as measured by a Dutch translation of the PANAS (Watson et al., 1988) was found to predict free recall in older adults aged 65 to 82 (Hill, van Boxtel, Ponds, Houx & Jolles, 2005). The authors concluded that when un-supported task conditions were present that placed heavy demand on internal processing respources, positive affect facilitated episodic memory in older adults.

In the present research the PANAS (Watson et al., 1988) was used to measure positive and negative affect. The relationship with age and memory performance was assessed.

1.8 Witness Interviewing

1.8.1 The Cognitive Interview

The outcome of the investigative interview can mark the difference between successful and unsuccessful investigations (Fisher, 2010). The type of questions asked, the manner in which they are asked and the structure of the interview can be crucial in terms of both the quantity (amount) and quality (accuracy) of information obtained (e.g., see Loftus, 1975; Milne & Bull, 2001; Tulving, 1991). However, interviewing is a complex skill. Prior to the early 1990s police officers throughout the world generally received limited witness interview training (see Milne & Bull, 2001). Sanders (1986) found that only 2% of his sample of US police officers had undergone any witness interview training. George (1991) surveyed several UK police forces and found that some provided no witness interview training at all while others provided just one day (Dando & Milne, 2009).

In Fisher et al.,'s (1987) analysis, closed questions (the who, what, where, when, why and how) accounted for approximately 90% of all the questions asked. These questions tend to reduce the amount of information provided and encourage witnesses to guess even if they are unsure of the answer. Furthermore, the use of leading or suggestive questions led the witness to provide incorrect information (Fisher et al., 1987; George & Clifford, 1996). (The use of leading questions on witness performance is addressed in Chapter 2 of this thesis).

Incorporating principles from cognitive and social psychology (Fisher & Castano, 2008), the CI was developed in the early 1980s to improve on this aspect of the investigative process. The CI and ECI are primarily designed for use with co-operative eyewitnesses (Fisher & Geiselman, 1992; Hernández-Fernaud & Alonso-Quecuty, 1997) and in cases where most of the evidence comes from eyewitness reports (e.g., assault or armed robbery) rather than physical evidence (Fisher & Geiselman, 1992).

The *cognitive* component of the CI is an application of four basic memory mnemonic strategies directed at explicit memory processes, where an individual is consciously or intentionally recollecting a specific episode (Schacter, 1990). In a CI witnesses are asked to reconstruct in their mind's eye the contextual features (i.e.

environmental, physiological and affective states) that were present at the time of the event in question (*mental context reinstatement*), to report all the details about an event without editing information that is considered trivial or incomplete irrespective of the level of subjective confidence associated with the information (*report everything*), and to recall the event in a different temporal order e.g., from the end to the beginning of the event (*reverse order recall*), and recall the event from a different perspective e.g., from a different witness' point of view (*change perspective*). The strategies are underpinned by two over-riding perspectives in cognitive theory regarding recollection from memory. The first perspective is the *encoding specificity principle* (Tulving & Thomson, 1973), according to which memory is enhanced when conditions present during retrieval match those that were present during encoding. The second perspective is the *multicomponent* view of the memory trace (Bower, 1967) which proposes that memories inaccessible with one type of retrieval probe can be made accessible with another (Holliday, Brainerd, Reyna, & Humphries, 2009). The CI was observed to be approximately 30% more effective when compared to a standard police interview in which interviewers used their normal everyday interview procedure, and other forms of forensic interviewing (e.g., Geiselman et al., 1984; Geiselman et al., 1985; Geiselman, Fisher, Cohen, Holland, & Surtes, 1986; for a review see Fisher, 2010).

In addition to the use of cognitive principles to increase the amount of information obtained, Fisher and Geiselman's (1992) ECI emphasises social and communication factors. Forensic interviews may reflect an imbalance of social status and power (e.g., a police officer and a child or vulnerable witness). Interviewers are therefore encouraged to develop a rapport with the witness from the outset. Control of the interview is transferred to the witness so that witnesses perceive themselves to be the *experts* and therefore the dominant person in the interview. Active witness participation is encouraged by asking open-ended questions and not interrupting the witness's narrative (Holliday et al., 2009). Altogether, the ECI includes thirteen basic skills: establish rapport, active listening, encourage spontaneous recall, use of open-ended questions, allow pauses, avoid interruptions, request detailed descriptions, encourage intense concentration, the use of mental imagery, context reinstatement, adopt the witness perspective, witness compatible questioning, and following the sequence of the CI (Fisher & Geiselman, 1992). Empirical investigation of the efficacy of the ECI found that when compared to the original CI, the ECI elicited 45% more correct items of information with no differences between the two conditions in the amount of errors (Fisher et al., 1987). As the original CI was found to be approximately 30% more effective than a standard police interview (Geiselman et al., 1985), it was concluded that the ECI produced 75% more correct recall when compared to a standard recall (Dando & Milne, 2009).

1.8.2 Efficacy of the Cognitive Interview

More than 100 studies have evaluated the CI procedure since its inception (Fisher & Castano, 2008). Both forms of the CI have been shown to improve witnesses' correct recollection of events with a number of different populations e.g., children (Akehurst, Milne, & Köhnken, 2003; Hayes & Delamothe, 1997; Holliday, 2003a, 2003b; Holliday & Albon, 2004; McCauley & Fisher, 1995, 1996; Memon et al., 1996, Milne & Bull, 2002; 2003), children with mild learning difficulties (Robinson & McGuire, 2006), young adults (Fisher et al., 1987; Fisher, Amador, & Geiselman, 1989), adults with mild learning disabilities (Milne, 1999) and in a number of languages (other than English) including German (Köhnken, Schimossek, Aschermann, & Hofer, 1995), Portuguese (Stein & Memon, 2006) and Spanish (Hernández-Fernaud & Alonso-Quecuty, 1997; Campos & Alonso-Quecuty, 1999). The CI procedures have also proved beneficial when evaluated with adult victims and witnesses of real crime (Fisher et al., 1989; George, 1991, cited in George & Clifford, 1996; Fisher & Castano, 2008). Some studies have reported a slight increase in the number of errors recalled using the ECI. However, this has not affected the overall accuracy rates calculated as the proportion of correct details relative to the total number of details reported (for a meta-analysis see Köhnken, Milne, Memon, & Bull, 1999; Memon, Meissner, & Fraser, 2010).

Thus from an applied perspective, there is a significant body of research to support the superiority of both the original and the enhanced CI procedures when compared with a standard police interview procedure. However from a theoretical perspective, as standard police interviews have been found to be less than adequate, as described above, it may be that the ECI superiority effect is simply as a result of its comparison to such poorly conducted standard interviews (Köhnken, Thurer, & Zoberbier, 1994). A more theoretical approach tends to employ a structured interview as a control. This is a variant of the ECI procedure minus the *cognitive* mnemonic components. Similar results were observed in terms of enhancing correct recall without a concomitant increase in errors (Köhnken et al., 1994; Memon, Wark, Bull, & Köhnken, 1997). Therefore irrespective of the control interview, researchers have consistently found that the CI/ECI enhances the quantity of information recalled by witnesses without jeopardising its quality (Dando & Milne, 2009).

1.9 Interviewing the Older Eyewitness

Most studies of age-related differences have analysed older adult eyewitness judgements in the context of line-up identifications (Havard & Memon, 2009; Memon & Bartlett, 2002; Memon & Gabbert, 2003a, 2003b; Rose, Bull, & Vrij, 2005; Searcy, Bartlett, & Memon, 2000; Wilcock, Bull, & Vrij, 2005, 2007; Yarmey, 1993). In general, older witnesses demonstrate poorer identification performance. Havard and Memon (2009) found that older adults (aged 61-83 years) performed more poorly in target present and absent line-ups than young adults (aged 18-35 years). While Searcy et al., (2000) found that older participants made more false choices of a line-up ‘foil’ than did younger participants. Although false identifications have and are likely to be a major source of miscarriages of justice, as described in the opening paragraphs of this thesis (e.g., Sello, 1911), Sporer (2008) proposes that person descriptions and recall of event details should be investigated more thoroughly as done by Stern (1903-1906) and Lipmann (1935). Age differences however, have also been found in terms of both the quantity and quality of older eyewitness accounts (Wilcock, 2010).

In research using a mock witness paradigm in which participants view a staged crime event and their memory for the event is later tested, a poorer memory for details relating to the perpetrator, the victim, what happened, and the surroundings was observed in older witnesses when compared to younger witnesses (Yarmey, 1982;

Yarmey & Kent, 1980; Yarmey, Jones, & Rashid, 1984). When the findings of these three studies were averaged, older adults (mean 70 years) were 20% less accurate in free recall, 13% less accurate in cued recall, and 15% less complete in their descriptions of the perpetrator than younger adults (mean 21 years) (Yarmey, 2001).

Qualitative differences in the recall between younger and older adults have also been observed. For example, Yarmey et al., (1984) showed mock witnesses a crime event in which the perpetrator was seen carrying a knife. The knife was reported by 80% of younger witnesses compared to just 20% of older witnesses. The same event included an eleven year old girl who had long hair worn in a pony tail. The girl was misidentified as a boy by 75% of older witnesses, whereas no younger witnesses made that mistake. Both of these findings relate to *central* aspects of a crime, important for a police investigation. Furthermore, descriptive details relating to the physical characteristics of a perpetrator and details relating to clothing are more likely to be absent in older witnesses accounts compared to younger witnesses accounts (Brimacombe, Quinton, Nance, & Garrioch, 1997).

1.9.1 The Cognitive Interview with Older Witnesses

Relatively few studies have assessed a CI's effectiveness to facilitate the event memory of older adults (Dornburg & McDaniel, 2006; Holliday et al., in press; McMahon, 2000; Mello & Fisher, 1996; Prescott, Milne, & Clarke in press; Wright & Holliday, 2007). This research has resulted in inconsistent findings.

Mello and Fisher (1996) were the first to investigate the effectiveness of the CI with older adults. Using a mock witness paradigm (previously described), younger and older adults were interviewed with either a standard police interview, a CI, (both as described above), or a modified CI (MCI) in which the change perspective strategy was omitted. Additionally, MCI interviewers slowed the pace of the interview and used simple vocabulary. The results revealed no age effect on recall. Additionally, no significant differences in performance between the modified CI and the CI were observed. However, the CI led to more information compared with the standard police interview. The advantage of the CI over the standard police interview was greater in the recall of older compared with younger witnesses. Of note, Mello and Fisher (1996) recruited the older adult participants from an institute which offered educational courses for older adults who may not be representative of the older population as a whole (Mello & Fisher, 1996).

Dornburg and McDaniel (2006) reported increased recall of correct details with CI instructions when young and older adults were tested repeatedly on their memories for a story. Participants were not interviewed using the phased approach characteristic of CI protocols and there was no question phase. Additionally, they were given one CI instruction per retrieval attempt over some weeks. Moreover, recall of a story involves retrieval from semantic memory. Recall 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 (Beail, 2002; Cardone & Dent, 1996; Scullin & Ceci, 2001). Semantic memory is less sensitive than episodic memory to impairment by age (e.g., Nyberg et al., 2003).

Wright and Holliday (2007) examined the performance of three age groups (17-31; 60-74; 75-90) when given a standard police interview, the ECI or an MCI whereby the change perspective strategy was omitted. There was a significant effect of age group: the recall of 75-95 year olds was less complete and less accurate compared with 60-74 year olds, which was in turn less complete and less accurate compared with 17-31 year olds. The ECI and MCI increased the number of correct person, action, object and surrounding details reported across every age group, without increasing the number of incorrect or confabulated details recalled (Wright & Holliday, 2007).

More recently Holliday et al., (in press) employed a mock witness paradigm to examine the effect of a Modified CI (MCI) on young and older adults' recall and subsequent reporting of misinformation. The CI was modified in that the instruction to form a mental image was given prior to the questioning phase and not as is usual in CI protocols, in the free recall phase of the interview. When compared to a Structured Interview, the MCI elicited more correct details and improved overall accuracy in both age groups. Older adults interviewed with an MCI were not susceptible to misinformation effects.

The Structured Interview (as previously described) was developed by Memon and her colleagues (Memon et al., 1996; Memon & Stevenage, 1996). The advantages associated with the CI usually have been reduced when compared with the structured interview (Fisher, 1998). McMahon (2000) found the CI to be no more effective compared with the structured interview in eliciting correct information. However, a small participant sample (N = 40) split into four groups and large participant age ranges (i.e. younger adults = 32 years; older adults = 38 years) limit the generalisability of their findings to a wider population (McMahon, 2000).

In summary, research that has assessed a CI's effectiveness to facilitate the event memory of older adults has resulted in inconsistent findings. This may be explained by the differing methodologies employed including different control interviews (Standard *vs.* Structured), different versions of the CI (complete *vs*. versions in which the change perspective mnemonic, or both the change perspective and change order mnemonics were omitted), small sample sizes and an older adult sample who may not be representative of the older population as a whole (Mello & Fisher, 1996). Additionally, because interviewers were aware of the experimental hypotheses, their expectations may have resulted in systematic differences in the length of interviews, or in the number and quality of questions asked in the interviews (Hayes & Delamothe, 1997).

As standard police interviews have been found to be less than adequate, as described above, the use of the standard interview as a control has been criticised (Memon & Stevenage, 1996). Mello and Fisher's (1996) participants interviewed with the standard interview were asked to provide a narrative of what they remembered about the event. Interviewers subsequently probed statements in the narrative when elaboration was deemed necessary. Interviewers did not interrupt participants but long pauses were not allowed. Direct questions followed shortly after a response to a previous question. As previously discussed, early research investigating the efficacy of the CI typically compared this technique with a standard interview in which the interviewers were neither trained in the use of the retrieval strategies that are central to the CI nor taught strategies to build rapport, or facilitate the obtaining of information from interviewees (e.g., use of open-ended questions). The strongest effects of the CI have been reported in studies in which it has been compared with this type of standard interview (Memon & Stevenage, 1996; Fisher 1998). In light of the inconsistent findings from the research presented, we are unable to draw a conclusion as to whether the CI is beneficial in the recall of older adults.

1.10 Quantity-based Measures vs Accuracy-based Measures of Performance

Traditionally memory research has favoured *input-bound*, quantity-based measures, reflecting the likelihood that each input item is correctly remembered. Therefore, measures of memory performance have typically been calculated conditional on the input, by expressing the number of items recalled or recognised as the proportion or percentage of the total number of items presented. More recently, increasing interest has been directed to *memory accuracy* i.e. the extent to which a person's memory report accords with that presented. These measures are known as *output-bound proportion correct*. The number of items correctly recalled or recognised is expressed as a proportion of the total number of items reported. Input-bound measures hold the rememberer responsible for what he or she fails to report, whereas output-bound measures hold the rememberer accountable only for what he or she does report (Koriat & Goldsmith, 1996).

The output bound assessment of memory accuracy is particularly suited to situations such as eyewitness testimony, where a high premium is placed on obtaining memory reports that can be relied upon (see e.g., Deffenbacher, 1991, 2008; Fisher et al., 1989; Hilgard & Loftus, 1979; Loftus, 1979; Poole & White, 1991, 1993; Wells & Lindsay, 1985; Wells & Loftus, 1984).

1.11 Report Option

Output-bound accuracy and input-bound quantity measures can be distinguished operationally under conditions of *free report*, that is, when rememberers are implicitly or explicitly given the option either to volunteer a piece of information or to abstain (e.g., respond "I don't know" Neisser, 1988). Most everyday situations are of this sort. Given the option of free report, e*yewitness motivation* refers to the fact that people are motivated to give an open and honest account (Undeutsch, 1982) so that their performance is mediated by a decision process used to avoid incorrect answers or illusions of familiarity (Burgess & Shallice, 1996; Goldsmith & Koriat, 1999, Kelley & Jacoby, 1996; Klatzky & Erdelyi, 1985; Koriat & Goldsmith, 1994; Schacter et al., 1998). In contrast, when memory is tested through a forced-report procedure, in which people are required to answer each and every question (as in standard forced-choice recognition tests), memory quantity and accuracy measures are operationally equivalent. This is because the number of output items is the same as the number of input items. The option of free report is essential when the focus is on output-bound memory accuracy (Koriat & Goldsmith, 1994, 1996).

1.12 Test Format

The manner in which questions are asked may affect the accuracy of eyewitness reports. In production questioning (open-ended or recall) the witness produces his or her own answers, whereas in selection testing (specific questioning or recognition) the witness chooses a response from options provided by the interviewer. The testing procedures involving recognition or directed questioning can have contaminating effects on memory (e.g., see Brown, Deffenbacher, & Sturgill, 1977; Fisher & Patterson, 2004; Gorenstein & Ellsworth, 1980; Hilgard & Loftus, 1979; Ibabe & Sporer, 2004; Koriat & Goldsmith, 1996 – Experiment 1; Loftus, 1979, 1982; Loftus & Hoffman, 1989). For example Fisher and Patterson (2004) found that responses to free recall probes (e.g., Describe the robber), were almost perfectly accurate whether witnesses were tested after a few minutes (proportion correct = 0.97) or after two weeks (0.94). By comparison, responses to cued recall tests (e.g., What colour was the robber's jacket?) were considerably less accurate (0.70) and (0.54) after two weeks. Furthermore, accuracy for responses to multiple-choice recognition tests (e.g., What colour was the robber's jacket: blue, white, green, red?) were also poor; (0.74) and (0.64) after two weeks (Fisher & Patterson, 2004).

However, in traditional list-learning experiments, superior performance has been found with recognition testing (e.g., Brown, 1976; Shepard, 1967; but see Tulving & Thomson, 1973). As laboratory experiments have historically focused on memory quantity rather than accuracy, Koriat and Goldsmith, (1994, 1996) propose that such a discrepant pattern of results reflects an interaction between *test format* (recall-production *vs*. recognition) and *memory property* (accuracy *vs*. quantity): recognition testing is superior to recall testing in terms of memory quantity performance, but recall testing yields greater memory accuracy (see Hilgard & Loftus, 1979; Lipton, 1977; Neisser, 1988). Testing procedures that differ in test format often also differ in report option. In free-recall testing people produce their own answers (production format) and report only what they feel they actually remember (free report) whereas in forced-choice recognition testing, people choose between the alternatives presented (selection format) and are also required to answer each item (forced report). When the factors of memory property, report option and test format are contrasted, they can be seen to have different effects on quantity-based and accuracy-based measures of memory performance: Recognition tasks elicit more complete (quantity) information, while at the same time reducing the degree of accuracy of the information provided (Clifford & Scott, 1978; Yuille & Cutshall, 1989). Memory accuracy depends primarily on report option (free more accurate than forced) (Koriat & Goldsmith, 1994).

1.13 Type of Content

Memory performance for events may also vary according to the type of content. Within a legal context a distinction can be made between *central* and *peripheral* information. Ibabe and Sporer (2004) note that decision makers are usually concerned with central details that are of legal relevance i.e. who did what to whom and can include the actions and descriptions of the main perpetrators. Peripheral information can be described as all other background information. The distinction between central and peripheral information is also important as researchers have found that mock jurors are more likely to believe witnesses who answered questions about peripheral details correctly (Wells & Lieppe, 1981). Central information is remembered better than peripheral information (Burke, Heuer, & Reisberg, 1992; Christianson & Loftus, 1987, 1991; Heuer & Resiberg, 1990; Ibabe & Sporer, 2004), and irrespective of the emotional content of the material (Heath & Erikson, 1998; Migueles & Garcia-Bajos, 1999; Wright & Stroud, 1998). The type of content as distinguished between *action vs. descriptive* details can further influence eyewitness reports (Aizpurua, Garcia-Bajos, & Migueles, 2009; Burke et al., 1992; Clifford & Scott, 1978; Garcia-Bajos & Migueles, 2003; Greenberg, Westcott, & Bailey, 1998; Ibabe & Sporer, 2004; List, 1986; Tichner & Poulton, 1975).

Ibabe and Sporer (2004) investigated memory for centrality of information (central *vs.* peripheral) and type of content (action *vs.* descriptive details) as a function of three question forms (open-ended, true-false (T-F) and four-alternative-forced-choice (4-AFC), based on a violent robbery scenario. Central action details were defined as those behaviours relating to central characters and that were contemporaneous to the critical event. Peripheral action details included behaviours of non central characters, or of central characters whose actions did not take place during the critical event. Central descriptive details were defined as physical characteristics of scenes, persons and objects related to the critical event. While peripheral details referred to descriptive information unrelated to the event itself.

The results revealed that memory accuracy was greater when participants reported action compared with descriptive details. A greater number of central action details were remembered compared with peripheral action details. Centrality of information made no difference to the accuracy of descriptive details. Confidence in correct and incorrect responses was significantly higher when participants reported action compared with descriptive details and in respect of central compared with peripheral information (Ibabe & Sporer, 2004). Such findings were explained in terms of the weapon focus effect (cf. Loftus, Loftus, & Messo, 1987; Steblay, 1992) whereby attentional focus is narrowed with central information captured first. Furthermore, it was suggested that action details are more likely to be the natural focus of attention in an ongoing event (Ibabe & Sporer, 2004).

Research investigating the event memory of younger and older adults has revealed a discrepant pattern of results. Aizpurua et al., (2009) investigated memory where participants reported actions, people and details of a robbery scenario. Actions were defined as verbal behaviours (e.g., came in shouting ‘this is a stick up!’) or non-verbal behaviours (e.g., pointed a gun at someone’s face). The people related content included general physical characteristics (e.g., age) or more specific features (e.g., goatee) of the people involved, as well as descriptions of their clothing and accessories (e.g., jeans). Details referred to primary object characteristics (e.g., red sports bag) and circumstantial information surrounding the event.

Aizpurua et al., (2009) found that participants accepted more false actions, thus achieving higher recognition accuracy for people and details. Participants also categorised false alarms in respect of actions more often as remember than as know or guess judgments. This pattern of results was more pronounced in the recollection of older adults. Such findings were explained in terms of the prior knowledge participants have about crimes in the form of situational schemata or scripts which can bias memory for the information of an event, particularly actions (e.g., Garcia-Bajos & Migueles, 2003; Greenberg et al., 1998; List, 1986; Migueles & Garcia-Bajos, 2004). Thus participants and particularly older adults displayed a greater bias when recognising actions than other contents (Aizpurua et al., 2009). However, it should be noted that the differing conceptualization of type of content as well as the differing test formats used in the two experiments detailed above may have contributed to the variance in results.

An alternative explanation is provided by the associate deficit hypothesis (Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2003). As introduced in Chapter 1, the associate deficit hypothesis attributes age-related memory deficits to the inability to encode and bind together items or features of an event. Old and Naveh-Benjamin (2008) investigated the associate deficit hypothesis to displays of people and their performance of everyday actions. Younger adults (aged 18-31) and older adults (aged 65-81) viewed a series of video clips each showing a different person performing a different action. Performance was measured as memory for individual people, individual actions, and the person-action associations. Older adults displayed a deficit in memory for people bound with their actions when compared with memory for individual people or actions.

This opening chapter will now introduce a witness's use of metacognitive judgments to determine whether information is volunteered or withheld during an initial recall.

1.14 Retrospective Memory Monitoring

1.14.1 The Accuracy-Confidence Relation

Confidence expressed about a memory can be used to infer its accuracy, both by the general public and by legal professionals (Cutler, Penrod, & Stuve, 1988; Lieppe, 1980; Lindsay, Wells, & O'Connor, 1989; Luus & Wells, 1994; Penrod & Cutler, 1995; Potter & Brewer, 1999). Additionally, witness confidence in his or her testimony appears to be a strong determinant of the perceived credibility of the eyewitness (Lieppe, Manion, & Romanczyk, 1992; Lindsay et al., 1989). As a consequence people who are less confident in their own memory may be erroneously treated as less credible witnesses (Van Bergen, Jelicic, & Merckelbach, 2009). Eyewitness research on the accuracy-confidence (AC) relation has heavily focused on person identification. From a judicial perspective, the Neil vs Biggers (1972) judgement in the US Supreme Court indicated that eyewitness confidence should be taken as an important guide to accuracy. However, a decision given with high confidence may also be incorrect (Brewer & Weber, 2008).

The traditional measurement of the AC relation in eyewitness studies is to calculate a point-biserial correlation coefficient. Researchers have shown the AC relationship to be weak (Bothwell, Deffenbacher, & Brigham, 1987; Penrod, Loftus, & Winkler, 1982; Sporer et al., 1995; Wells & Murray, 1984). For example Sporer et al., (1995) reported an average coefficient of .29 over 30 studies ranging from 1985 to 1994 and concluded that 'experts probably should, at a minimum, advise jurors that witness confidence is one, but only one, indicator of witness accuracy' (p. 324). However, Sporer (2008) also notes that the AC relation for memory of event details may be higher. Witnesses may be able to differentiate between correct and incorrect details (Sporer, 2008).

Researchers have argued that a calibration approach should be used to analyse the AC relationship as it provides more forensically relevant information about the AC relation than a correlation coefficient (Weingardt, Leonesio, & Loftus, 1994; Juslin, Olsson, & Winman, 1995). Specifically, calibration reflects the overall relation between confidence rating and accuracy over different confidence levels. The proportion of correct identification responses made within 100% confidence, 90% confidence and so on is determined. The data is represented in a calibration graph. Confidence (x – axis) is calculated separately for different confidence levels and is then plotted against accuracy (y – axis). A number of aspects of the AC relationship can be examined. A calibration statistic (C) indicates the degree that a response made with a particular level of confidence was accurate (Brewer & Wells, 2006; Juslin et al., 1995). Over- or underconfidence (OU) can be calculated (Brewer & Wells, 2006; Juslin et al., 1995). Finally, a resolution statistic (NRI; see Brewer & Wells, 2006) which indexes the extent to which responses discriminate between correct and incorrect responses can also be calculated. In contrast to the correlation coefficient, knowledge of the AC calibration in a given situation can allow the use of a confidence judgement as a direct estimate of the probability of accuracy of any single eyewitness identification decision (Brewer & Weber, 2008).

Adopting a calibration approach in both eyewitness identification and face recognition paradigms, researchers found that calibration curves indicated a positive linear AC association, when confidence was assessed immediately after the identification response. The AC relationship was evident for choosers i.e. participants who made a positive identification response, but not for non-choosers. Children (aged 10-13) showed substantial overconfidence and poor resolution. Participants displayed overconfidence with interventions designed to reduce overconfidence, resulting in improved calibration for adults but not children (Brewer & Day, 2005; Brewer, Keast, & Richworth, 2002; Brewer & Wells, 2006; Keast, Brewer, & Wells, 2007; Weber & Brewer, 2003, 2004, 2006).

The AC relation for witnesses' descriptions of people and events has revealed mixed results. Researchers have found that both the between- and the within-subjects AC relations are weak (Smith, Kassin, & Ellsworth, 1989). However, meaningful AC relationships at both the between- and within-subjects levels have also been observed (Robinson & Johnson, 1996; Ibabe & Sporer, 2004). Others have reported reliable relationships only at the within-subjects level (Perfect, 2002). For example, Ibabe and Sporer (2004) investigated memory accuracy and confidence in young adult participants (mean 19 years) as a function of three question forms (open-ended, true-false (T-F) and four-alternative-forced-choice (4-AFC), centrality of information (central *vs.* peripheral) and type of content (action *vs.* descriptive details), based on a violent robbery scenario. Participants were significantly more confident when reporting correct compared with incorrect answers. Confidence in correct answers in response to open-ended questions was lower compared with (T-F) and (4-AFC) questions. Confidence was significantly higher in respect of correct and incorrect action compared with descriptive details and in respect of central compared with peripheral information (Ibabe & Sporer, 2004).

As well as demonstrating the malleable nature of eyewitness memory during the interviewing process (cf. Loftus, 1975, 1989; Malpass, 1996), research has shown that witness confidence judgments are also malleable. The positive and negative effects of repeated retrieval or retrieval practice are highly relevant to the study of eyewitness reports. As described in the opening paragraph of this thesis, an initial interview is often followed by additional interviews during later stages of the investigation. Witnesses may provide new information during follow-up questioning, when information not remembered initially may be remembered (Turtle & Yuille, 1994). However, repeated interviewing may also introduce memory distortion, and it offers witnesses the opportunity to practice retrieval of their memories. Retrieval practice may also affect the level of confidence that is expressed by witnesses about the accuracy of their memory (Granhag, 1997; Granhag, Stromwall, & Allwood, 2000; Hastie, Landsman, & Loftus, 1978; Odinot, Wolters, & Lavender, 2009; Shaw, 1996; Shaw & McClure, 1996; Turtle & Yuille, 1994).

A repeated test-schedule can both increase (Hastie et al., 1978; Odinot et al., 2009; Shaw, 1996; Shaw & McClure, 1996) and decrease (Granhag, 1997; Granhag et al., 2000; Turtle & Yuille, 1994) the degree of overconfidence. For example, Odinot et al., (2009) found a significant inflation of witness confidence in respect of retrieval practise items, regardless of whether the answers were correct or incorrect. Such findings were explained in terms of the *retrieval fluency* hypothesis, according to which post-event questioning about an episodic memory leads to strengthening and consequently an increase in retrieval fluency for the information that is recalled. This in turn may increase confidence. As repeated post-event questioning affects retrieval fluency irrespective of the correctness of response, increased confidence occurs both for correct and incorrect memories (Odinot et al., 2009). A significant decrease in overconfidence found by Granhag (1997) could result from participants being provided with their previously made statements, as they were asked about confidence for the second time. Granhag et al., (2000), found that when participants were asked to repeat confidence judgments, confidence and overconfidence was significantly reduced between Sessions 1 and 2. Between-condition analysis also showed that relative to controls, participants in the repeat condition gave significantly lower confidence judgements (Granhag et al., 2000). A further dissociation between confidence and accuracy has been found when witnesses are provided with post-identification feedback which confirms their response as being correct. This feedback substantially increases witness confidence without any increase in accuracy (Semmler, Brewer, & Wells, 2004; Wells & Bradfield, 1998, 1999).

The reliability of the AC finding has led eyewitness researchers to argue strongly against the probabtive value of confidence assessments volunteered in a courtroom. However, Brewer and Weber (2008) argue that the AC relation is important to applied research as it provides decision-makers with an understanding of witness behaviour; Rather than merely a predictor of accuracy, confidence actually determines what is reported. For example, Koriat and colleagues (Koriat & Goldsmith, 1996; Goldsmith, Koriat, & Pansky, 2005; Goldsmith, Koriat, & Weinberg-Eliezer, 2002) have shown that a person's confidence in memory for general knowledge details actually determines whether they report (or withhold) some details, as well as the level of detail reported. Knowing that eyewitnesses control their memory reports in this way illustrates that studying the AC relationship is important in understanding why eyewitnesses report what they do (Brewer & Weber, 2008). The framework developed by Koriat and Goldsmith (1996) and empirical support for the model, which guided Experiment 1 of this thesis is described in Chapter 3.

1.14.2 Age and the Accuracy-Confidence Relation

There is conflicting evidence, about older adults' ability to assess the accuracy of remembered information (e.g., Dodson & Krueger, 2006; Kelley & Sahakyan, 2003; Pliske & Mutter, 1996). Semantic memory is less sensitive than episodic memory to impairment by age (e.g., Nyberg et al., 2003). Older and younger adults are comparably adept or older adults even more so, at assessing the likely accuracy of responses to questions about general knowledge (e.g., "Who wrote *Alice in Wonderland"?*) that measure well-learned or frequently encountered information (cf. Dahl, Allwood, & Hagberg, 2009; Dodson, Bawa, & Krueger, 2007; Hanson, Rönnlund, Juslin, & Nilsson, 2008; Marquie & Huet, 2000; Perlmutter, 1978; Pliske & Mutter, 1996). For example, Pliske and Mutter (1996) compared 21-year-olds (n = 22) and 68 year-olds (n = 21) and found a tendency for the older group to show less overconfidence and significantly better discrimination (using the gamma coefficient – previously described). Dahl et al., (2009) analysed the calibration measures, and resolution measures (as described above) of 60-93 year olds (n = 1,384) in response to general knowledge questions. No age differences were found for either measure.

In contrast however, older adults are impaired at judging the likely accuracy of responses to questions about recently acquired or episodic memories (e.g., Dodson et al., 2007; Dodson, Bawa, & Slotnick, 2007, Dodson & Krueger, 2006; Kelley & Sahakyan, 2003; Shing, Werkle-Bergner, Li, & Lindenberger, 2009). For example, Dodson, et al., (2007) found that older adults (aged 61-76 years) were poorer than young adults (aged 18-26 years) at judging the accuracy of their responses on a source identification task (i.e. who said what). A *misrecollection* account is proposed to explain the pattern of results. This suggests that age-related memory impairments are due to older adults' vulnerability in making high-confidence errors when answering questions that require memory for specific details about recently learned events (e.g., recollective information). The occurrence of these high confidence errors on some tasks and not on others (e.g., well-learned or frequently encountered information) explains when older adults will and will not show poorly calibrated metamemory (i.e. the ability to monitor and assess the likely accuracy of what they have remembered) (Dodson et al., 2007).

The misrecollection account builds on theories of cognitive aging (as described above) that attribute memory impairments to changes in the capacity to bind together and associate items or features of an event (e.g., Chalfonte & Johnson, 1996; Henkel, et al., 1998; Koutstaal, et al., 2001; Kroll, et al., 1996; Light, et al, 2004; Mitchell et al., 2000, Naveh-Benjamin, et al, 2003; Schacter, et al., 1998). The reduced ability to recollect specific features about past events increases the likelihood that older adults will guess a response or rely on less specific information such as familiarity (Dodson et al., 2007). Support for the misrecollection hypothesis has come from Shing et al., (2009).

1.15 Conclusion

Based on the above review of the witness's age, police interviewing protocols and related empirical research, a number of conclusions can be drawn. The age of the witness is one factor that affects eyewitness performance. Age differences have been observed in both the quantity and quality of older eyewitness accounts. Cognitive performance may deteriorate with age and deficits to episodic memory are more pronounced. Such change has implications for eyewitness testimony which is an example of episodic memory. However, researchers have begun to move beyond conceptualizing cognitive aging merely as a population-level phenomenon. There is a growing appreciation for person-to-person individual differences in the cognitive aging process. The context in which the behaviour occurs, and individual variability in cognitive functioning as well as areas such as health, mood, sensory deficits and memory beliefs have accounted for age-related deficits in memory performance. It is noted that where they exist, issues relating to the physical and mental health of witnesses have the potential to influence their behaviour.

The type of questions asked, the manner in which they are asked, and the structure of the police interview can be crucial in terms of both the quantity (amount) and quality (accuracy) of information obtained from witnesses. Psychological principles of memory, social and communication factors have led to the development of new more effective interview procedures. However, the efficacy of such methods to facilitate the older adult eyewitness testimony has proved inconclusive. The differing methodologies, or small sample sizes and an older adult sample who may not be representative of the older adult population as a whole may provide some explanation.

Alternatively, when given the option of free report, participants may use metacognitive monitoring and control processes to *filter out* their responses. Deficits to either monitoring or control may adversely affect the quantity and / or quality of the older eyewitness accounts. The experiment to be reported in Chapter 3 investigates age-related differences in the monitoring and control procedures within a framework developed by Koriat and Goldsmith (1996). Memory accuracy when reporting action and descriptive details of a to-be-remembered video-taped staged event were investigated.

However older adults may also be psychologically vulnerable to the demand characteristics of the police interview. Such vulnerabilities may impact on the reliability of their statements during police interviews or when giving evidence in court. Both situational factors (the nature of the interaction between circumstances / contextual factors, custodial factors, vulnerability factors, and support factors for example the presence of a lawyer or appropriate adults) and dispositional factors (vulnerability), are relevant in assessing the capacity of an interviewee to cope with police interviews. The experiment to be reported in Chapter 4 assesses the effects of personality and situational determinants within the interrogative context on interviewee responses. Interrogative pressure in the form of e.g., *leading questions*, *feedback, or repeated questioning* can be brought to bear on witnesses that may interact with personality factors and the vulnerabilities of some interviewees. In Chapter 2, the Gudjonsson and Clark (1986) model of interrogative suggestibility, interpersonal pressure and associated research is introduced.

Chapter 2: Interrogative Suggestibility and Interrogative Pressure

2.1 Defining Interrogative Suggestibility

Susceptibility to suggestion and interpersonal pressure (IP) within an interrogative context has been termed *interrogative suggestibility*. Binet (1900, 1905) and Stern (1910, 1938, 1939), are credited with the introduction of the concept and early attempts at its measurement (Gudjonsson, 2003). Both researchers presented participants with static pictures and asked leading questions about them. Leading questions are defined as questions phrased in such a way as to communicate expectations and premises to interviewees and hence to suggest a particular answer (Stern, 1970). Although the use of static pictures (Binet, 1900, 1905; Stern 1910, 1938, 1939), may limit the forensic relevance of the material (Davies, Flin, & Baxter, 1986), such work demonstrated the propensity of suggestive questioning to distort memory recall and testimony. Subsequent research has supported the earlier findings (Burtt, 1948; Cohen & Harnick, 1980; Powers, Andricks, & Loftus, 1979; Trankell, 1958).

The introduction of post-event information contained in leading questions can significantly affect the accuracy of eyewitness accounts (Ainsworth, 1998; Kebbell & Giles, 2000; Loftus, 1975, 1979; Memon & Wright, 2000). Post-event information (PEI) is where a witness may hear or read information about a crime previously viewed, which is then incorporated into their memory. PEI can be encountered in various ways: by being asked about the event by police officers or lawyers, reading about the crime in a newspaper, or talking to other witnesses (Wright & Davies, 1999). Due to the reconstructive nature of memory, information presented after an event can change or add to aspects of a memory report (Hyman et al., 1995; Loftus & Palmer, 1974; Loftus et al., 1978; Loftus & Pickrell, 1995; Wright & Stroud, 1998).

In terms of the formal police interview, Gudjonsson and Clark (1986, p. 84) 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”. 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, and (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 socially interactive nature of interviewing was described in Chapter 1. The second component relates to the information gathering process by which one or more individuals seek through questioning, to elicit to-be-remembered details 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. Leading questions are one type of suggestive stimulus. It is assumed that the acceptance or rejection of mis-leading information is mediated by cognitive mechanisms such as *discrepancy detection* i.e. where differences between post-event suggestions and memory for the original event are detected (Schooler & Loftus, 1986). The use of leading questions in police investigations although not eliminated (Fisher, 2010; George & Clifford, 1992) is widely recognised as problematic (Fisher, 2010), as their use has been shown to produce distorted responses (Fisher et al., 1987; George & Clifford, 1996). This is particularly so if a witness has some *psychological vulnerability* (Gudjonsson, 2003; Redlich, 2004).

2.1.1 Vulnerable Witnesses

Bull (2010) states that there is no internationally agreed definition of the word *vulnerable* with regard to witnesses. In his article he focuses on vulnerable groups such as children and those with learning difficulties. Gudjonsson (2006) defines psychological vulnerabilities as "psychological characteristics or mental state which render a witness prone, in certain circumstances, to providing information which is inaccurate, unreliable or misleading" (p. 68). In this context, psychological vulnerabilities represent potential *risk factors* rather than definitive markers of unreliability. This definition is consistent with the letter and spirit of Code C of Practise of the Police and Criminal Evidence Act (Home Office 2008).

Gudjonsson (2003) suggests that psychological vulnerabilities should not be interpreted in isolation to other surrounding factors. Gudjonsson and MacKeith (1997) argue that the capacity of the interviewee to cope with police interviews depends in some cases on medical, psychiatric, and psychological factors and may include: *circumstances* (e.g., the nature and seriousness of the crime, pressure on the police to solve the crime), *interactions* (e.g., complex interactions between the interviewee, the police, and other persons present in the interview), *personality* (i.e., enduring qualities of the interviewee), and *health* (physical and mental health, mental state). The emphasis is on the nature of the interaction between circumstances / contextual factors, custodial factors, vulnerability factors, and support factors (e.g., presence of a lawyer, appropriate adult), (Gudjonsson, 2003).

Furthermore, Gudjonsson (2006) argues that there are typically four types of psychological vulnerabilities relevant to the psychological or psychiatric evaluation of victims, witnesses and suspects in criminal cases. These are labelled *mental disorder* (i.e., mental illness, learning disabilities, personality disorder), *abnormal mental state* (e.g., anxiety, mood disturbance, phobias, bereavement, intoxication or withdrawal from drugs or alcohol), *intellectual functioning* (e.g., borderline IQ scores) and *personality* (e.g., suggestibility, compliance, acquiescence).

The fourth component, acceptance of the stimulus, emphasises that individuals must perceive the message in the suggestive stimulus to be credible. 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 (1986) definition of interrogative suggestibility, individuals must respond, verbally or nonverbally, to indicate acceptance or non-acceptance of the suggestion.

2.1.2 Primary and Secondary Suggestibility

Interrogative suggestibility bears limited relation to traditional types of suggestibility (Gudjonsson, 2003). On the basis of correctional and factor analytical work, Eysenck and Furneaux (1945) proposed two independent types of suggestibility termed *primary* and *secondary*. 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). Measured using ideo-motor tests such as Hull’s (1933) Body Sway test, primary suggestibility is thought to be associated with motor processes and receptivity to hypnotic induction (Gudjonsson, 2003). Secondary suggestibility appears to encompass a more diffuse set of phenomena than primary suggestibility. Eysenck (1947) proposed that 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 by the persuasion of 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.

Fundamental to the comprehension of secondary suggestibility and therefore interrogative suggestibility, is McDougall’s (1908) emphasis on both motivation and cognitive factors such as the person’s knowledge and confidence in that knowledge. Implicit in the concept of suggestibility is the idea that it refers to some stable tendency of an individual to respond in a particular way to a given situation, such that Prideaux (1919) viewed suggestibility as a general personality trait. However, it has been argued that situational factors also affect suggestibility, whereby the specific nature of the test situation is assumed to be more important than a person's psychological disposition (Baxter, 1990; Krech & Crutchfield, 1948).

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, and (iii) it has a strong uncertainty component which relates to the cognitive processing capacity and functioning of the individual (Gudjonsson & Clark, 1986).

2.2 The Gudjonsson-Clark Theoretical Model

Gudjonsson and Clark (1986) developed a detailed psycho-social 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 individual coping strategies employed to manage these variables. Gudjonsson and Clark's (1986) approach is concerned in part with individual differences in suggestibility and takes into account that responses to police investigative interviewing can vary significantly across individuals.

Within the model, uncertainty, occurs when the interviewee is not sure of the correct answer to a question (e.g., where memory for an event is poor due to a long delay between encoding and retrieval, and / or limited cognitive processing capacity). Interpersonal trust reflects the degree to which the interviewee perceives the intentions and behaviour of the investigator to be genuine and without trickery. Interviewees who are suspicious of the investigator'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, when 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 an investigator's questions. 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 investigator, together with an interviewee's cognitive appraisal and coping strategies that influence an interviewee's susceptibility to suggestions (Gudjonsson, 2003).

Gudjonsson and Clark (1986, p. 88) suggest that police officers enter an interview with a particular cognitive set, "an 'event model'…. and seek to extract information consistent with this model." Therefore, police interviewers may have expectations and premises which influence the type of questions asked, the manner in which they are asked as well as the questioning content. Where questioning derives from too rigid a cognitive set, such that these expectations and premises are erroneous, it is possibile that interviewers may bias or cue a witness to respond inaccurately (Gudjonsson, 2003).

Interviewees also enter an interview with a general cognitive set comprising mood, thinking and expectations; influenced in part by their past experiences. Those who have experience of police interviewing are likely to have a different cognitive set to individuals who are unfamiliar with investigative interviewing procedures (Gudjonsson & Singh, 1984a, 1984b). 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 co-operative on the other. Based on their particular general cognitive set, witnesses will adopt a *general coping strategy* to deal with the interview which can facilitate either a *suggestible* or *resistant* set of responses (Gudjonsson, 2003).

Finally, Gudjonsson and Clark's (1986) model incorporates both *leading questions* and *negative feedback* aspects of suggestibility (Gudjonsson, 1983, 1984). As previously mentioned leading questions 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). The Gudjonsson-Clark theoretical model is shown in Appendix A.

2.2.1 Negative Feedback

Examples of explicit negative feedback include the investigator openly stating that the witness has made a mistake, is lying, and making critical perhaps abusive personal remarks about the witness or suspect. It may be that such direct negative feedback is less prevalent in witness compared with suspect forensic interviewing. However, negative feedback can also 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). Gudjonsson and Clark (1986) propose that interviewees receiving negative feedback will appraise it, then reject or accept it. Negative feedback reduces interviewees' confidence in their own memories and this makes them less likely to compare the suggestions of the interviewer with their own recollection (Schooler & Loftus, 1986).

Rejection of negative feedback may be more likely in a forensic population with a particular cognitive set familiar with investigative interviewing procedures. For example, Gudjonsson and Singh (1984a) investigated the relationship between interrogative suggestibility and the number of previous convictions among a sample of delinquent boys. The results revealed a negative correlation between the number of convictions and the tendency to accept interrogative suggestibility in the form of negative feedback, showing that the greater the number of previous convictions, the more likely to resist attempts at pressure. Distrust of the interviewer may lead interviewees to resist further suggestions. Acceptance or rejection of negative feedback is also dependent on the perceived credibility of the person giving the feedback. For example, Skagerberg and Wright (2009) found that susceptibility to post-identification feedback was only observed when co-witness responses were attributed to a high credibility source.

According to the Gudjonsson and Clark (1986) model, acceptance of negative feedback can result in increased anxiety, reduced self-esteem and increased interviewee uncertainty. In order to deal with uncertainty and the potential threat of inadequate performance, interviewees may employ inefficient coping mechanisms (e.g., a greater reliance upon the interviewer for guidance as to whether a question has been answered correctly (Gudjonsson, 1988b; Emmett, Clifford, & Gwyer, 2003), particularly through facial cues and interviewer demeanour (Baxter & Boon, 2000; Baxter, Jackson, & Bain, 2003). 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 interrogative pressure and a tendency to change initial responses.

Negative feedback is also likely to increase the *psychological distance* between the interviewer and the 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.

2.3 Measuring Interrogative Suggestibility

Judges, litigators, and legal scholars deem witness consistency to be one of the most important features of witness credibility (Brewer, Potter, Fisher, Bond, & Luszcz, 1999; Fisher, Brewer, & Mitchell, 2009; Potter & Brewer, 1999). In the US, a standard federal instruction on witness credibility directs jurors to attend to whether "the witness testified inconsistently while on the witness stand, or if the witness said or did something, or failed to say or do something, at any other time that is inconsistent with what the witness said while testifying" (Committee on pattern Jury Instructions, Sixth Circuit Criminal Pattern Jury Instructions, No. 107, 2005). However, the fallibility of human memory, has been well documented in the literature (e.g., Bartlett, 1932; Clifford & Bull, 1978; Crombag, Wagenaar, & van Koppen, 1996; Cutler & Penrod, 1995; Jelicic et al., 2006;

Johnson et al., 1993; Loftus & Palmer, 1974; Smith, 1930). Another issue is the reliability of evidence given during police questioning by individuals who may be particularly susceptible to suggestion (e.g., Gudjonsson & Gunn, 1982). As a result, Gudjonsson (1984) developed a psychometric measure, the Gudjonsson Suggestibility Scale (GSS 1) in order to assess the degree of susceptibility to either leading questions or negative feedback.

2.3.1 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. Intended primarily to measure the extent an interviewee will yield to leading questions, the scale also measures the extent to which they give in to interrogative pressure in the form of negative feedback. To increase the scale's reliability, the true purpose of the test is not immediately apparent. A broad range of scores is obtained from the general population, forensic populations and people with intellectual disabilities. Testing often takes place in forensic settings, therefore administration of the scale should be 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 both UK subjects (Gudjonsson, 1997) and 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). Research has further investigated the viability of a shortened version of the GSS (Smeets, Leppink, Jelicic, & Merckelbach, 2007) and a computer-administered version (Gorassini, Harris, Diamond, & Flynn-Dastoor, 2006).

2.3.2 GSS Procedure and Specific Measures Obtained

The key outcome measures of the GSS 1 and GSS 2 are *Yield* and *Shift.* Yield is the number of suggestive questions that interviewees give in to, and Shift is 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. The interviewee is asked to verbally free-recall as much as possible about the narrative and the number of *ideas* recalled gives an immediate recall score. A further delayed free-recall score is obtained 50 minutes after the initial recall. Interviewees are then subsequently questioned.

The interviewee is then asked 20 specific questions about the narrative, 15 of which are suggestive questions. *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 repeated 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 administration of negative feedback. *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).

Two forms of interrogative suggestibility proposed by Gudjonsson and Clark (1986) can be measured: (1) susceptibility to misleading questions prior to and post-negative feedback, and (2) vulnerability to negative feedback i.e. the extent to which interviewees change their answers when they are told they have made a number of errors. Individual differences in interrogative suggestibility can thus readily be measured. Use of factor analysis has shown that the two types of suggestibility are relatively independent of each other Gudjonsson (1984; 1991; 2003). (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

In addition to use with clinical and forensic populations, the GSS 1 and GSS 2 were designed as research instruments to investigate the nature and mechanisms of interrogative suggestibility. Interrogative suggestibility is related to a number of other psychological constructs. Scores on the GSS 1 and 2 have been shown to correlate negatively with intelligence showing that the lower the level of intelligence, the greater the vulnerability to interrogative suggestibility (Gudjonsson, 1983; 1988a; Polczyk, 2005; Pollard et al., 2004). Suggestibility scores also tend to correlate negatively with immediate and delayed memory recall (Gudjonsson, 1983; Polczyk, 2005).

Gudjonsson (1988b) found a positive correlation between anxiety and interrogative suggestibility, showing the higher the level of interviewee anxiety, the greater the level of 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. A number of studies have found trait anxiety to correlate poorly with suggestibility (e.g., Gudjonsson, 1983; McGroarty & Baxter, 2007; cf. Gudjonsson, Rutter, & Clare, 1995) lending support to the importance of situational factors (Baxter, 1990; Krech & Crutchfield, 1948). Gudjonsson (1988b) also notes 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).

High scores on the GSS have been associated with low self-esteem (Sing & Gudjonsson, 1984a), 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 high self monitoring styles, (Bain, Baxter, & Ballantyne, 2007). High scores on the GSS were also found to be associated with negative life events (Drake, Bull, & Boon, 2008), while Register and Hihlstrom (1988) found the interrogative suggestibility of a group of college students was not related to their level of hypnotisability as measured by the Stanford Hypnotic Susceptibility Scale, Form C (SHSS-C; Weitenhoffer & Hilgrad, 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 previously noted, levels of suggestibility measured by the GSS depend upon three investigator-led influences: the delivery of negative feedback, the use of leading questions, and the repetition of questions. Acceptance or rejection of mis-leading information contained in leading questions, GSS Yield measure, is assumed to be mediated by cognitive mechanisms such as discrepancy detection (as defined above) (Schooler & Loftus, 1986). In contrast, the GSS Shift measure has been assumed to be a measure principally of the effect of negative feedback, depending primarily upon social factors (Gudjonsson, 1997). The GSS procedure includes leading questions and so does not assess the impact that negative feedback alone may have on interviewee responding. Baxter, Boon, and Marley (2006) adapted the GSS 2 to comprise only *minimally leading* questions, i.e. the overtly leading aspects of the GSS2 questions were removed. Baxter et al., (2006) avoided the term *non-leading* noting that any question may hold implications and so may have the potential to *lead*. Following Baxter (2004), Baxter et al., (2006) argued that interview-based evidence obtained through the use of leading questions may be challenged by defense lawyers in court on the basis that it has been obtained using manipulative interviewing techniques (see also Baxter et al., 2007). The results revealed that negative feedback alone continued to be associated with response change (2.15 i.e. 10.75%), although this change was less than the GSS norm (3.00) as would be expected since the component of interrogative pressure associated with leading questions was absent. The authors suggested that while negative feedback is indeed the primary influence on *Shift*, the leading questions of the GSS 2 make a small but significant difference to this measure. A similar percentage of response change following negative feedback in the absence of leading questions (i.e. 10.81%) was reported by McGroarty and Baxter (2007).

This finding may be of some significance for forensic interviewing. 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, and can lead to a witness providing incorrect information (Fisher, 2010). This is particularly the case if a witness has some psychological vulnerability (Gudjonsson, 2003; Redlich, 2004).

Interviewer behaviour is an important influence on interrogative suggestibility (Bain & Baxter, 2000; Bain et al., 2004; Baxter & Boon, 2000; Baxter et al., 2003). For example in a study by Baxter and Boon (2000), participants received negative feedback delivered by interviewers in 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 demeanor, or one which was unfriendly and critical. It is suggested 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 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 (McGroarty & Baxter, 2007).

2.4.2 Studies using Adapted GSS Protocols

A warning about the possible presence of misinformation in the questioning procedure reduces the suggestive effect of leading questions (Bain et al., 2004; Boon & Baxter, 2000, 2004; Greene, Flynn, & Loftus 1982). For example, Greene et al., (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 displayed 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; 2004) 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, 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, 2004).

Varying the type of feedback results in significant differences to Shift scores (Baxter, under review; Baxter et al., 2006; Boon & Baxter, 2000; McGroarty & Baxter, 2007; Register & Kihlstrom, 1988). For example Register and Kihlstrom (1988) using an amended version of the GSS 1 procedure in which no negative feedback was delivered, found that despite the absence of feedback, participants changed a number of their answers during requestioning. The authors suggested that the repetition of questions without explanation perhaps communicates to participants that their previous responses are in some sense incorrect or inappropriate and should be changed. Alternatively, inconsistencies between first and second answers may occur simply as a result of memory failure (Gudjonsson, 2003). Boon and Baxter (2000) showed that removal of negative feedback lowers Total Suggestibility scores significantly below those obtained using the standard GSS 2 procedure. As described in the opening paragraph of this chapter, the use of certain stimulus materials (e.g., picture slides or the spoken narrative of the GSS 1 and 2) may limit the generalisability of findings to the dynamics of real police interviews concerned with events. In particular, recall of the verbal material presented in the GSS narrative involves retrieval from semantic memory. In contrast, recall of a past event places demands on episodic memory (as described in Chapter 1) (Beail, 2002; Cardone & Dent, 1996; Scullin & Ceci, 2001). As a result, using a mock witness paradigm, described in Chapter 1, researchers McGroarty and Baxter (2007) examined the effects of negative feedback on recall of a videotaped event. The questions were not overtly misleading. The results showed that when compared to neutral feedback, negative feedback resulted in more response changes, higher reported state anxiety and higher ratings of interview difficulty.

2.5 Compliance and Interrogative Suggestibility

Compliance is a concept related to suggestibility. It can be defined as “a tendency of the individual to go along with propositions, requests or instructions, for some immediate instrumental gain” (Gudjonsson, 2003, p. 370). Whereas suggestibility assumes that people accept the information provided, the same does not apply for compliance. Compliance is associated with poor self-esteem (Gudjonsson, Hannnesdottir, Petursson, & Bjornsson, 2002). Furthermore, compliance may be viewed as an ineffective coping mechanism during tasks perceived as stressful or interpersonal conflict (Costa & McCrae, 1992). For example, Gudjonsson and Sigurdsson (2003) investigated the relationship of compliance with coping strategies and self-esteem. The results showed that low self-esteem and denial coping were associated with compliance in both male and female participants. Female participants reported lower self-esteem, were more compliant and used different coping strategies when confronted with a stressful situation.

2.6 Subjective Memory Beliefs and Interrogative Suggestibility

As introduced in Chapter 1, there are large individual differences in how people evaluate their own memory. Some people believe that their memory is much poorer than that of others from their own age group (Crombag et al., 2000; Magnussen et al., 2006). In a forensic setting, pessimistic ideas about one's own memory might have far-reaching consequences. Given the weight that triers of fact attach to confidence (Lieppe et al., 1992) as introduced in Chapter 1, people with pessimistic ideas may erroneously be treated as less credible witnesses (Van Bergen et al., 2009).

People who judge their memory to be very poor because they suffer from the *memory distrust syndrome* are thought to be especially prone to memory distortions (Gudjonsson & MacKeith, 1982). Memory distrust is "a condition where people develop profound distrust of their memory recollections, as a result of which they are particularly susceptible to relying on external cues and suggestions" A number of court cases in which defendants suffering from memory distrust developed false memories, eventually resulted in false confessions (Gudjonsson, 2003, p. 196; Gudjonsson, Kopelman, & MacKeith, 1999). Van Bergen et al., (2009) make the distinction between state and trait memory distrust, the former referring to the cases previously described and the latter manifesting itself as a personality trait.

Of particular interest, Van Bergen et al., (2009) investigated the relationship between subjective memory beliefs and suggestibility, compliance, false memories, and objective memory performance. A Dutch translation of the GSS (Gudjonsson, 1984) was used to measure suggestibility (Merckelbach, Muris, Wessel, & Van Koppen, 1998). Subjective memory beliefs were measured with the Squire Subjective Memory Questionnaire (SSMQ; Squire, Wetzel, & Slater, 1979; Van Bergen, Brands, Jelicic, & Merckelbach, 2010). The SSMQ is a self-report questionnaire, consisting of 18 items that are answered on a 9-point scale (- 4 = disastrous, + 4 = perfect). Sample items are "My ability to recall things when I really try is". Scores are summed to obtain a total SSMQ score (varying from -72 to 72), with a negative score indicating a negative subjective evaluation of one's own memory. The results revealed that subjective memory problems correlated significantly with compliance, however contrary to expectations did not correlate with suggestibility (both Yield and Shift) or false recollections, in an adult sample ranging in age from 17 to 46. The authors suggested that the low level of pressure associated with the GSS explained the results. When people suffering from memory distrust are exposed to high interrogative pressure, it is possible that they become suggestible (Van Bergen et al., 2009).

2.7 Age and Interrogative Suggestibility

Within the eyewitness misinformation paradigm, the literature is inconsistent with regard to age-related changes in susceptibility to post-event information. On the one hand older adults are more susceptible to misinformation than young adults (Cohen & Faulkner, 1989; Gabbert, Memon, Allan, & Wright 2004; Karpel, Hoyer, & Toglia 2001; Loftus, Levidow, & Deunsing 1992; Mitchell, Johnson, & Mather 2003; Schacter, Koutstaal, Johnson, Gross, & Angel 1997). In contrast older adults have not been observed to be more suggestible (Bornstein, Witt, Cherry, & Greene 2000; Coxon & Valentine, 1997; Searcy et al., 2000). Age differences between the studies and diverse methodologies may account for the varying pattern of results.

Within the interrogative suggestibility paradigm as developed by Gudjonsson and Clark (1986), research has primarily focused on comparing the performance of children, adolescents, and adults (e.g., Danielsdottir, Sigurgeirsdottir, Einarsdottir, & Haraldsson, 1993; Gudjonsson & Henry, 2003; Gudjonsson & Singh, 1984; Richardson, Gudjonsson, & Kelly, 1995; Singh & Gudjonsson, 1992; Warren, Hulse-Trotter, & Tubbs, 1991; Young, Powell, & Dudgeon, 2003) which generally indicates that the relationship between age and interrogative suggestibility is curvilinear: children and adolescents are more suggestible than younger adults, who in turn are generally less suggestible than older adults. Specifically, Yield is the highest in children. It is lower in adolescents aged between 12 and 18 years and in adults, and rises again in the older age group. In contrast, Shift is highest in children, decreases in adolescents, but differs from adults, whose scores on Shift are lower, and do not differ from older adults (e.g., Polczyk et al., 2004; for a review see Gudjonsson, 2003).

To the present author's knowledge only one study has investigated individual differences in interrogative suggestibility across younger and older adults. Polczyk et al.,

(2004) compared younger (mean 22.3 years) and older adults (mean 64.1 years) on interrogative suggestibility as measured by the GSS2 (Gudjonsson, 1987a, 1997; Polish version: Polczyk, 2000). Negative feedback was administered. Compared to younger adults, older adults scored higher on measures of Yield, but not on Shift. The participants also completed measures of memory performance (Wechsler Memory Scale), (WMS, Wechsler, 1945; Polish version: Choynowski, 1959) and a self-assessment of memory (Memory Assessment Clinics Self-Rating Scale), (MAC-S, Crook & Larrabee, 1990). Regression analysis revealed that memory performance and memory self-assessment were independent predictors of yielding to suggestive questioning in the group of older adults, when controlling for the effects of age. The authors suggested that older adults may not be more susceptible to negative feedback than young adults because they are more self-confident and less dependent on authority (Polczyk et al., 2004).

The experimental study reported in Chapter 4 follows those cited immediately above in seeking to further develop the Gudjonsson and Clark (1986) theoretical model, by investigating individual differences in response change within a mock witness paradigm across younger, middle-aged and older adults.

Chapter 3: Experiment 1

3.1 Introduction

As introduced in Chapter 1, the output-bound assessment of memory accuracy where the rememberer is held accountable only for what he or she reports (Koriat & Goldsmith, 1996) is particularly suited to situations such as eyewitness testimony. Given the number of miscarriages of justice (described in the opening paragraphs of Chapter 1) a high premium is placed on obtaining memory reports that can be relied upon (see e.g., Deffenbacher, 1991; Fisher et al., 1989; Hilgard & Loftus, 1979; Loftus, 1979; Poole & White, 1991, 1993; Wells & Lindsay, 1985; Wells & Loftus, 1984). Chapter 1 also introduced a number of factors which may affect eyewitness memory performance: The age of the witness (younger *vs.* older adults, Wilcock et al., 2008), test format (free recall *vs.* forced-choice recognition questions, e.g., see Brown et al., 1977; Fisher & Patterson, 2004; Gorenstein & Ellsworth, 1980; Hilgard & Loftus, 1979; Ibabe & Sporer, 2004; Koriat & Goldsmith, 1996 – Experiment 1; Loftus, 1979, 1982; Loftus & Hoffman, 1989), and the type of content (action *vs.* description details, Aizpurua et al., 2009; Burke et al., 1992; Clifford & Scott, 1978; Garcia-Bajos & Migueles, 2003; Greenberg et al., 1998; Ibabe & Sporer, 2004; List, 1986; Tichner & Poulton, 1975) may all affect the accuracy of witness information.

The current experiment assessed the way in which these factors interact and relate to both the quantity of correct information available to the rememberer and the monitoring of this information within a general theoretical model developed by Koriat and Goldsmith (1996). The model specifies that when given the option to report or withhold information, participants use metacognitive monitoring and control processes to strategically regulate their memory accuracy. To the researcher's knowledge the effects of test format, and type of content across younger and older adults within a general theoretical model which specifies the mediating role of metacognitive processes, has not previously been investigated. A second question of interest was whether individual differences in participant education, crystallised intelligence, subjective memory beliefs, health, sensory deficits, and affective states mediated the effect of age and performance either directly or indirectly via monitoring and control processes. (Further information about these psychological constructs and their effect on memory performance were detailed in Chapter 1).

The Koriat and Goldsmith (1996) framework will next be described. A schematic illustration of the model is provided in *Figure 1* below.

3.2 The Strategic Regulation of Memory Reporting: A Metacognitive Framework

As introduced in Chapter 1, metacognition refers to the knowledge or awareness of one's cognitive processes. Researchers have examined the manner in which individuals use metacognitive judgments to monitor the validity of their memories and the accuracy of the monitoring process (e.g., Koriat, 1993; Koriat, Lichtenstein, & Fischoff, 1980; Schwartz, 1994). Other researchers have examined the way in which monitoring processes are used to control the process of remembering, and in regulating memory performance (e.g., Barnes, Nelson, Dunlosky, Mazzoni, & Narens, 1999; Goldsmith &

Koriat, 1999, 2008; Higham, 2002, 2007; Koriat & Goldsmith, 1994, 1996; Koriat, Goldsmith, & Halamish, 2008; Nelson & Narens, 1990). Most relevant to the current research, the examination of monitoring and control processes operating during memory retrieval has provided important insights regarding memory deficits in old age (e.g., Henkel, Johnson, & de Leonardis, 1998; Jacoby, Bishara, Hessels, & Toth, 2005; Kelley & Sahakyan, 2003; Koutstaal 2006; Pansky, Goldsmith, Koriat, & Pearlman-Avnion, 2009; Rhodes & Kelley, 2005).

Also introduced in Chapter 1, report option refers to the implicit or explicit instruction that rememberers can either volunteer or withhold a piece of information. As eyewitness are motivated to provide an open and honest account (Undeutsch, 1982), the accuracy of memory reporting is mediated by a decision process used to avoid incorrect answers. Koriat and Goldsmith (1996) proposed a model of the metamemory *monitoring* and *control* processes that underlie the strategic regulation of memory performance in free report situations. The model is based on signal detection theory (Green & Swets, 1966; Swets, Tanner, & Birdshall, 1961), a framework used to investigate the decision making processes underlying forced-report recognition memory (Banks, 1970; Bernbach, 1967; Kintsch, 1967; Lockhart & Murdock, 1970; Murdock, 1974; 1982; Norman & Wickelgren, 1969). As shown in *Figure 1* below, Koriat and Goldsmith's model implies a separation between three components of memory performance: (a) *retention* i.e. the amount of information that can be retrieved, (b) *monitoring effectiveness* i.e. the extent to which the subject's confidence distinguishes correct and incorrect potential responses, and (c) *a control mechanism* that determines whether the candidate answer is volunteered or withheld (Koriat & Goldsmith, 1996). Central to the model is the two-phase, forced-free paradigm, combined with the elicitation of confidence judgements in the forced-report phase.

Koriat and Goldsmith (1996) suggest that performance at free report depends on three factors: *monitoring effectiveness*, *control sensitivity* and *response criterion setting*. Participants monitor the correctness of their response by assigning it a probability of being correct (P_a). A response criterion (P_{rc}) is then set which controls responding. If (P_a) exceeds or is equal to (P_{rc}) an answer is volunteered, otherwise the response is withheld. The criterion (P_{rc}) can vary based on the level of accuracy incentive. This can be adjusted upward where large losses are associated with a commission error (wrong answer) or downward if there is no penalty for a commission error and the emphasis is placed on the quantity of correct answers. Under high incentives for accuracy compared to low or moderate incentives, people will be able to increase accuracy if they have effective monitoring of the probability of the correctness of candidate answers, good control sensitivity, and effective response criterion setting.

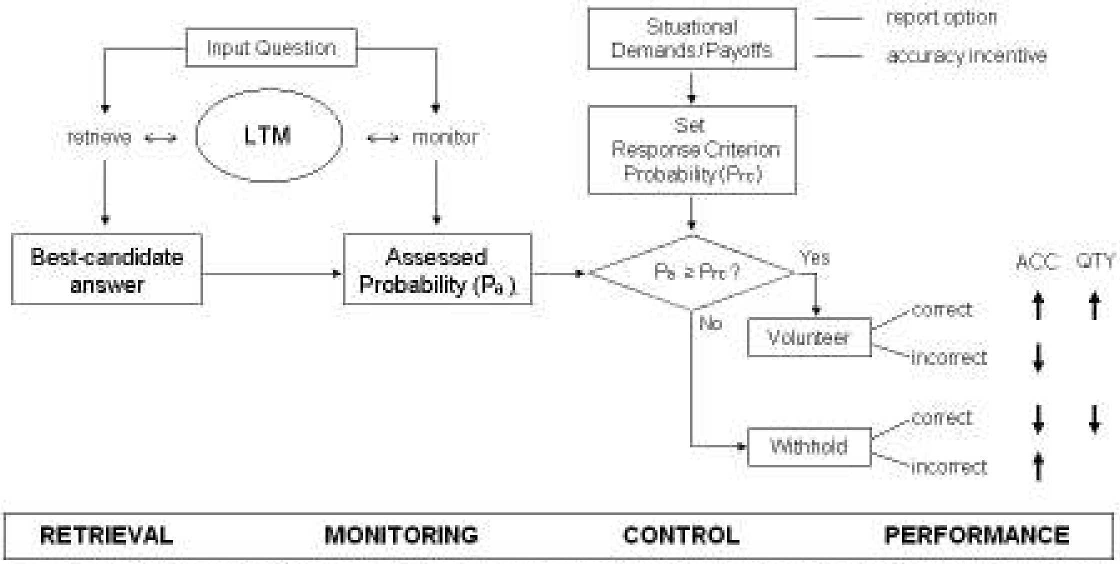


Figure 1. A schematic model of the strategic regulation of memory accuracy quantity performance, utilising the option of free report. The upward and downward pointing arrows on the right side of the figure signify positive and negative performance outcomes.
LTM = long term memory; ACC = accuracy; QTY = quantity; P_a = assessed probability; P_{rc} = response criterion probability
(adapted from Koriat & Goldsmith, 1996)

Monitoring effectiveness is the extent to which assessed probabilities of being correct (P_a) successfully distinguish correct and incorrect responses and is measured by within-subject gamma correlations (Nelson, 1984). Control sensitivity refers to the extent the free-report volunteering and withholding decisions are based on the monitoring output (i.e. confidence ratings at forced-report). Finally, response criterion setting is the minimum level of confidence required by the participant before he / she is willing to volunteer an answer. This can be estimated by finding a cut-off on the confidence ratings which best separate the items volunteered and those withheld at free report. As described in Chapter 1, confidence judgements also provide information about monitoring *per se*: its absolute levels, its calibration (e.g., over / under confidence), and the extent to which it discriminates correct and incorrect candidate answers (monitoring resolution). Resolution (relative correspondence) is critical for subject control mechanism (Koriat & Goldsmith, 1996).

Additionally, monitoring effectiveness is decomposed into two factors, *polarisation* and *correspondence*. Polarisation refers to the distribution of probability assessments. Variability is necessary for the monitoring process to be useful for control. The other extreme is high polarisation, where each response is assigned a probability of only 0 or 100. High polarization will be an effective basis for control, but only if there is also good correspondence between those assessed probabilities and actual probabilities of correctness. People could differ in their monitoring effectiveness because of differences in polarisation and / or differences in correspondence (Koriat & Goldsmith, 1996).

Most previous investigations of the recall criterion have focused on the control processes alone (e.g., see Klatzky & Erdelyi, 1985). A quantity-accuracy trade-off is generally predicted. To the extent that the confidence ratings reasonably reflect the correctness of the candidate answers, and the answers are volunteered or withheld on such a basis, then raising the response criterion will result in fewer volunteered answers, a higher percentage of which are correct (increased accuracy) but a lower number of which are correct (decreased quantity). Because raising the response criterion is assumed to increase accuracy at the expense of quantity, the strategic control of memory performance should require the rememberer to weigh the relative pay-offs for accuracy and quantity in reaching an optimal criterion setting. The extent of any quantity-accuracy trade-off is dependent upon the quality of the participant memory monitoring and control processes (Koriat & Goldsmith, 1996). In sum, when monitoring is poor selective reporting should yield a greater quantity-accuracy trade-off. At the extreme, the withholding of answers on the basis of invalid subjective probabilities could fail to improve memory accuracy and only reduce memory quantity performance (Koriat & Goldsmith, 1996).

The framework distinguishes between quantity and accuracy based measures of performance. Specifically, quantity is measured as the number of correct responses elicited during forced-report, at which time a confidence judgement for each response is also obtained. When given the option to volunteer or withhold a response during free report, participants can regulate the accuracy of their output by exercising monitoring and control processes. Accuracy is therefore measured as the proportion of correct responses volunteered out of the responses offered.

3.3 Empirical Research

Koriat and Goldsmith (1996 – Experiment 1) administered a 60-item general knowledge test in either a recall or recognition format to young adult participants. The test was first completed under forced-report instructions and participants provided confidence judgments as to the correctness of each answer. Immediately afterward they completed the same test again under free-report instructions with either a moderate or high accuracy incentive. The results were generally consistent with the model. First, participants were successful in monitoring the correctness of their answers. The tendency to report an answer under free-report conditions was strongly correlated with subjective confidence in the correctness of the answer. This tendency was also sensitive to accuracy incentive. High incentive participants set a stricter criterion than did the moderate-incentive participants. A quantity-accuracy trade-off was also observed. The extent of the quantity costs relative to the improved accuracy increased, relative to the higher the criterion. Recall participants were more effective than recognition participants in discriminating correct from incorrect answers. Analysis of the response criterion values revealed that recognition participants were significantly more conservative than recall participants in their control policies however, the increased correctness of their candidate answers allowed them to volunteer as many answers as did the recall participants. The net result appeared to be an advantage in recognition quantity performance, achieved at no disadvantage in accuracy compared with recall (Koriat & Goldsmith, 1996).

Koriat and Goldsmith (1996) collected the forced and free report data in two separate phases: first under forced report instructions and then again under free report instructions (or in reverse order – Experiment 2). Alternatively, the free and forced report data can be collected on an item-by-item basis, by first forcing the participant to provide an answer, then eliciting a confidence judgement, and finally having the participant decide whether to volunteer the answer or not (Kelley & Sahakyan, 2003; Rhodes & Kelley, 2005). A consistent pattern of results has been observed across different variations (Goldsmith & Koriat, 2008). Results from various studies have provided support for the model (Danion, Gokalsing, Robert, Massin-Krauss, & Bacon, 2001; Goldsmith et al., 2002; Goldsmith et al., 2005; Higham, 2002; Higham & Tam, 2005; Kelley & Sahakyan, 2003; Koren et al., 2004; Koriat & Goldsmith, 1996; Pansky, et al., 2009; Payne, Lambert, & Jacoby, 2002; Rhodes & Kelley, 2005; for a review see Goldsmith & Koriat, 2008).

3.4 Age and the Strategic Control of Memory Reporting

Of particular interest to the current experiment are studies which have focussed specifically on age-related differences in participants' ability to strategically regulate memory accuracy in free report (Kelley & Sahakyen, 2003; Pansky et al., 2009; Rhodes & Kelley, 2005). Kelley and Sahakyen, (2003, Experiment 1) used Koriat and Goldsmith's (1996) model and methodology together with an associative interference paradigm developed by Kato (1985) to investigate the strategic regulatory processes of younger and older adults. Accuracy in respect of control word-pairs (not expected to elicit associative interference), at forced report was superior by younger adults. The older adults made gains in memory accuracy from forced to free-report, but did so at the expense of greater losses in quantity correct. The age difference in accuracy became larger under free-report than under forced report in respect of 'deceptive' word-pairs (in which the retrieval cues evoke a highly accessible associate that competes with the target, thereby presenting a difficult challenge to memory monitoring). This interactive pattern was explained in terms of monitoring effectiveness. Older adults showed lower levels of monitoring resolution in respect of both deceptive and control items. Although both young and older adults were highly overconfident in the correctness of their responses to the deceptive items, this was more pronounced in the judgments of older adults. Young and older adults showed excellent control sensitivity and based their decision to volunteer items on the assessed probability of that item being correct. Age differences in memory control did appear in the setting of response criterion. Only younger adults responded to the high incentive for accuracy by setting their response criterion higher. Additional experiments suggested that the impaired monitoring of the older participants derived from impoverished encoding. When the encoding of the younger participants was disrupted by divided attention at encoding, a similar pattern of results in terms of both memory accuracy and memory monitoring, to that of the older participants was observed. Thus, Kelley and Sahakyan (2003) suggested that the poorer memory monitoring of older adults derives primarily from their greater reliance on the familiarity of candidate responses rather than on recollection of details of the study experience (Jacoby, 1999; Jacoby, Debner, & Hay, 2001), which in turn stems at least in part from poor encoding.

A similar conclusion was reached by Rhodes and Kelley (2005), who used the same approach to investigate age differences in memory performance, but now tying these to neuropsychological measures of executive functioning. In their study, path analysis supported a model in which aging impairs executive functioning, which in turn impairs retention (forced-report performance – a product of both encoding quality and retrieval), which in turn impairs free-report memory accuracy, both directly, and also indirectly by way of impaired monitoring.

Pansky et al., (2009) examined age-related differences in memory performance for a narrated slide show consisting of 27 colour photographs, specifically differences in metacognitive control. Replicating previous results, younger adults showed superior memory quantity and accuracy performance compared to older adults, an effect which remained when young adults performed under conditions of divided attention. The performance of older adults was explained in terms of less effective memory monitoring and also from differences in two aspects of control: a more liberal report criterion, and reduced control sensitivity. Older adults displayed a greater tendency to report incorrect as well as correct answers and less reliance on subjective monitoring as a basis for responding. Moreover, across both age groups, control sensitivity was highly correlated with two measures of executive functioning, perseverative errors on the Wisconsin Card Sorting Task (Hart, Kwentus, Wade, & Taylor, 1988) and the number of words generated on the FAS word fluency test (Benton & Hamsher, 1976). It was further suggested that diminished control sensitivity could derive from a general age-related difficulty in inhibiting prepotent responses. This capacity has been attributed to the frontal lobe (see e.g., Miyake et al., 2000; Rabbitt, Lowe, & Shilling, 2001), and as introduced in Chapter 1 is thought to be impaired among older adults (see e.g., Zacks et al., 2000).

3.5 Overview of Experiment

As discussed in the opening paragraphs of Chapter 2, the use of certain stimulus materials (e.g., picture slides or tests of general knowledge) may limit the generalisability of findings to the dynamics of real police interviews concerned with events. In particular, tests of general knowledge (Koriat & Goldsmith, 1996, Experiment 1) involve retrieval from semantic memory. Recall 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 episodic memory (as described in

Chapter 1) (Beail, 2002; Cardone & Dent, 1996; Scullin & Ceci, 2001). Furthermore, presenting *actions* as slides (Pansky et al., 2009) is problematic as actions involve movements which cannot be perceived naturally within this method of presentation (Ibabe & Sporer, 2004). To address these issues, the current experiment employed a more ecologically valid mock witness paradigm (also as described in Chapter 1) to examine the monitoring and control processes employed for details of a staged videotaped event. Although common within eyewitness testimony research, the use of such stimulus material had not previously been used in the investigation of this area of witness behaviour.

Participants viewed a 130 second video clip which featured a neutral scene of a passerby asking for directions. An Event Memory Questionnaire (EMQ), consisting of 30 - items based on the content of the film, was presented in either cued-recall or in multiple choice recognition format. Monitoring was manipulated as the 30 questions contained items relating to action, description and dialogue details.

Report option was manipulated as follows: In the first phase, the participants were asked to take the test under free report conditions, deciding which answers to volunteer and which to withhold, with an incentive for accurate reporting. Accuracy was manipulated by one of three incentives: low, moderate or high incentive. In the second phase participants were asked to take the same test again under forced report instructions, in that participants were required to answer all 30 questions (guessing if necessary) and to indicate their confidence in the correctness of each answer. In contrast to Kelley and Sahakyan, (2003) and Rhodes and Kelley (2005), who implemented the two phases on an item by item basis, the current study followed Koriat and Goldsmith's (1996 - Experiment 2) procedure by collecting the forced and free report data in two separate phases. The volunteering of information under free-report instructions was immediately followed by forced-report instructions.

3.5.1 Experimental Hypothesis

3.5.1.1 Memory Accuracy in Free Report

In light of previous findings, detailed above, it was predicted that there would be three main effects of age, test format, and type of content. Memory accuracy in free report was expected to be greater from younger adults, greater in response to cued-recall questions, and greater when reporting descriptive compared with action details. Age-related decrements have been found to be larger in recall compared with recognition (Craik & McDowd, 1987; Rabinowitz 1984, 1986) and greater in recall compared with cued-recall (Ceci & Tabor 1981; Craik et al., 1987), therefore, an interaction may be observed between age and test format. Participants in general and older adults in particular display a greater bias when recognising actions than other contents (e.g., Aizpurua et al., 2009) therefore a further interaction may be observed between age and type of content. Interactive patterns of results may also reflect differences in participant monitoring and control processes.

3.5.1.2 Memory Monitoring and Control

Monitoring Effectiveness

Two indices of monitoring effectiveness were calculated. Both absolute correspondence (calibration) and relative correspondence (resolution) were evaluated. Calibration reflects the overall relation between accuracy and confidence ratings. In a calibration graph, confidence (*x* axis) is calculated separately for different confidence levels and is then plotted against accuracy (*y* axis). The calibration graph indicates over / under confidence (i.e. the extent to which the confidence judgments are higher or lower than the actual proportions correct). Monitoring resolution is the degree to which assessed probabilities of correctness successfully differentiate correct and incorrect candidate answers, and is measured by within-subject Kruskal-Goodman gamma correlations (Nelson, 1984).

Memory bias resulting from situational schemata particularly affects memory in respect of action details (Garcia-Bajos & Migueles, 2003; Migueles & Garcia-Bajos, 2004), that may however be held in high confidence (Ibabe & Sporer, 2004). The second experimental hypothesis predicted that participants' confidence judgments in respect of action items would generally be undiagnostic of the correctness of their answers. The second experimental hypothesis predicted that calibration and monitoring resolution would be relatively good in respect of descriptive items but poor in respect of action items. An interaction was expected as accuracy in respect of action items might be significantly poorer by older adults (Aizpurua et al., 2009).

Control Sensitivity – Subjective Confidence and the Decision to Respond

If participants are distrustful of their memory monitoring as well as their memory *per se*, this might produce a weaker relationship between confidence assessments and the decision to respond or withhold a potential response. Reduced control sensitivity might be observed in older adults (Kelley & Sahakyen, 2003; Rhodes & Kelley, 2005; Pansky et al., 2009).

Response Criterion

According to Koriat and Goldsmith's (1996) framework, the response criterion can vary based on the level of accuracy incentive. This can be adjusted upward where large losses are associated with a commission error (wrong answer) or downward if there is no penalty for a commission error and the emphasis is placed on the quantity of correct answers. Under high incentives for accuracy compared to low or moderate incentives, people will be able to increase accuracy if they have effective monitoring of the probability of the correctness of candidate answers, good control sensitivity, and effective response criterion setting.

People who believe that they have a strong memory for particular stimuli expect to be able to remember those stimuli and set a higher response criterion (Stretch & Wixted, 1998; Morrell, Gaitan, & Wixted, 2002). Conversely, people who believe the opposite (i.e. a weak memory) relax their criterion (Stretch & Wixted, 1998). Several results indicate age-related differences in measures of control. Older adults adopt a more lenient response criteria than young adults (Aizpurua et al., 2009; da Silva & Sunderland, 2010; Kelley & Sahakyan, 2003; Koutstaal & Schacter, 1997; Pansky et al., 2009), and are less likely than younger adults to withhold answers when given the option of free report, even though this leads them to higher rates of false memory (Jacoby et al., 2005; Pansky et al., 2009). Therefore, a liberal report criterion is likely to result in a greater number of volunteered incorrect (and correct) answers.

It might be hypothesised therefore, that older adults would be expected to set a more liberal response criterion and volunteer a significantly greater number of correct (and incorrect) responses in free report.

However, prior research has manipulated accuracy incentive by imposing a monetary penalty for each incorrect candidate response (Koriat & Goldsmith, 1994, 1996; Kelley & Sahakyan, 2003; Pansky et al., 2009). Younger adults respond to such an incentive structure, however older adults do not (Kelley & Sahakyan, 2003; Pansky et al., 2009). Older adults have been shown to respond to various pay-off matrixes regarding bias on a recognition test, although not to the same degree as younger adults (Baron & Surdy, 1990). The present study makes the novel contribution of manipulating accuracy incentive by offering a reward for correct candidate responses, as opposed to a penalty for incorrect candidate responses. In the high incentive condition a monetary payment, payable to a charity of the candidate's choice was offered to the candidate who recorded the most accurate performance. It might therefore be predicted that participants, particularly older adults set a more conservative response criterion, in an effort to improve their accuracy performance. These rival hypotheses will be investigated within the current study.

3.6 Methodology

3.6.1 Design

An individual differences analysis was used to assess the relationship between objective memory performance (accuracy and quantity) and participant years of education, crystallised intelligence, self-reported measures of health, sensory deficits, subjective memory beliefs, and positive and negative affect across younger and older adults, within a general theoretical model which specifies the mediating role of metacognitive processes. The monitoring effectiveness, control sensitivity, and response criterion setting of younger and older adults were measured, as they responded to action, descriptive and dialogue items contained within a to-be-remembered event. The individual difference measures are summarised in *Table 1* in the Results Section. Participants responded to either cued-recall or multiple choice recognition tests, under conditions of high, moderate or low incentive.

A 2 (Age-group: younger adults aged 18-35, older adults aged 60-85) x 2 (Test format: cued-recall, recognition) x 3 (Accuracy incentive: low, moderate, high) factorial design was employed. 106 younger adults and 94 older adults were randomly assigned to one of 6 experimental conditions: 1) Cued-recall, low incentive, 2) Cued-recall, moderate incentive, 3) Cued-recall, high incentive, 4) Recognition, low incentive, 5) Recognition, moderate incentive, 6) Recognition, high incentive.

The between-subjects factors were age group, test format and accuracy incentive. Within-subject measures were accuracy and confidence judgements in respect of correct and incorrect action and descriptive responses at free and forced report. Participant self report ratings on the Positive and Negative Affect Scale (PANAS; Watson et al., 1988) were also within-subject measures.

3.6.2 Participants

The participant sample consisted of 106 younger adults (mean age 23.22, SD, 5.63, range = 18-35) and 94 older adults (mean age 70.22, SD, 5.38, range = 60-85). Younger adults were recruited from the student population of the University of Strathclyde, Glasgow. Older adults were recruited from the University of Strathclyde Centre for Lifelong Learning, community organisations and social clubs within the Glasgow Area. 39 undergraduate students were awarded Psychology course credits. The remaining student and older adult names were entered into a cash prize draw. A condition of participation specified that all participants should be proficient in English language. From an original recruitment sample of 200, 5 (2 younger adults and 3 older adults) were eliminated: 2 young adults and 2 older adults did not complete the experiment in full, and one older adult was not able to view the filmed event from a standard distance. Data analysis was conducted on a sample size of 195. The research was approved by the standing Ethics Committees of the Department of Psychology and University of Strathclyde.

3.6.3 Materials

3.6.3.1 Participant Details Questionnaire

Participants completed a demographics questionnaire providing age, gender and years of education attained. Self-ratings were also provided in respect of everyday vision, hearing, physical and mental health on a 5-point Likert-type scale ranging from (1) *excellent* to (5) *very poor*. The participant details questionnaire is reproduced in Appendix D.

3.6.3.2 Self-Report Scales

Memory beliefs

Memory beliefs were assessed with two measures; the Metamemory in Adulthood Questionnaire (MIA; Dixon et al., 1988); and four scales from the Memory Self-Efficacy Questionnaire (MSEQ; Berry et al., 1989).

The MIA has 108 items distributed among 7 factor derived scales. The 7 scales are:-

1) Strategies reflect the knowledge of one's remembering abilities such that performance in given circumstances is potentially improved. This knowledge includes the reported use of mnemonics and external memory aids. (+ = high use) (e.g., do you keep a list or otherwise note important dates such as birthdays and anniversaries),
2) Task, is a knowledge of basic memory processes, especially of how most people perform (+ = high knowledge) (e.g., for most people, facts that are interesting are easier to remember than facts that are not),

3) Capacity, is the perception of memory capacities, as measured by a predictive report of performance on a given task (+ = high capacity) (e.g., I am good at remembering conversations I have had),
4) Change, is the perception of memory abilities as generally stable or subject to long-term decline (+ = stability) (e.g., I am less efficient at remembering things now than I used to be),
5) Anxiety, is the respondent's rating of the influence that anxiety and stress have on performance (+ = high knowledge) (e.g., I get anxious when I am asked to remember something),
6) Achievement, is the perceived importance of having a good memory and of performing well on memory tasks (+ = high achievement) (e.g., it is important to me to have a good memory), and
7) Locus, is the individuals's perceived personal control over remembering abilities (+ = internal locus of control) (e.g., as long as I exercise my memory it will not decline).

Items were statements followed by a 5-point Likert-type scale. Particpants indicated the extent to which they agreed with the statement from *agree strongly*, *agree*, *undecided*, *disagree* to *disagree strongly*. The strategy subscale items asked participants to indicate their response from *never*, *rarely*, *sometimes*, *often* to *always*. All responses could be positive or negative. Participants were instructed as to the questionnaire completion both verbally and in writing; instructions were printed on the front of the questionnaire together with item examples and information relating to the 5-point Likert-type scale.

Completion of the questionnaire took approximately 45 minutes. This questionnaire has been shown to be internally consistent with Cronbach's alpha α ranging from .71 to .93 (Hultsch et al., 1987). The alpha reliabilities for each subscale for the current sample were Strategy (α = .82), Task (α = .70), Capacity (α = .76), Change (α = .90), Anxiety (α = .84), Achievement (α = .71), and Locus (α = .75).

Memory Self- Efficacy Questionnaire (MSEQ)

Four scales of the Memory Self- Efficacy Questionnaire (MSEQ; Berry et al., 1989) were presented:-

1) household items recall (e.g., If I placed 18 common everyday objects in different locations at home, a few minutes later I could remember where I had put all 18 of the items),
2) shopping list recall (e.g., If I went to the store the same day, I could remember 18 items from a friend's shopping list of 18 items, without using a list),
3) photograph recall (e.g., If someone showed me the photographs of 10 people and told me their names once, I could identify 10 persons by name if I saw the pictures again a few minutes later), and
4) story recall (e.g., If I had just read part of a story (about 10 sentences), I could correctly remember the main points from 2 sentences).

Each scale had five questions representing five levels of difficulty for one task, presented in descending order. Participants indicated how confident they would be about performing a specific task at that level on a scale of 0 (*certain I cannot do it* to 100% (*certain I can do it*). Self efficacy level (SEL) scores were calculated by summing the number of responses made with at least 20% confidence. This is a reflection of individual assessment of his or her basic memory skill level. Responses were summed to form SEL scales for the four specific tasks (five items for each task). Self efficacy strength scores (SEST), which represented an overall indicator of memory self-efficacy, were calculated as an individuals' average level of confidence across the 20 items. Scores on the four scales (five items for each task) were also averaged because, in factor analysis, the MSEQ factors as a single scale (Berry, West, & Cavanaugh, 1996), and shows similar age relationships across scales and moderate to high intercorrelations among scales (Berry et al., 1989; West & Berry, 1994). The alpha reliability for the current sample was (α = .90) when each individual question (N = 20) was entered into the analysis. Participants were instructed as to the questionnaire completion both verbally and in writing; instructions were printed on the front of the questionnaire together with question examples and information relating to the response scale. Completion of the questionnaire took approximately 20 minutes.

In previous research, the order of presentation of two measures (MSEQ and a general memory self evaluation) had no impact on performance or self ratings (West, Welch, & Thorn, 2001), so these items were presented to participants in the same order in this study, memory beliefs measured by the MIA followed by the MSEQ.

Positive and Negative Affect Scale (PANAS)

The PANAS Scale (Watson et al.*,*1988) is a self report measure consisting of 20 adjectives. Participants were asked to rate the extent to which they felt the adjective at that moment in time on a 5-point Likert-type scale from (1) *very slightly or not at all*, (2) *a little*, (3) *moderately*, (4) *quite a bit*, (5) *extremely.* The measure consists of 10 positive affect (PA) adjectives (e.g., interested, excited) and 10 negative affect (NA) adjectives (e.g., irritable, distressed). PA and NA are not considered to be opposites of a single factor, and the lower extreme of each dimension is typified by the absence of the relevant characteristic rather than the presence of the other. The values of PA and NA ranged from a minimum of 10 to a maximum of 50. PANAS questionnaires were completed by all participants both at the outset of the test and on conclusion of the experimental manipulation. The reliabilities of the present data were good. The Cronbach's α value for PA at outset and conclusion was .84 and .88 respectively. The Cronbach's α value for NA at outset and conclusion was .84 and .83 respectively.

Wechsler Test of Adult Reading (WTAR)

All participants completed the Wechsler Test of Adult Reading (WTAR; Wechsler, 2001). This reading test involves asking the client to read out loud 50 words that have atypical grapheme to phoneme translations. The WTAR is a reliable and valid test that allows an initial estimation of pre-morbid intellectual and memory functioning for individuals aged 16 to 89 years. Normative data has been collected with both UK and US sample populations. Administration of the WTAR takes less than 10 minutes and the total score is the number of words read correctly.

Mini Mental State Exam

Older adults completed the Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975) which consists of 11 questions with scores ranging from 0 to 30. A score of 23 or less indicates cognitive impairment. The test takes 5 to 10 minutes to complete, and is designed to assess orientation, attention, language abilities, immediate and short term recall, as well as the ability to follow simple verbal commands. The MMSE was not given to young adults because this group was unlikely to present symptoms of cognitive decline.

3.6.4 Apparatus / Materials

3.6.4.1 Videotaped Event

The stimulus material consisted of a 130 second video recording of a neutral scene. The film opened with a general panorama of a residential area within the University of Strathclyde Campus, and then settled on a café / bar area. There were four main characters: a waiter, two women who conversed and a passerby who stopped to ask the women for directions. The incident contained information relating to objects, actions, descriptions and dialogue.

Overall the template consisted of 187 units of information which were further sub-divided into 128 units of person information, 35 units of surrounding information, and 24 units of dialogue information (The film summary template is included in Appendix E).

3.6.4.2 Event Memory Questionnaire (EMQ)

Normative data was obtained in a pilot study. Pilot study participants did not take part in the main experiment. Thirty questions were devised. Replicating the research of Ibabe and Sporer (2004), action details were defined as character non verbal behaviours. Descriptive details were defined as physical characteristics of scenes, persons and objects.12 questions concerned action details (i.e. visual, spatial; e.g., *How many chairs did the waiter re-position at the table?),* and 12 questions related to descriptive details (i.e. person, object; e.g., *How many pedestrians stood beside the parked cars?*). The remaining 6 questions related to dialogue information contained within the event (e.g., *What did Girl 2 order?*). Due to insufficient power, statistical analysis was restricted to responses in respect of action and description questions. The action, description and dialogue questions are reproduced in Appendix F (1-3).

Six versions of the EMQ were used in the experiment. Participants responded to either cued-recall questions or 4-Alternative Forced Choice (AFC) questions under conditions of high, moderate, or low incentive. The questions for the two tests were identical, but in the cued-recall version a blank line was provided below each question for recording the response, whereas in the recognition version the correct answer plus three foils were listed for selection. The foils were designed to be as plausible as possible. The correct answer was presented in a random order throughout the 30 items. The EMQ is reproduced in Appendix G (1 – 6).

3.7 Procedure

Prior to taking part in the experiment, all participants first gave verbal consent. Participants then read and signed a consent form. The researcher (a female in her late 40's) conducted all experiments. Participants were given verbal and written instructions for all tasks. Older adults required just slightly more explanation of the tasks, and the interviewer proceeded only when she was sure that participants understood the tasks.

Participants first completed the PANAS, MIA and MSEQ questionnaires then viewed the videotaped event. Participants were informed that they would view a non violent scene containing people, actions, objects and dialogue. They were instructed to watch carefully and pay attention to everything that happened as if they were a real eyewitness (Campos & Alonso-Quecuty, 1999). The researcher announced that she had not seen the film, explaining that this was an experimental control designed to prevent her from 'leaking' information about the video during the subsequent testing (Wright & Holliday, 2007). The researcher stood outside the testing room with the door closed while the video was showing. The video was operated via a Panasonic Video cassette recorder and shown on a 51cm Panasonic television monitor. Participants were seated approximately 4 feet from the television monitor. An immediate free recall was obtained and scored to ensure no differences in conditions prior to experimental manipulation. Delayed testing took place after 10 minutes during which time all participants completed the WTAR and Participant Details Questionnaire. Throughout the testing phases of the experiment, time limits were not imposed instead participants were advised to take as long as they needed to complete each task.

There were two experimental phases. At Phase 1 participants completed either a cued-recall or 4-AFC version of the 30-item EMQ under free report conditions. Participants were told that memory accuracy was the main interest in the study and accuracy was manipulated by one of three incentives: low, moderate or high incentive. In the low incentive condition, participants were instructed that they need not volunteer all the answers. The moderate and high incentive conditions included a pay-off schedule. Participants were advised that they could choose to give a response or pass on to the next item without being penalised or rewarded for omitted responses. In the moderate incentive condition one point would be awarded for a correct answer volunteered, but one point deducted for each incorrect answer volunteered. In the high incentive condition participants were awarded five points for a correct answer volunteered and one point would be deducted for each incorrect answer volunteered. High incentive participants were also advised that a payment of £25.00, payable to a charity of the participant's choice, would be made to the participant with the most accurate performance.

After completing Phase 1, participants started Phase 2. They completed the same 30-item EMQ under forced-report conditions. Participants were not reminded of the answers provided in the previous phase. Participants were asked to answer each item (guessing if necessary), and to indicate their level of confidence in the correctness of their answer on a 0 -100% scale (with 100% being absolutely certain). They were encouraged to use the entire scale range. Participants were also told that responding in this phase did not affect point totals because the point system was no longer operative. All participants then completed a second PANAS questionnaire and older adults completed the MMSE. The participants were then de-briefed and thanked for their participation. All tasks were self paced, and all testing was completed in one session, lasting between 90 and 120 minutes. Studies conducted by West, Bagwell, and King (2006) found no relationship between group size and memory performance or memory beliefs. Therefore, testing was completed either individually or in age segregated groups of up to 4 people. Groups of participants completed the same experimental condition. All data was collected independently. All instructions were printed on the experimental materials and were also read aloud to participants.

3.8 Scoring Procedure

3.8.1 Free Narrative

Recall was a secondary measure intended only to allow a basic comparison between the event-memory available to each group, prior to experimental manipulation. The participant free narrative reports were scored by two independent raters. The total number of correct and incorrect items of information were calculated. Two further scores were also calculated: (1) the proportion of correct details relative to the total number of details reported (i.e. percentage accuracy) and (2) the proportion of correctly recalled details relative to the total possible number of details that could have been recalled (i.e. quantity) (Dent, 1986; Milne, Clare, & Bull, 1999). Inter-rater reliability was calculated on the total number of correct and incorrect details for 36 participants: 18 (3 x Recall, low incentive, 3 x Recall, moderate incentive, 3 x Recall, high incentive, 3 x Recognition, low incentive, 3 x Recognition, moderate incentive, and 3 x Recognition, high incentive) from each of the younger and older adult age groups. The Pearson's correlations of the two raters' scores for the number of correct details was ($rho = 0.99$, $N = 36$, $p < 0.01$) and incorrect details was ($rho = 0.93$, $N = 36$, $p < 0.01$). The raters were provided with the free report narratives and coding frame. Otherwise they were experimentally blind.

3.8.2 Memory Accuracy and Quantity in Free Report

The 30-item EMQs were scored by two independent raters. The total number of *correct*, *incorrect* and *not answered* questions were recorded. Accuracy is scored as the number of correct responses divided by the number of responses offered at free report, when participants were permitted to leave items blank. Quantity is scored the proportion of correctly recalled details relative to the total possible number of details that could have been recalled (i.e. quantity) (Dent, 1986; Milne et al., 1999). Accuracy and quantity measures were also calculated for action and descriptive details. Inter-rater agreement was calculated on the total number of questions answered correctly, incorrectly and not answered, for 36 participants: – 18 (3 x Recall, low incentive, 3 x Recall, moderate incentive, 3 x Recall, high incentive, 3 x Recognition, low incentive, 3 x Recognition, moderate incentive, and 3 x Recognition, high incentive) from each of the younger and older adult age groups. The Pearson's correlations of the two raters' scores for the number of correct details was (*rho* = 0.99, $N = 36$, $p < 0.01$), incorrect details was (*rho* = 0.99, $N = 36$, $p < 0.01$), and questions not answered was (*rho* = 1.00, $N = 36$, $p < 0.01$).

3.8.3 Memory Accuracy and Quantity in Forced Report

The 30-item EMQs were scored by two independent raters. The total number *correct* and *incorrect* responses were recorded. Quantity and accuracy are equivalent under forced report conditions as participants must produce a response to every item; the likelihood of remembering each item is equal to the likelihood that each reported item is correct. Inter- rater agreement was calculated on the total number of questions answered correctly, and incorrectly, for 36 participants: – 18 (3 x Recall, low incentive, 3 x Recall, moderate incentive, 3 x Recall, high incentive, 3 x Recognition, low incentive, 3 x Recognition, moderate incentive, and 3 x Recognition, high incentive) from each of the younger and older adult age groups. The Pearson's correlations of the two raters' scores for the number of correct details was (*rho* = 0.98, $N = 36$, $p < 0.01$), and incorrect details was (*rho* = 0.96, $N = 36$, $p < 0.01$).

3.8.4 Monitoring Effectiveness

Both absolute correspondence (calibration) and relative correspondence (resolution) are evaluated. Calibration reflects the overall relation between accuracy and the level of the confidence ratings. In a calibration graph, confidence (x axis) is calculated separately for different confidence levels and is then plotted against accuracy (y axis). The calibration graph indicates over / under confidence (i.e. the extent to which the confidence judgments are higher or lower than the actual proportions correct). The calibration curves based on the forced report performance, by age group and test format for action, and descriptive details are presented in *Figures 4* to *7* in the Results Section. Monitoring resolution is the degree to which assessed probabilities of correctness successfully differentiate correct and incorrect candidate answers, and is measured by within-subject gamma correlations (Nelson, 1984).

Control Sensitivity: Subjective Confidence and the decision to respond.

Control processes influence accuracy by determining whether a response is volunteered or withheld. According to the model, this decision is based on (a) the output of the monitoring mechanism and (b) the incentive for accuracy.

Response criterion setting

Response criterion setting is the minimum level of confidence that is required by the participant before he / she is willing to volunteer an answer. This can be estimated by finding a cutoff on the confidence ratings that best separates the items volunteered from those that were withheld at free report. Response criterion estimates were calculated as the mean probability of the items that were volunteered and withheld at Phase 1, and the proportion of items volunteered at free report (Pansky et al., 2009).

3.9 Results

Analysis can be broadly categorised into analysis of free recall, individual difference measures, memory performance at free report (accuracy and quantity), and on four component measures: 1) memory retention at forced report, 2) monitoring effectiveness (the ability to discriminate between correct and incorrect answers), 3) control sensitivity (the extent to which a volunteered response is based on the monitoring output), and 4) free-report volunteering rate as an indirect measure of control policy. Finally path models of accuracy show the direct and indirect effects of the predictor variables on free-report accuracy. Each of these components of the data analysis will be considered in turn. Recall was a secondary measure intended only to allow a basic comparison between the event-memory available to each group, prior to experimental manipulation. Analysis revealed no significant differences between conditions. An individual differences analysis was used to assess whether variability in participant education, crystallised intelligence, health, sensory deficits, subjective memory beliefs, and positive and negative affect mediated the effect of age and objective memory performance, either directly or indirectly within a general theoretical model which specifies the mediating role of metacognitive processes. Individual difference measures are summarised in *Table 1* below, as well as the results of *t*-tests concerning differences between the samples. Pre-and post-test measures of positive and negative affect were analysed by Analysis of Variance (ANOVA). Memory performance (accuracy and quantity) first at free report and then at forced report was then investigated. Analysis of the monitoring processes followed in terms of how well calibrated people's confidence judgments were, and how well they discriminated correct from incorrect potential responses as indicated by measures of monitoring resolution. The memory control processes, in terms of both the relation between confidence and the decision to report or withhold a candidate response and the setting of participant response criteria were investigated.

Based on Koriat and Goldsmith's (1996) model, Experiment 1 was designed to compare the monitoring effectiveness (i.e. how well people calibrate their confidence judgments, and how well they can discriminate correct from incorrect potential responses as indicated by measures of monitoring resolution), the control sensitivity (i.e. the relation between confidence and the decision to report or withhold a potential response), and the response criterion setting for action and descriptive details within a to-be-remembered event. Participants responded to either cued-recall or multiple choice recognition tests, under conditions of high, moderate or low incentive. Memory performance (accuracy and quantity) was measured first under free-report followed by forced-report instructions.

The means and standard deviations (SD) of overall memory performance measures are provided in *Table 4*. Memory accuracy was calculated as the proportion (%) of correct responses provided out of the number of responses offered. Memory quantity was calculated as the number of correct responses as a proportion (%) of the total number of responses available. Memory quantity and accuracy are equivalent in the forced report stage (Phase 2) as participants were required to produce a response to every item. The memory monitoring data are provided in *Table 5*.

Finally, the current study also examined whether individual difference measures provide a unique variance as a predictor of free report accuracy performance beyond age, and memory monitoring and control processes. This question was considered separately in respect of action and descriptive details using Baron and Kenny's (1986) test for mediator effects. The path models are provided in *Figures 10* and *11*.

Kolmogorov-Statistical tests of normality were computed on the distribution of scores on all performance measures prior to analysis. Unless otherwise noted, the assumptions to allow statistical analysis with parametric statistics, ANOVA were met. Main effects analysis included a *Bonferroni* correction. In view of the large number of planned comparisons, unless otherwise noted, the value of α for ANOVA statistical tests was set at (0.01), (Sheskin, 2007).

3.9.1 Individual Difference Measures

Table 1: Means and (SD) of individual difference measures of younger and older adults

Variable	Younger Adults	Older Adults	*p*
N	102	95	
Education	15.56 (2.75)	14.37 (3.06)	($p < 0.01$)
WTAR[a]	35.74 (10.33)	44.46 (4.32)	($p < 0.01$)
Physical Health	1.97 (0.87)	2.19 (0.90)	*ns*
Mental Health	1.69 (0.77)	1.97 (1.05)	($p < 0.05$)
Vision	2.28 (1.24)	2.7. (0.94)	($p = 0.01$)
Hearing	1.98 (0.90)	2.76 (1.05)	($p < 0.01$)
MIA[b]			
Strategy	62.72 (9.52)	63.71 (9.85)	*ns*
Task	63.82 (6.38)	62.39 (5.36)	*ns*
Capacity	59.28 (9.28)	54.89 (9.09)	($p < 0.01$)
Change	63.07 (10.16)	48.47 (11.59)	($p < 0.01$)
Anxiety	45.69 (9.98)	45.19 (8.78)	*ns*
Achievement	62.12 (8.56)	59.25 (6.78)	($p = 0.01$)
Locus	30.18 (5.40)	28.84 (5.51)	*ns*
MSEQ Sel[c]	67.71 (21.27)	62.32 (21.93)	*ns*
MSEQ Sest[d]	70.25 (14.04)	61.98 (16.88)	($p < 0.01$)
Positive Affect (1)	31.39 (6.28)	33.12 (6.89)	
Positive Affect (2)	31.95 (8.96)	33.04 (8.13)	
Negative Affect (1)	14.42 (4.86)	11.45 (2.98)	
Negative Affect (2)	15.33 (4.44)	13.83 (3.70)	
MMSE[e]		29.59 (0.65)	

[a] Wechsler Test of Adult Reading (Wechsler, 2001) score represents the total number of correctly pronounced irregular words out of a total of 50 words. [b] Metamemory in Adulthood Questionnaire (MIA; Dixon et al., 1988), scales: Strategy, Task, Capacity, Change, Anxiety, Achievement, Locus. [c] Memory Self Efficacy Questionnaire, Self-efficacy Level (MSEQ; Berry et al., 1989). [d] Memory Self Efficacy Questionnaire, Self-efficacy strength (MSEQ; Berry et al., 1989).[e] Mini Mental State Examination (MMSE; Folstein et al., 1975) score represents the number of points earned out of a possible total of 30.

Initial analysis by independent samples *t*-tests were carried out on the following measures: years of education, WTAR, physical and mental health, vision, hearing and the MIA scales of strategy, task, capacity, change, anxiety, achievement and locus, and MSEQ level and strength. The means and standard deviations (SD) of younger and older adults are provided in *Table 1* above. Significant age effects were observed in the individual difference measures of education (t (193) = 2.85, p < 0.01), WTAR (t (193) = 6.91, p < 0.001), mental health (t (193) = 2.15, p < 0.05), vision (t (193) = 2.61, p = 0.01), and hearing (t (193) = 5.61, p < 0.01), Older adults attained fewer years of education, correctly pronounced more irregular words, and reported poorer mental health, vision and hearing. The analysis further revealed that significant age effects appeared in the MIA scales of capacity (t (193) = 3.32, p = 0.001), change (t (193) = 9.34, p < 0.001), achievement (t (193) = 2.58, p = 0.01), and MSEQ strength (t (193) = 3.74, p < 0.001), Younger adults were significantly more positive about their memory. To determine whether the afore-mentioned individual difference variables were to be included as covariates in the further analysis, partial correlations were computed between the individual difference measures and measures of objective memory performance. The correlation matrix is included in *Table 2* below. The correlations were computed after partialling out age.

Table 2: Correlation matrix between individual difference measures and objective memory performance (accuracy and quantity) at free and forced report

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Education	1.0	-.17	.01	.08	-.04	.03	.04	.00	.06	.01	.06	-.02	.10	-.05	-.18
2. WTAR		1.0	.01	.02	.11	-.13	-.11	-.22	-.13	.13	-.02	.11	-.11	.15	.11
3. Mental Health			1.0	.14	**..28**	**-.20**	**-.27**	**.22**	-.12	-.01	.07	-.04	.14	.01	.17
4. Vision				1.0	**.33**	-.03	-.03	.07	-.15	.05	-.03	.05	.07	.06	.01
5. Hearing					1.0	-.14	-.10	.08	-.16	-.02	-.11	-.01	-.07	.01	-.08
6. Capacity						1.0	**.53**	-.00	**.35**	-.04	-.04	.03	-.04	.01	-.07
7. Change							1.0	-.06	**.20**	-.17	.01	-.07	-.03	-.04	-.04
8. Achievement								1.0	.03	-.09	-.02	-.08	.06	-.01	.03
9. SEST									1.0	-.05	-.04	.02	-.10	-.06	.00
10 FR Action Accy										1.0	-.02	**.61**	-.05	**.61**	-.08
11 FR Desc Accy											1.0	-.11	**.64**	-.13	-.06
12. FR Action Qnty												1.0	-.11	**.68**	.09
13. FR Desc Qnty													1.0	-.16	-.01
14.Forced Action														1.0	.06
15 Forced Description															1.0

All correlations in bold type-face were significant at $p < 0.01$ or better

1. Years of Education attained; 2. Wechsler Test of Adult Reading (WTAR; Wechsler, 2001) ; 3. Mental Health: Self-report measure on 5-point scale; 4. Vision: Self-report measure on 5-point scale; 5. Hearing: Self-report measure on 5-point scale; 6. Metamemory in Adulthood Questionnaire (MIA; Dixon et al., 1988) – Capacity; 7. Change; 8. Achievement; 9. Memory Self Efficacy Questionnaire, (MSEQ; Berry et al., 1989) - Self-efficacy strength; 10. Free Report Action Details Accuracy; 11. Free Report Description Details Accuracy; 12. Free Report Action Details Quantity; 13. Free Report Description Details Quantity; 14. Forced Report Action Details; 15 Forced Report Description Details Accuracy.

Self-report measures of mental health correlated with different variables. Mental health measures were significantly positively correlated with self-report measures of hearing ($r = .28$, $p < 0.001$), and the MIA scale of achievement ($r = .22$, $p < 0.001$). Self-report measures of mental health were also significantly negatively correlated with MIA scales of capacity ($r = -.20$, $p = 0.005$) and change ($r = -.27$, $p < 0.001$). Memory self-efficacy strength was positively correlated with the MIA subscales of capacity ($r = .35$, $p < 0.001$) and change ($r = .20$, $p = 0.005$) showing that overall memory self-efficacy was associated with stable memory and a high perceived sense of control over memory skills. Correlations between the individual difference variables and the memory performance variables did not reach significance at $p < 0.01$ or better.

Is objective memory performance affected by positive and negative affect?

Pre- and post-test measures were taken of both positive and negative affect. The data are summarised in *Table 1* above. An age group x positive affect mixed model ANOVA was carried out on the pre- and post-test positive affect scores. The results revealed no significant main effects of age group, positive affect, or interaction ($p > 0.05$).

The same analysis was carried out on the pre- and post-test negative affect scores. Box's Test of Equality of Covariances was significant ($p < 0.001$). Levene's Test of Equality of Error Variances was significant in respect of both pre- and post-test negative affect scores ($p < 0.01$). The value of α for this analysis was therefore set at (0.001), (Stevens, 1996). The main between-subjects effect of age group was significant (F (df, 1, 193) = 17.67, MS Error = 27.56, $p < 0.001$, $\eta^2 = .08$). Younger adults reported higher levels of negative affect (M = 14.88) compared with older adults (M = 12.64). The main effect of negative affect was also significant (F (df, 1, 193) = 44.68, MS Error = 5.89, $p < 0.001$, $\eta^2 = .19$). The interaction between age group and negative affect was not significant (F (df, 1, 193) = 8.86, MS Error = 5.89, $p > 0.001$, $\eta^2 = .04$).

To determine whether the variables of pre- and post-test positive affect and / or negative affect were to be included as covariates in the further analysis partial correlations were computed on all measures of memory performance (accuracy and quantity) at free and forced report. The correlation was computed after partialling out age. The correlation matrix is included in *Table 3* below.

Table 3: Correlation matrix between pre and post-test measures of positive and negative affect and objective memory performance (accuracy and quantity) at free and forced report

	1	2	3	4	5	6	7	8	9	10
1. Positive Affect (1)	1.0	**.72**	.11	.05	-.02	-.07	.06	-.00	.03	.06
2. Positive Affect (2)		*1.0*	.11	-.06	.01	-.10	.08	.01	.03	.07
3. Negative Affect (1)			*1.0*	**.65**	-.05	-.04	-.02	-.05	-.09	-.05
4. Negative Affect (2)				*1.0*	-.14	-.06	-.10	-.07	-.09	-.00
5. FR Accuracy – Action					*1.0*	-.02	**.61**	-.06	**.61**	-.08
6. FR Accuracy – Description						*1.0*	-.11	**.65**	-.12	.05
7. FR Quantity – Action							*1.0*	-.11	**.68**	.08
8. FR Quantity – Description								*1.0*	-.15	.01
9. Forced Accuracy – Action									*1.0*	.06
10. Forced Accuracy – Description										*1.0*

All correlations in bold type-face were significant at $p < 0.01$ or better.

1. Pre-test Positive Affect; 2. Post-test Positive Affect (PANAS; Watson et al., 1988); 3. Pre-test Negative Affect; 4. Post-test Negative Affect (PANAS; Watson et al., 1988); 5. Free Report Accuracy – Action; 6. Free Report Accuracy – Description; 7. Free Report Quantity – Action; 8. Free Report Quantity – Description; 9. Forced Report Accuracy – Action; 10. Forced Report Accuracy – Description.

Self-report measures of pre-test positive affect were positively correlated with post-test positive affect ($r = .72$, $p < 0.001$). Pre-test negative affect were positively correlated with post-test measures of negative affect ($r = .65$, $p < 0.001$). Correlations between the individual difference variables of pre- and post-test positive and negative affect and the memory performance variables did not reach significance at $p < 0.01$ or better.

3.9.1.1 Summary of Preliminary Analysis

To summarise the preliminary analysis, older adults attained fewer years of education, correctly pronounced more irregular words, and reported poorer mental health, vision and hearing. Significant age effects appeared in the MIA scales of capacity, change, achievement, and MSEQ strength. Younger adults reported significantly more positive beliefs about their memory. Analysis further revealed that correlations between the aforementioned individual difference variables and the outcome performance variables did not reach significance at $p < 0.01$ or better and therefore were not included as covariates in the between-subjects further analysis.

3.9.2 Memory accuracy in Free Report

Do age, test format and content affect memory accuracy in free report?

The next analysis focuses on the outcome of the retrieval, monitoring, and control processes by looking at memory accuracy in free report. Tests of normality were carried out on memory performance by content (action, description). The Kolmogorov-Statistics of both young and older adult recall of description details were significant ($p < 0.05$), suggesting a violation of the ANOVA assumption that observations are normally distributed on the dependent variable in each group. No observations with scores in excess of 3.29 standard deviations from the mean ($p < 0.001$) were recorded. The consequences of violating this assumption were reviewed by Glass, Peckham and Saunders (1972) who found that both skewness and kurtosis have slight effects on the level of significance or power. Stevens (1996), suggests that this is due to the *Central Limit Theorem,* which states that as the number of observations increase, the sum of independent observations approaches a normal distribution. Furthermore, Bock (1975, p.111) noted that "even for distributions which depart markedly from normality, sums of 50 or more observations approximate to normality". In the present study the observations of younger and older adults numbered 102 and 93 respectively, and as no extreme scores were identified, the distribution of all observations at free report were considered normal and no further action was necessary (Tabachnick & Fidell, 2001).

Accuracy was scored as the number of correct responses divided by the number of responses offered at free report. The means and standard deviations (SD) of memory performance (accuracy and quantity) at free and forced report are provided in *Table 4* below.

Table 4: Means and standard deviations (SD) of quantity and accuracy scores at free and forced report by age group, test format and content

Age Group	Test Format	Content	Report Option		
			Free Report		Forced Report
			Quantity	Accuracy	Quantity
			Mean (SD)	Mean (SD)	Mean (SD)
Younger	Recall	Action	.54 (.14)	.71 (.16)	.63 (.14)
		Description	.52 (.16)	.72 (.18)	.61 (.14)
	Recognition	Action	.57 (.16)	.69 (.13)	.65 (.12)
		Description	.56 (.14)	.75 (.16)	.75 (.13)
Older	Recall	Action	.47 (.13)	.67 (.17)	.57 (.10)
		Description	.58 (.14)	.82 (.14)	.50 (.14)
	Recognition	Action	.39 (.17)	.52 (.17)	.50 (.14)
		Description	.61 (.16)	.79 (.16)	.67 (.14)

An age group (younger, older adults) x test format (cued-recall, recognition questions) x incentive (high, moderate, low) x content (action, description) mixed model ANOVA, with content being the within-subjects measure was carried out. The results revealed no significant main effects or interactions arising from the incentive conditions ($p > 0.05$) therefore the data was collapsed over the remaining groups. An age group x test format x content mixed model ANOVA was carried out.

Age group

The main between-subjects effect of age group was not significant ($p > 0.01$). A significant interaction effect between age group and test format was observed (F (df, 1, 191) = 10.22, MS Error = .03, $p < 0.005$, $\eta^2 = .05$). Analysis of main effect means by Tukey HSD *post hoc* comparisons for unequal sample sizes revealed greater older adult accuracy in response to cued-recall questions. Younger adult performance was greater in response to recognition questions. The interaction between age group and content was also significant, (F (df 1, 191), = 29.21, MS Error = .03, $p < 0.001$, $\eta^2 = 0.13$). Planned comparisons showed that the difference in accuracy when reporting descriptive (M = .81) compared with action (M = .60) details was significant in the recall of older adults (t (92) = -8.93, $p < 0.001$), while not significant in the recall of younger adults (t (101) = -1.27, $p > 0.01$). The triple interaction between age group, test format and content was not significant ($p > 0.01$).

Test Format

The main between-subjects effect of test format was significant (F (df, 1, 191) = 6.96, MS Error = .03, $p < 0.01$, $\eta^2 = .04$). Accuracy was greater in response to cued-recall compared with recognition questions.

Content

The main within-subjects effect of content was significant (F (df, 1, 191) = 53.16, MS Error = .03, $p < 0.001$, $\eta^2 = .22$). Accuracy was greater when participants reported descriptive compared with action details.

3.9.3 Memory Accuracy at Forced Report

Do age, test format and content affect retrieval at forced report?

The analysis next turns to how well retrieval operated as shown by performance under forced report. Quantity and accuracy are equivalent in the forced report stage as participants must produce a response to every item. The means and standard deviations (SD) of memory performance (accuracy and quantity) at free and forced report are provided in *Table 4* above.

Tests of normality were carried out on memory retrieval by content (action, description). The Kolmogorov-Statistics of both young and older adults in respect of all measures were significant ($p < 0.05$), suggesting a violation of the ANOVA assumption that observations are normally distributed on the dependent variable in each group. No observations with scores in excess of 3.29 standard deviations from the mean ($p < 0.001$) were recorded. As discussed on page 115 the distribution of all observations were considered normal and no further action was necessary (Tabachnick & Fidell, 2001).

An age group x test format x incentive x content (action, description) mixed model ANOVA, with content being the within-subjects measure was carried out. The results revealed no significant main effects or interactions arising from the incentive conditions ($p > 0.05$) therefore the data was collapsed over the remaining groups. An age group x test format x content (action, description) mixed model ANOVA was carried out. Levene's Test of Equality of Error Variances in respect of accuracy when reporting action details was significant ($p < 0.5$). The value of α for this analysis was therefore set at (0.001), (Stevens, 1996).

Age group

The main between-subjects effect of age group was significant (F (df, 1, 191) = 46.41, MS Error = .02, p < 0.001, η^2 = .20). Younger adult accuracy (M = .66, SE .01) was significantly greater when compared with older adult accuracy (M = .56, SE .01). The interaction between age group and test format was not significant (p = 0.31). The interaction between age group, test format and content was not significant (p = 0.01, η^2 = 0.03).

Test Format

The main between-subjects effect of test format was significant, (F (df 1, 191), = 23.17, MS Error = .02, p < 0.001, η^2 = 0.11). Retrieval at forced report was significantly greater in response to recognition (M = .64, SE .01) when compared with cued-recall questions (M = .58, SE .01).

Content

The main within-subjects effect of content was significant (F (df, 1, 191) = 12.06, MS Error = .02, p < 0.002, η^2 = .06). Retrieval was significantly greater in response to descriptive (M = .63, SE .01) compared with action details (M = .59, SE .01).

3.9.4 Gains in Accuracy from Forced to Free Report

Do age, test format and content affect the gains in memory accuracy from forced to free report? Were participants able to use monitoring and control processes to improve their accuracy between forced and free report? The means and standard deviations (SD) of memory performance (accuracy and quantity) at free and forced report are provided in *Table 4* above.

An age group x test format x incentive x report option x content (action, description) mixed model ANOVA, with report option and content the within-subjects factors, was carried out on memory accuracy scores at free and forced report. The results revealed no significant main effects or interactions arising from the incentive conditions ($p > 0.05$) therefore the data was collapsed over the remaining groups. An age group x test format x report option x content mixed model ANOVA was carried out. Levene's Test of Equality of Error Variances was significant ($p < 0.05$). The value of α for the between-subjects analysis was therefore set at (0.001), (Stevens, 1996).

Age group

The results showed a significant between-subjects main effect of age group (F (df, 1, 191) = 23.49, MS Error = .03, $p < 0.001$, $\eta^2 = .11$). Younger adult accuracy was significantly greater compared with older adult accuracy. An interaction between age group and test format approached significance (F (df, 1, 191) = 8.31, MS Error = .03, $p = 0.004$, $\eta^2 = .04$). The younger adult accuracy in response to recognition questions was greater (M = .71, SE .01) compared with cued-recall questions (M = .67, SE .01). The opposite pattern of results was observed in the recall of older adults. Performance in response to cued-recall questions (M = .64, SE .01) was significantly greater when compared with recognition questions (M = .62, SE .01). The interaction effect between age group and the within-subjects factor of report option was significant (F (df, 1, 191) = 16.37, MS Error = .02, $p < 0.001$, $\eta^2 = .08$). Significantly greater gains from forced to free report were realised by older adults (M = 1.00, t (92) = 11.01, $p < 0.001$) compared with younger adults (M = .09, t (101) = 11.59, $p < 0.001$). The interaction between age group, test format and report option also approached significance (F (df, 1, 191) = 4.59, MS Error = .02, p = 0.03, η^2 = .02). The interaction between age group and the within-subjects factor of content was significant, (F (df 1, 191), = 17.71, MS Error = .02, p < 0.001, η^2 = 0.09). Analysis of the main effect means by Tukey HSD *post hoc* comparisons for unequal sample sizes, revealed that young adult accuracy was significantly greater (p < 0.01) when reporting descriptive (M = .67) compared with action details (M = .71). A similar pattern of results was obtained in older adult accuracy (p < 0.01), although the difference between descriptive (M = .70) compared with action details (M = .56) was more pronounced. The triple interaction between age group, report option and content was also significant (F (df 1, 191), = 23.01, MS Error = .02, p < 0.001, η^2 = 0.11). Younger (M = .06, t (101) = 5.44, p < 0.001), and older adults (M = .06, t (92) = 3.73, p < 0.001) realised significant gains from forced to free report when reporting action details. Significant gains in descriptive details were also realised by younger (M = .06, t (101) = 2.52, p = 0.01), and older adults (M = .22, t (92) = 9.22, p < 0.001), although the gain was more pronounced in the older adult performance (see *Figure 2* below for the interaction). The interaction between age group, test format and content approached significance (F (df 1, 191), = 5.99, MS Error = .02, p = 0.02, η^2 = 0.03). The interaction between age group, test format, report option, and content was not significant (p > 0.01).

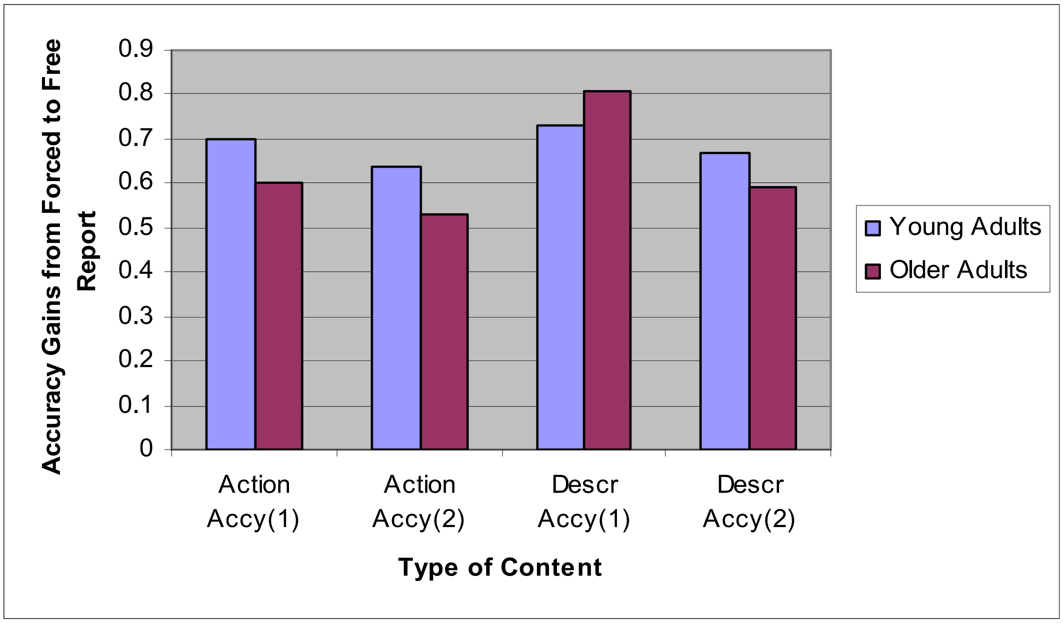


Figure 2: Interaction between age group, report option, and content in respect of memory accuracy gains from forced to free report.

Test Format

The main between-subjects effect of test format was not significant ($p > 0.001$).

Report Option

The within-subjects main effect of report option was significant (F (df, 1, 191) = 48.57, MS Error = .02, $p < 0.001$, $\eta^2 = .20$). All participants were able to significantly improve the accuracy of their memory accounts from forced to free report.

Content

The main effect of content was significant (F (df, 1, 191) = 108.40, MS Error = .02, $p < 0.001$, $\eta^2 = .36$). Accuracy was significantly greater when reporting descriptive compared with action details.

3.9.5 Losses in Quantity from Forced to Free Report

Do age, test format and content affect the losses in memory quantity from forced to free report? If participants cannot perfectly distinguish which of their candidate responses in forced report are correct and which are incorrect, then increases in accuracy in free report may come at the expense of quantity correct. Quantity and accuracy are equivalent in the forced report stage as participants must produce a response to every item. The means and standard deviations (SD) of memory performance (accuracy and quantity) at free and forced report are provided in *Table 4* above.

Tests of normality were carried out on measures of content (action, description). At free report, the Kolmogorov-Statistics of both young and older adults on all measures were significant ($p < 0.05$). No observations with scores in excess of 3.29 standard deviations from the mean ($p < 0.001$) were recorded, therefore as discussed on page 115 the distribution of all observations were considered to be normal (Tabachnick & Fidell, 2001).

In comparing overall memory performance (quantity) under free and forced instructions, an age group x test format x incentive x report option x content (action, description) mixed model ANOVA, with report option and content being the within-subjects measures, was carried out on memory quantity scores at free and forced report. The results revealed no significant main effects or interactions arising from the incentive conditions ($p > 0.05$) therefore the data was collapsed over the remaining groups. An age group x test format x report option x content mixed model ANOVA was carried out. Levene's Test of Equality of Error Variances in respect of action details at forced report was significant ($p < 0.05$). The value of α for the between-subjects analysis was therefore set at (0.001), (Stevens, 1996).

Age group

The results showed a significant between-subjects main effect of age group (F (df, 1, 191) = 34.14, MS Error = .03, $p < 0.001$, $\eta^2 = .15$). Younger adults retrieved a significantly greater number of correct responses compared with older adults. The interaction effect between age group and report option was significant (F (df, 1, 191) = 17.16, MS Error = .02, $p < 0.001$, $\eta^2 = .08$). Greater losses from forced to free report were realised by older adults (M = .09, t (92) = 11.63, $p < 0.001$) compared with younger adults (M = .06, t (101) = 9.72, $p < 0.001$). The triple interaction between age group, test format and report option was not significant ($p = 0.02$, $\eta^2 = 0.03$). The interaction between age group and content was significant, (F (df 1, 191), = 12.39, MS Error = .02, $p = 0.001$, $\eta^2 = 0.06$). Analysis of main effects by Tukey HSD *post hoc* comparisons for unequal sample sizes revealed that both young and older adults reported a significantly greater number of correct action details ($p < 0.01$). The difference was greater in the younger adult performance. The triple interaction between age group, report option and content was also significant (F (df 1, 191), = 23.01, MS Error = .02, $p < 0.001$, $\eta^2 = 0.11$). From forced to free report, younger adults lost a greater number of correct descriptive (M = 13, (t (101) = 6.37, p < 0.001) compared with action details (M = .09) (t (101) = 7.50, $p < 0.001$). Older adults lost a greater number of action (M = .11)

(t (92) = 10.04, $p < 0.001$), compared with descriptive details (M = .01, (t (92) = .35, $p <$ 0.01) (see *Figure 3* below for the interaction). The interaction effect between age group and test format approached significance ($p = 0.04$, $\eta^2 = 0.02$). The triple interaction between age group, test format and content was not significant ($p > 0.05$, $\eta^2 = 0.01$). The four-way interaction between age group, test format, report option, and content, was not significant ($p > 0.05$, $\eta^2 = 0.00$).

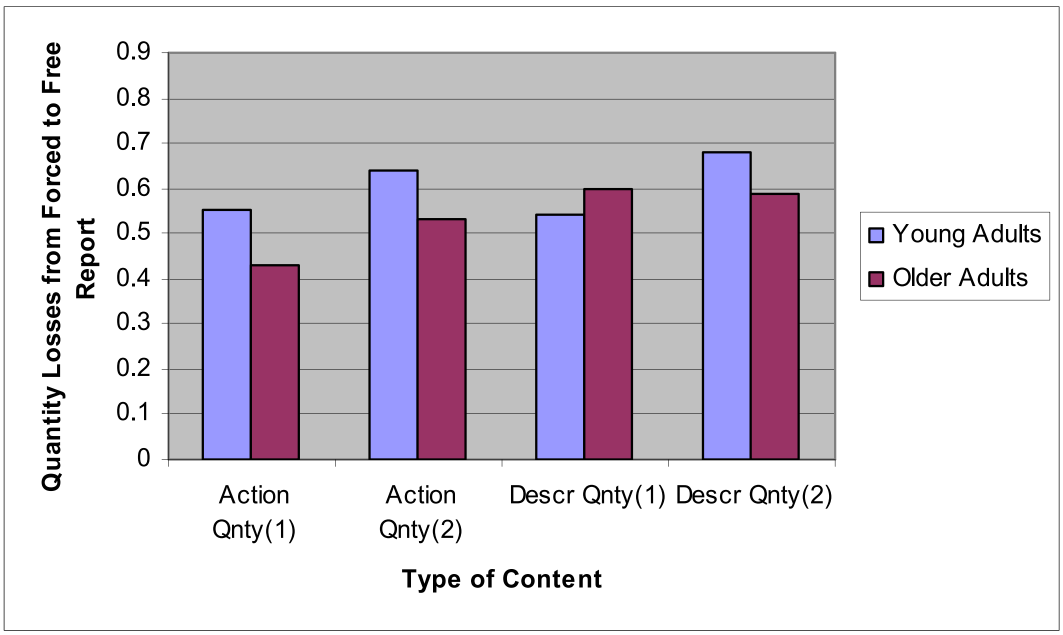


Figure 3: Interaction between age group, report option and content in respect of quantity losses from forced to free report

Test Format

The main effect of test format was significant (F (df, 1, 191) = 10.65, MS Error = .03, p = 0.001, η^2 = .05). Recognition questions generated a significantly greater number of correct responses compared with cued-recall questions.

Report Option

The main effect of report option was significant (F (df, 1, 191) = 27.06, MS Error = .02, $p < 0.001$, $\eta^2 = .12$). Exercising the option of free report led to significant gains in accuracy from forced to free report with a trade-off in quantity lost.

Content

The main effect of content was significant (F (df, 1, 191) = 85.10, MS Error = .02, $p < 0.001$, $\eta^2 = .31$). A significantly greater number of correct action compared with correct descriptive details were reported.

3.9.6 Summary of Memory Performance

Participants demonstrated considerable gains in accuracy from forced (.62) to free report (.70), at the expense of quantity correct from (.60) to (.54). When reporting descriptive items, younger adults showed a gain of 0.06 percentage points in accuracy from forced to free report, with a 0.01 percentage point loss in quantity. When reporting action items, they showed a gain of 0.06 percentage points, with a loss of 0.04 in quantity. The older adults achieved a substantial .21 percentage point gain in accuracy when reporting descriptive items, with a .11 percent point loss on quantity. Older adult gains were smaller when reporting action items, only .5 percentage points, with a loss of .5 percentage points in quantity.

At free report, the between-subjects analysis of memory accuracy revealed no main effect of age. Within-subjects analysis of content revealed a significant age x content interaction. Older adult memory accuracy was significantly greater when reporting descriptive compared with action details. No significant differences were observed in the younger adult content performance. At forced report between-subjects analysis revealed a main effect of age. Older adults retrieved significantly fewer correct items when compared with younger adults. No significant age x content interactions were observed. Within-subjects analysis of the gains in memory accuracy performance from forced to free report revealed that the older group gained significantly more than younger adults from the opportunity to screen out potential incorrect responses. Significantly greater gains were realised in respect of descriptive details (see *Figure 2* above for the interaction). Gains in older adult accuracy came at the expense of greater losses in quantity correct. Older adults lost a greater number of correct action details compared with descriptive details. Younger adults lost a greater number of correct descriptive compared with action details (see *Figure 3* above for the interaction). Gains in accuracy from forced to free report may reflect the efficacy of monitoring and control processes which will be examined in the next section.

3.9.7 Memory monitoring and the likelihood of being correct

The confidence rating on each item obtained immediately following forced report is equivalent to the assessed probability that the response is a memory, which is then assumed to be the basis for the decision to volunteer or withhold it. On 1% of the items, participants changed their response between free and forced report, and those items are omitted from the following analysis. Participants provided an immediate confidence rating for each item given during forced report. Confidence ratings were grouped into 12 levels (e.g. 0, .01 - .10, .11 - .20, .21 - .30, ...91 - .99, 1.0). As the incentive condition was applied only to the free recall performance, the data for the probability judgements was collapsed across the other groups. Using these confidence ratings, memory monitoring effectiveness was quantified in terms of calibration and monitoring resolution (see *Table 5 below*). Two indices of monitoring effectiveness were calculated. Both absolute correspondence (calibration) and relative correspondence (resolution) were evaluated. Calibration reflects the overall relation between accuracy and the level of the confidence ratings. In a calibration graph, confidence (*x* axis) is calculated separately for different confidence levels and is then plotted against accuracy (*y* axis). The calibration graph indicates over / under confidence (i.e. the extent to which the confidence judgments are higher or lower than the actual proportions correct). Monitoring resolution is the degree to which assessed probabilities of correctness successfully differentiate correct and incorrect candidate answers, and was measured by within-subject Kruskal-Goodman gamma correlations (Nelson, 1984).

Table 5: Memory monitoring data (and SD) by age group, test format and content

Measure	Test Format	Content	Younger Adults	Older Adults
Calibration Error	C-Recall	Action	.32 (.18)	.38 (.17)
		Description	.41 (.19)	.34 (.19)
	Recognition	Action	.36 (.20)	.37 (.17)
		Description	.43 (.21)	.40 (.16)
y – Correct[a]	C-Recall	Action	.61 (.28)	.60 (.25)
		Description	.66 (.29)	.63 (.27)
	Recognition	Action	.52 (.32)	.42 (.27)
		Description	.75 (.28)	.68 (.27)
y – Response[b]	C-Recall	Action	.89 (.15)	.82 (.23)
		Description	.92 (.19)	.92 (.13)
	Recognition	Action	.88 (.20)	.88 (.16)
		Description	.94 (.15)	.89 (.20)
Response Vol. Rate[c]	C-Recall	Action	.77 (.16)	.72 (.20)
		Description	.74 (.17)	.71 (.14)
	Recognition	Action	.83 (.16)	.75 (.20)
		Description	.76 (.13)	.77 (.13)
P_{RC}[d]	C-Recall	Action	.75 (.13)	.71 (.18)
		Description	.81 (.12)	.82 (.12)
	Recognition	Action	.63 (.15)	.69 (.16)
		Description	.66 (.17)	.74 (.15)

[a] Monitoring Resolution; [b] Relationship between confidence and decision to respond;
[c] Proportion of potential responses volunteered under free report; [d] Response Criterion

3.9.7.1 Calibration

The calibration curves based on the forced report performance, by age group and test format in respect of action, and descriptive details are presented in *Figures 4* to *7* below. The proportion correct (% accuracy) is plotted against the mean assessed probability across participants; the diagonal line with an intercept of .0 and a slope of 1.0 indicates perfect calibration.

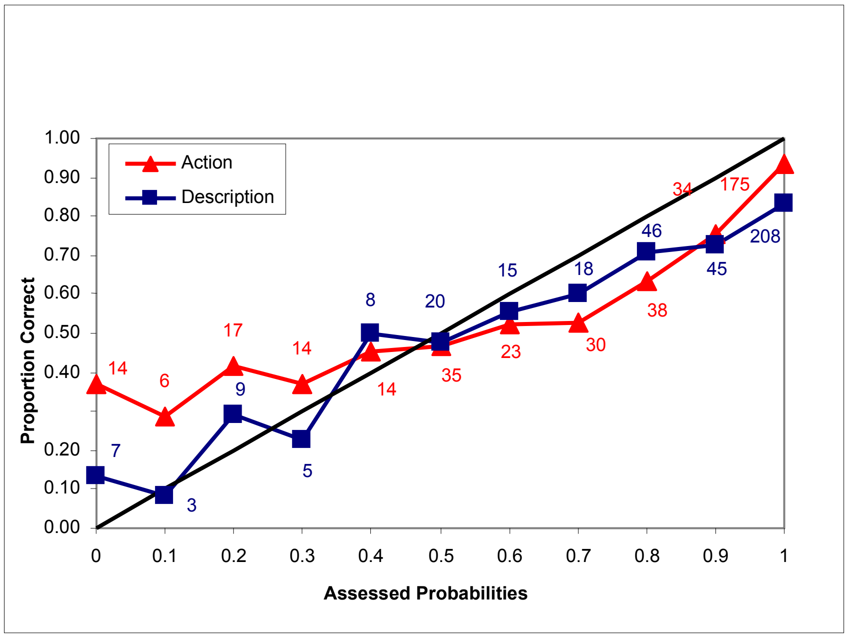


Figure 4: Calibration curves in respect of action and descriptive items by young cued-recall participants

Diagonal line represents perfect calibration

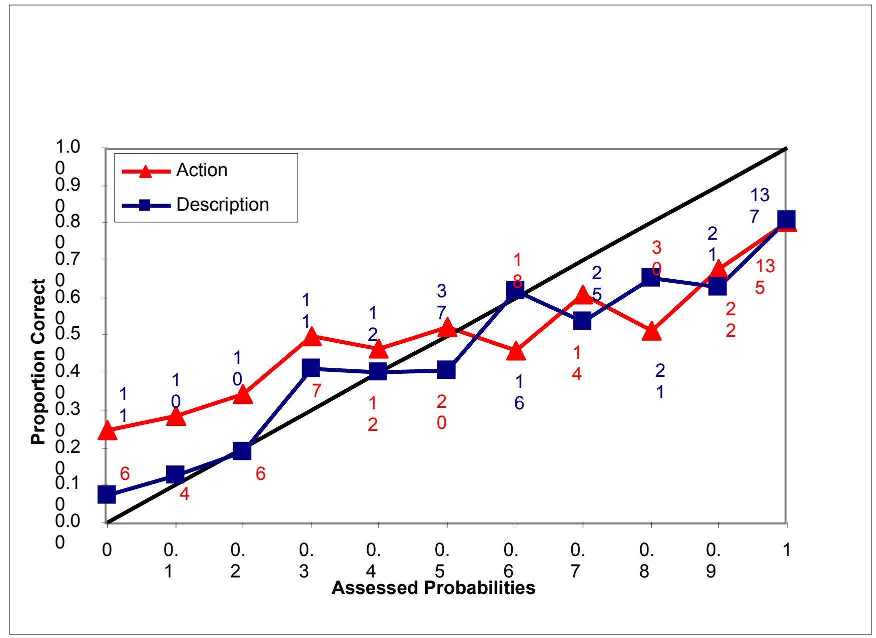


Figure 5: Calibration curves in respect of action and descriptive items by older cued-recall participants

Diagonal line represents perfect calibration

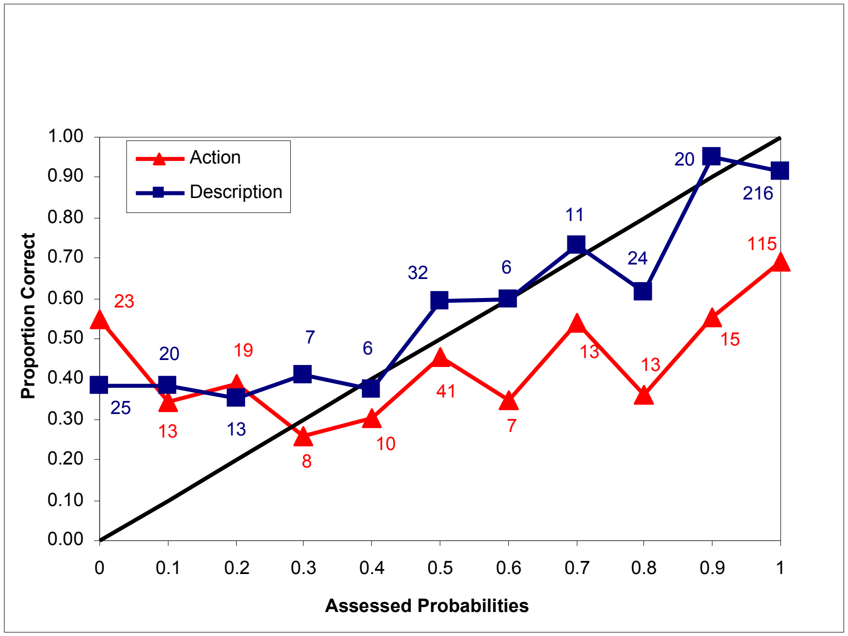


Figure 6: Calibration curves in respect of action and descriptive items by young recognition participants

Diagonal line represents perfect calibration

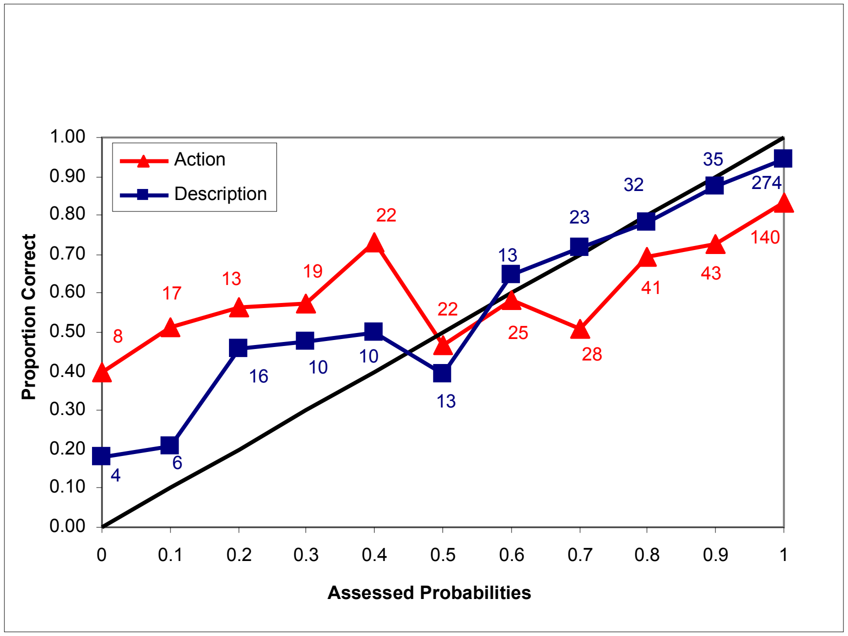


Figure 7: Calibration curves in respect of action and descriptive items by older recognition participants

Diagonal line represents perfect calibration

Young adults were well calibrated when reporting both action and descriptive items. However, older adults demonstrated considerable overconfidence in their reporting of both action and descriptive items. Older adults' confidence in respect of action items averaged .70 whereas their actual proportion correct was .53. In respect of descriptive items the assessed confidence averaged .78 whereas the actual proportion correct was .59. Recognition participants were well calibrated in response to both action and descriptive items. However cued-recall participants demonstrated considerable overconfidence in response to both action and descriptive items. Cued-recall participants' confidence in respect of action items averaged .74 whereas the actual proportion correct was .60. In respect of descriptive items, the assessed confidence averaged .81 whereas the actual proportion correct was .56. Participants' individual calibration error scores were computed as the weighted mean of the absolute difference between the actual proportion correct and the mean assessed probability for each category. The means and standard deviations (SD) of memory monitoring data are provided in *Table 5* above.

An age group x test format x content (action, description) mixed model ANOVA, with content being the within-subjects measure, was carried out on the calibration error scores.

Age group

A significant interaction was obtained between age group and content calibration error (F (df, 1, 191) = 9.76, MS Error = 0.02, $p < 0.005$, $\eta^2 = 0.05$). Analysis of the main effect means by Tukey HSD *post hoc* comparisons for unequal sample sizes, revealed that young adults were significantly more overconfident ($p < 0.01$) when reporting descriptive (.42) compared to action items (.34). Overconfidence was significantly greater ($p < 0.01$) in the older adult performance when reporting action (.37) compared with descriptive items (.36).

Test Format

The results revealed no main effects or interactions ($p > 0.05$).

Content

A within-subjects main effect of content calibration error was obtained, (F (df, 1, 191) = 8.07, MS Error = 0.02, $p = 0.005$, $\eta^2 = 0.04$). Participants were significantly more overconfident when reporting descriptive (.39) compared with action details (.36).

3.9.7.2 Monitoring Resolution

The option of free report can enhance memory accuracy only to the extent that participants are able to discriminate correct from incorrect answers. Monitoring resolution depends upon having a polarized distribution of assessed probabilities, as well as good correspondence between assessed probability and correctness.

An age group x test format x content (action, description) mixed model ANOVA, with content being the within-subjects measure, was carried out on measures of polarization (use of extreme probability ratings of 0 and 100%).

Age group

The results revealed no significant main effects or interactions ($p > 0.01$).

Test Format

The results revealed a significant between-subjects main effect of test format (F (df, 1, 191) = 7.01, MS Error = 585.50, $p < 0.01$, $\eta^2 = 0.04$). The categories of 0 and 100% were used by cued-recall and recognition participants 46% and 40% of the time respectively.

Content

A within-subjects main effect of content polarization was also obtained, (F (df, 1, 191) = 100.81, MS Error = 180.49, $p = 0.001$, $\eta^2 = 0.35$). The categories of 0 and 100% were used when recalling action and descriptive details 36% and 50% of the time respectively.

Monitoring effectiveness was evaluated in terms of within-subjects Goodman-Kruskal gamma correlations between the assessed confidence and the correctness of each answer (Nelson, 1984). The *y* correlations between the assessed confidence and the correctness of each answer were computed for each participant. At times, *y* correlations were incalculable (e.g., when no incorrect responses were made, or when only one probability category was used) and the data was not included in the analysis. This occurred when reporting descriptive items by one younger adult. Monitoring resolution averaged .67 for young compared with .60 for older adults.

The means and standard deviations (SD) of memory monitoring data are provided in *Table 5* above.

An age group x test format x content (action, description) mixed model ANOVA, with content being the within-subjects factor was carried out on the *y* correlations. The results revealed no significant between-subjects main effects or interactions ($p > 0.01$). A within-subjects main effect of content resolution was obtained, (F (df, 1, 190) = 23.40, MS Error = 0.08, $p < 0.01$, $\eta^2 = 0.11$). Monitoring resolution when reporting action details (0.54) was significantly lower compared with descriptive details (.68).

Therefore according to this index, overall participants were quite effective in discriminating correct from incorrect answers when recalling descriptive details, but not so for action details. Differences in age or test format were not apparent.

3.9.7.3 Summary of Monitoring Effectiveness

Analysis of absolute correspondence (calibration) revealed a significant age x content interaction. Young adults displayed significantly greater overconfidence when reporting descriptive when compared with action details. Whereas overconfidence was significantly greater in the older adult performance when reporting action compared with descriptive items. Relative correspondence (resolution) was significantly poorer in response to action details, although this did not vary with age or test format. This did not appear to be as a result of polarization. The categories of 0 and 100% were used when recalling action and descriptive details 36% and 50% of the time respectively.

3.9.7.4 Control Sensitivity: Subjective confidence and the decision to respond

Control processes influence accuracy by determining whether a response is volunteered or withheld. According to the Koriat and Goldsmith (1996) model, this decision is based on (a) the output of the monitoring mechanism and (b) the incentive for accuracy.

The Contribution of Monitoring

To assess the contribution of monitoring, the link between assessed confidence and the decision to volunteer an answer was computed in the form of individual *y* correlations. At times, *y* correlations were incalculable (e.g., when no responses were withheld or only one probability rating was used) and the data was not included in the analysis. This occurred when reporting action items by 14 younger and 16 older adults and when reporting descriptive items by 10 younger and 10 older adults. *Y* correlations averaged .91 for young compared with .89 for older adults, indicating a tight link between confidence and the decision to report a candidate response.

An age group x test format x incentive x content (action, description) mixed model ANOVA was carried out on the volunteering *y* scores. Box's Test of Equality of Covariances was significant ($p < 0.001$). The results revealed no significant main effects or interactions arising from the incentive conditions, therefore the data was collapsed over the remaining groups. An age group x test format x content (action, description) mixed model ANOVA was carried out on the volunteering *y* scores. Box's Test of Equality of Covariances was significant ($p = 0.01$). Levene's Test of Equality of Error Variances was significant in respect of the action volunteering *y* scores ($p < 0.05$). The value of α for this analysis was therefore set at (0.001), (Stevens, 1996).

The results revealed no significant between-subjects main effects or interactions ($p > 0.001$). The within-subjects main effect was also not significant ($p > 0.001$).

Response Criterion Setting

Did changing the incentives for accuracy affect response decisions? In the moderate incentive condition, the loss for a commission error was the same as the gain for correct recall (1 point), however in the high incentive condition the potential gain for correct recall is 5 times that for a commission error. Did such a payoff structure induce participants to use a more conservative or liberal response criterion? Response criterion estimates were calculated as the mean probability of the items that were volunteered and withheld at Phase 1 (Pansky et al., 2009). The means and standard deviations (SD) of memory monitoring data are provided in *Table 5* above.

Confidence in volunteered responses

Overall the mean confidence rating of items volunteered in the free report stage was .74 compared with .46 on withheld items. An age group x test format x incentive x content (action, description) mixed model ANOVA was carried out on the confidence judgments of items volunteered at the free report stage. Box's Test of Equality of Covariances was significant ($p < 0.001$). Levene's Test of Equality of Error Variances was significant in respect of the mean confidence of descriptive items volunteered ($p < 0.005$). The interaction effect between age group, test format and incentive only approached significance (F (df, 2, 183) = 2.81, MS Error = 0.03, $p = 0.06$, $\eta^2 = 0.03$). (See *Figures 8* and *9* below for the interaction).

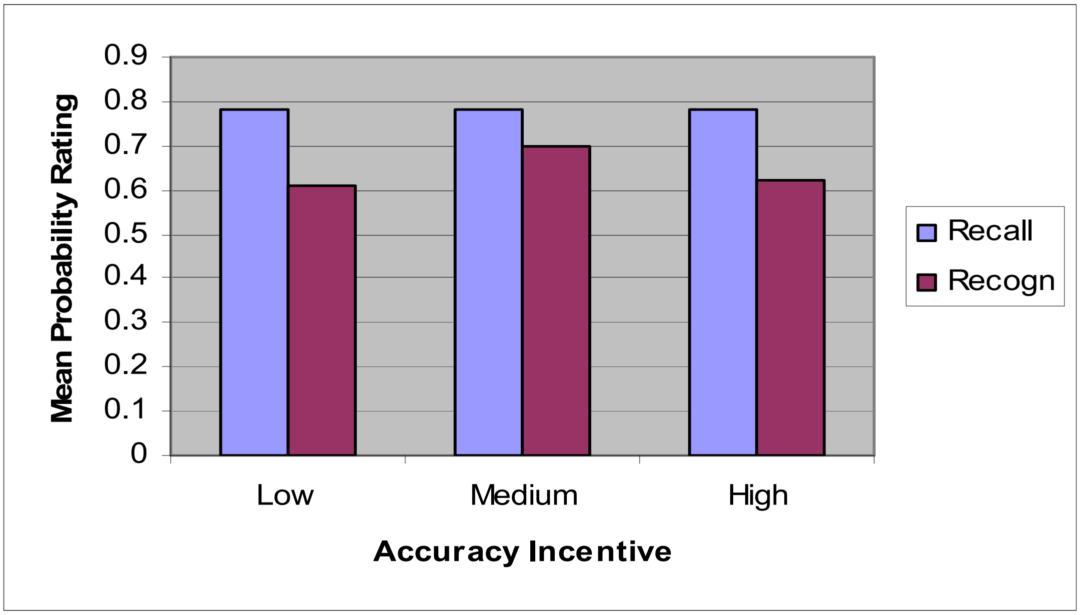


Figure 8: Younger adult interaction between test format and accuracy incentive in respect of responses volunteered at free report

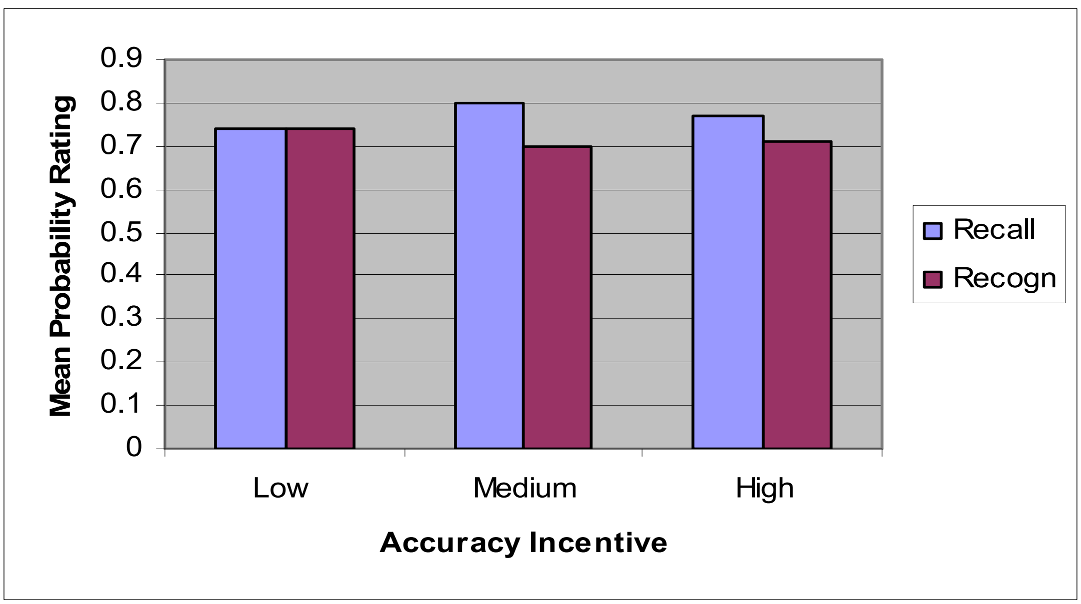


Figure 9: Older adult interaction between test format and accuracy incentive in respect of responses volunteered at free report

The data was collapsed over the remaining groups. An age group x test format x content (action, description) mixed model ANOVA was carried out. Box's Test of Equality of Covariances was significant ($p < 0.001$). Levene's Test of Equality of Error Variances was significant in respect of the confidence scores of volunteered descriptive items ($p < 0.005$). The value of α for this analysis was therefore set at (0.001), (Stevens, 1996).

Age group

The results revealed no significant main effects or interactions ($p = 0.09$).

Test Format

The results revealed a between-subjects main effect of test format (F (df, 1, 191) = 29.27, MS Error = 0.03, $p < 0.001$, $\eta^2 = 0.13$). Cued-recall participants were significantly more confident in their volunteered answers (.78) compared with recognition participants (.68).

Content

A within-subjects main effect of content was observed (F (df, 1, 191) = 26.26, MS Error = 0.01, $p < 0.001$, $\eta^2 = 0.12$). Participants set a higher response criterion when reporting descriptive (.76) compared with action details (.70). No other within-subjects main effects or interactions were observed ($p > 0.001$).

Confidence in withheld responses

An age group x test format x incentive x content (action, description) mixed model ANOVA was carried out on the confidence judgments in respect of items withheld at the free report stage. The results revealed no significant main effects or interactions arising from the incentive conditions ($p > 0.05$) therefore the data was collapsed over the remaining groups. An age group x test format x content (action, description) mixed model ANOVA was carried out. A trend was observed with test format only approaching significance (F (df, 1, 144) = 4.44, MS Error = 0.07, $p < 0.04$, $\eta^2 = 0.03$). The mean response criterion of withheld responses from cued-recall participants was 0.5 compared with .44 withheld from recognition participants. The results revealed no other main effects or interactions ($p > 0.01$).

Response Volunteering Rate

An age group x test format x incentive x content (action, description) mixed model ANOVA, was carried out on the proportion of items volunteered at free report. The results revealed no significant main effects or interactions arising from the incentive conditions ($p > 0.01$) therefore the data was collapsed over the remaining groups. An age group x test format x content mixed model ANOVA was carried out. Levene's Test of Equality of Error Variances in respect of the proportion of action details volunteered was significant ($p < 0.01$). The value of α for this analysis was therefore set at (0.001), (Stevens, 1996).

Between-subjects analysis observed a trend only in respect of main effects of age group (F (df, 1, 191) = 4.68, MS Error = 0.03, $p = 0.03$, $\eta^2 = 0.02$) and test format (F (df, 1, 191) = 5.55, MS Error = 0.03, $p = 0.02$, $\eta^2 = 0.03$). Younger adults volunteered proportionately more responses (M = .77, SE .01) compared with older adults (M = .74, SE .01). Recognition participants volunteered proportionately more responses (M = .77, SE .01) compared with cued-recall participants (M = .74, SE .01). The results revealed no other significant main effects or interactions ($p > 0.001$).

3.9.7.5 Summary of Control Processes

The assessed probability of volunteered items varied as a condition of test format and content. Cued-recall participants displayed significantly greater confidence in their volunteered answers compared with recognition participants. A higher response criterion was set in response to descriptive compared with action details. These main effects did not vary with age. The assessed probability of withheld items did not vary as a condition of age, test format or content. The confidence-volunteering relationships of cued-recall participants and in response to descriptive details were most sensitive to the assessed probabilites. The proportion of responses volunteered at free report did not differ significantly by age, test format or content. A trend only was observed in respect of age group in that younger adults chose to volunteer a greater proportion of responses compared with older adults. Thus in exercising the option of free report, participants seem to have relied upon their subjective confidence and not on monitoring resolution.

3.9.8 Mediational Analysis

The current study also examined the whether individual difference measures provide a unique variance as a predictor of free report accuracy performance beyond age, and memory monitoring and control processes. This question was considered separately for action and descriptive details using Baron and Kenny's (1986) test for mediator effects. Partial correlations were computed controlling for age. The correlation matrix is included in *Tables 6 and 7* below

Table 6: Correlation matrix between participant individual differences and monitoring and control measures

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.	Education	1.0	-.19	-.09	-.03	.04	-.06	.02	.10	.06	.11	.03	.03	.09	-.08	-.09	.02	-.04	.06
2.	WTAR		1.0	.06	.02	.02	.11	-.15	-.11	-.15	**-.21**	**-.21**	-.11	-.07	.04	-.03	-.17	-.04	**.20**
3.	Physical			1.0	**.44**	.09	**.20**	-.03	-.06	.12	.06	.13	-.02	-.06	-.09	.05	.09	.04	.06
4.	Mental				1.0	.12	**.23**	-.14	-.11	**.25**	.10	-.01	-.02	-.15	**-.30**	-.01	.12	.19	.04
5.	Vision					1.0	**.32**	-.14	-.12	.03	.04	.02	.01	.05	.03	-.02	.07	-.01	-.04
6.	Hearing						1.0	-.16	-.17	.05	.02	-.01	-.10	-.05	-.07	.07	.01	.02	-.04
7.	PA(1)							1.0	**.72**	.07	.04	.17	.02	**-.22**	-.05	.09	.01	.07	-.00
8.	PA(2)								1.0	.12	-.06	.14	.07	-.12	-.03	.10	.04	.06	.04
9.	NA(1)									1.0	**.62**	-.15	-.14	-.09	-.09	-.00	-.06	.03	.06
10.	NA(2)										1.0	-.03	-.11	-.14	-.08	.01	.01	.01	-.14
11.	Mean Conf - Action											1.0	**.44**	.04	.04	.12	-.03	.04	-.05
12.	Mean Conf - Description												1.0	-.03	.14	09	.04	.08	.08
13.	Act Resolution													1.0	-.02	-.18	-.04	-.10	-.09
14.	Desc Resolution														1.0	.07	.06	-.05	.04
15.	Response VR – Action															1.0	-.07	.11	.07
16.	Response VR - Description																1.0	-.03	-.15
17.	Control Sensitivity - Action																	1.0	**.20**
18.	Control Sensitivity - Description																		1.0

All correlations in bold type-face were significant at $p \leq 0.01$ or better.

1.Years of Education; 2. Wechsler Test of Adult Reading (Wechsler, 2001); 3. Physical Health – self report on 5-point scale; 4. Mental Health – self report on 5-point scale; 5. Vision – self report on 5-point rating scale; 6. Hearing – self report on 5-point rating scale; 7-10 Positive and Negative Affect (Watson et al., 1988); 11. Mean Confidence Rating – Action Details; 12.Mean Confidence Rating – Descriptive Details; 13. Monitoring Reolution – Action Details; 14. Monitoring Resolution - Descriptive Details; 15. Response Volunteering Rate – Action Details; 16. Response Volunteering Rate – Descriptive Details; 17. Control Sensitivity – Action Details; 18. Control Sensitivity – Description Details.

Table 7: Correlation matrix between participant individual differences and monitoring and control measures

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Strategy	*1.0*	.18	**-.44**	**-.30**	**.44**	**.27**	.13	**-.22**	**-.30**	-.04	.03	-.08	.11	-.03	.15	-.07	-.00
2. Task		*1.0*	-.12	-.10	.18	.08	-.16	.00	-.08	.11	.05	.00	-.05	-.06	.14	.05	.03
3. Capacity			*1.0*	**.54**	**-.51**	-.00	**.22**	**.23**	**.40**	.03	-.07	-.08	-.03	.07	.01	.12	-.02
4. Change				*1.0*	-.50	-.09	**.31**	**.27**	**.24**	.10	.10	.09	.05	.15	.00	.03	-.01
5. Anxiety					*1.0*	**24**	-.14	-.12	**-.41**	-.12	-.01	-.06	-.01	-.11	.12	-.02	-.07
6. Achievement						*1.0*	**.25**	-.00	-.00	.05	.00	-.11	-.07	.02	.11	.13	-.04
7. Locus							*1.0*	.13	**.24**	.01	.09	-.15	-.05	.11	.08	.04	.07
8. SEL								*1.0*	**.22**	.19	.13	.01	.00	.05	.15	-.10	-.11
9. SEST									*1.0*	-.02	.03	-.05	-.07	.16	-.11	.05	.09
10. CR Mean - Action										*1.0*	**.44**	.04	.03	.12	-.04	.04	-.05
11. CR Mean - Description											*1.0*	-.03	.14	.09	.04	.08	.08
12. Resolution – Act												*1.0*	-.01	-.18	-.04	-.11	-.09
13. Resolution – Des													*1.0*	.07	.05	-.05	.04
14. Response VR – Action														*1.0*	-.07	.11	.07
15. Response VR - Description															1.0	-.03	-.15
16. Control Sensitivity – Action																1.0	**.20**
17. Control Sensitivity – Description																	1.0

All correlations in bold type-face are significant at $p \leq 0.01$ or better

1 – 7 Metamemory in Adulthood Questionnaire (MIA; Dixon et al., 1988), scales: Strategy, Task, Capacity, Change, Anxiety, Achievement, Locus.
8. Memory Self Efficacy Questionnaire, (MSEQ; Berry et al., 1989) - Self-efficacy Level; 9 Self-efficacy strength ; 10. Mean Confidence Rating – Action Details; 11.Mean Confidence Rating – Descriptive Details; 12. Monitoring Reolution – Action Details; 13. Monitoring Resolution - Descriptive Details; 14. Response Volunteering Rate – Action Details; 15. Response Volunteering Rate – Descriptive Details; 16. Control Sensitivity – Action Details; 17. Control Sensitivity – Description Details.

A negative correlation was observed between scores on the WTAR and the mean confidence rating when reporting action details ($r = -.21$, $p = 0.01$), showing that the lower the participant's score on the WTAR the higher the mean confidence rating when reporting action details. A significant negative correlation was observed between self-report measures of mental health and monitoring resolution when reporting descriptive details ($r = -.30$, $p < 0.001$), showing that the poorer the participant's mental health, the less they were able to discriminate correct and incorrect descriptive details.

A significant positive correlation was observed between scores on the WTAR and the participant control sensitivity when reporting description details ($r = .20$, $p = 0.01$), showing that the higher the participant's score on the WTAR the greater the relationship between assessed confidence and the decision to volunteer an answer in respect of description details. A significant negative relationship was observed between pre-test positive affect and monitoring resolution in respect of action details ($r = -.22$, $p < 0.01$), showing that the more positive the participant's mood the more they were better able to discriminate correct and incorrect responses when reporting action details.

To assess the unique contribution of pre-test positive affect in explaining the monitoring resolution of action details, over and above age, multiple regression analysis was performed. Age and pre-test positive affect, were included as predictors. *Table 8* below gives the outcome of the regression analysis.

Table 8: Regression analysis – predictors of monitoring resolution in respect of action details

		B	*SE B*	β	*p*
Model					
	Constant	.89	.10		
	Age	.00	.00	-.07	$p > 0.05$
	Positive Affect(1)	-.01	.00	-.21	$p < 0.005$

$\Delta R^2 = .05$

To explain the monitoring resolution of action details, self-report measures of pre-test positive affect were identified as a unique predictor over and above age.

To assess the unique contribution of self-report measures of mental health in explaining the monitoring resolution of descriptive details, over and above age, multiple regression analysis was performed. Age and self-report measures of mental health were included as predictors. *Table 9* below gives the outcome of the regression analysis.

Table 9: Regression analysis – predictors of monitoring resolution in respect of descriptive details

		B	*SE B*	β	*p*
Model					
	Constant	.85	.06		
	Age	.00	.00	-.09	$p > 0.05$
	Mental Health	-.07	.02	-.22	$p < 0.05$

$\Delta R^2 = .05$.

To explain the monitoring resolution of descriptive details, self-report measures of mental health were identified as a unique predictor, over and above age.

To assess the unique contribution of scores on the WTAR, in explaining the relationship between assessed confidence and the decision to volunteer a descriptive response at free report, over and above age, multiple regression analysis was performed. Age and scores on the WTAR were included as predictors. *Table 10* below gives the outcome of the regression analysis.

Table 10: Regression analysis for the relationship between assessed confidence and the decision to volunteer a descriptive response at free report.

		B	*SE B*	β	*p*
Model					
	Constant	.95	.03		
	Age	-.00	.00	-.17	$p > 0.05$
	WTAR	.00	.00	.18	$p > 0.05$

$\Delta R^2 = .02$.

In explaining the relationship between assessed confidence and the decision to volunteer a descriptive response at free report, scores on the WTAR were not identified as a unique predictor over and above age.

To assess the unique contributions of scores on the WTAR, and pre-test positive affect in explaining the mean confidence rating of action details volunteered at free report, over and above age, multiple regression analysis was performed. Age, scores on the WTAR and self-report measures of pre-test positive affect were included as predictors. *Table 11* below gives the outcome of the regression analysis.

Table 11: Regression analysis – predictors of mean confidence rating in respect of action details volunteered at free report

		B	*SE B*	β	*p*
Model					
	Constant	.67	.08		
	Age	.00	.00	.03	$p > 0.05$
	WTAR	-.00	.00	.02	$p > 0.05$
	Positive Affect(1)	.00	.00	.01	$p > 0.05$

$\Delta R^2 = .01$.

To explain the mean confidence rating of action details volunteered at free report, neither scores on WTAR or pre-test positive affect were identified as unique predictors.

Do age and individual differences mediate the effects of monitoring and control on free report memory accuracy?

The extent to which the variables of age and individual differences in measures of pre-test positive affect, and mental health could contribute either directly or indirectly to memory accuracy at free report, and the way in which these are mediated by memory and metamemory components, a series of were conducted (see *Figures 10* and *11* below). Path analyses were conducted separately for memory accuracy when participants reported action and descriptive details. Based on Koriat and Goldsmith's (1996) model, each path model assumed that free report accuracy was affected directly by forced report quantity correct (the amount of correct information retrieved), monitoring resolution (within-participant gamma correlation between confidence rating and the actual correctness of each answer), mean confidence rating, response volunteering rate (the proportion of answers volunteered under free report), and control sensitivity (within-participant gamma correlation between confidence rating and the decision to volunteer / withhold the response). In addition, monitoring resolution was assumed to depend in part on the amount and quality of information retrieved (see Kelley & Sahakyan, 2003; Koriat, 1993, 1995; Pansky et al., 2009; Rhodes & Kelley, 2005). Therefore age or individual difference measures might affect monitoring resolution directly or indirectly via an effect on forced report quantity correct. Mean confidence was also assumed to be strongly determined by forced-report quantity correct. Therefore any additional effect of age or individual difference measures on the confidence mean would indicate an effect that could not be explained by differences in the correctness of responses, and would therefore reflect either differences in the use of the confidence scale (i.e. assigning numbers to subjective confidence levels) or differences in monitoring calibration bias (over-or under confidence). Similarly, the answer withholding rate was assumed to be strongly determined by the confidence mean. Any additional effect of age or individual difference measures on this measure would indicate a difference in the withholding rate that could not be explained by a difference in confidence mean alone, and would therefore reflect a more conservative or more liberal control policy.

Prior to analysis, participant scores on the variables of pre-test measures of positive affect, and mental health were centred by subtracting the mean of each variable from each person's score. *Table 12* below shows the descriptive statistics for these new variables. In order to test the interaction effect, a cross-product term was then created by multiplying the two variables of interest (Keith, 2006).

Table 12: Means and (SD) of pre-test positive affect, and mental health – centred by age group

Measure	Younger Adults	Older Adults
Positive Affect (1) – C	-.83 (6.28)	.89 (6.89)
Mental Health – C	-.09 (.77)	.10 (.95)

The direct and indirect effects of age, measures of pre-test positive affect, and mental health in respect of accuracy when reporting action details at free report, are examined in *Figure 10* below.

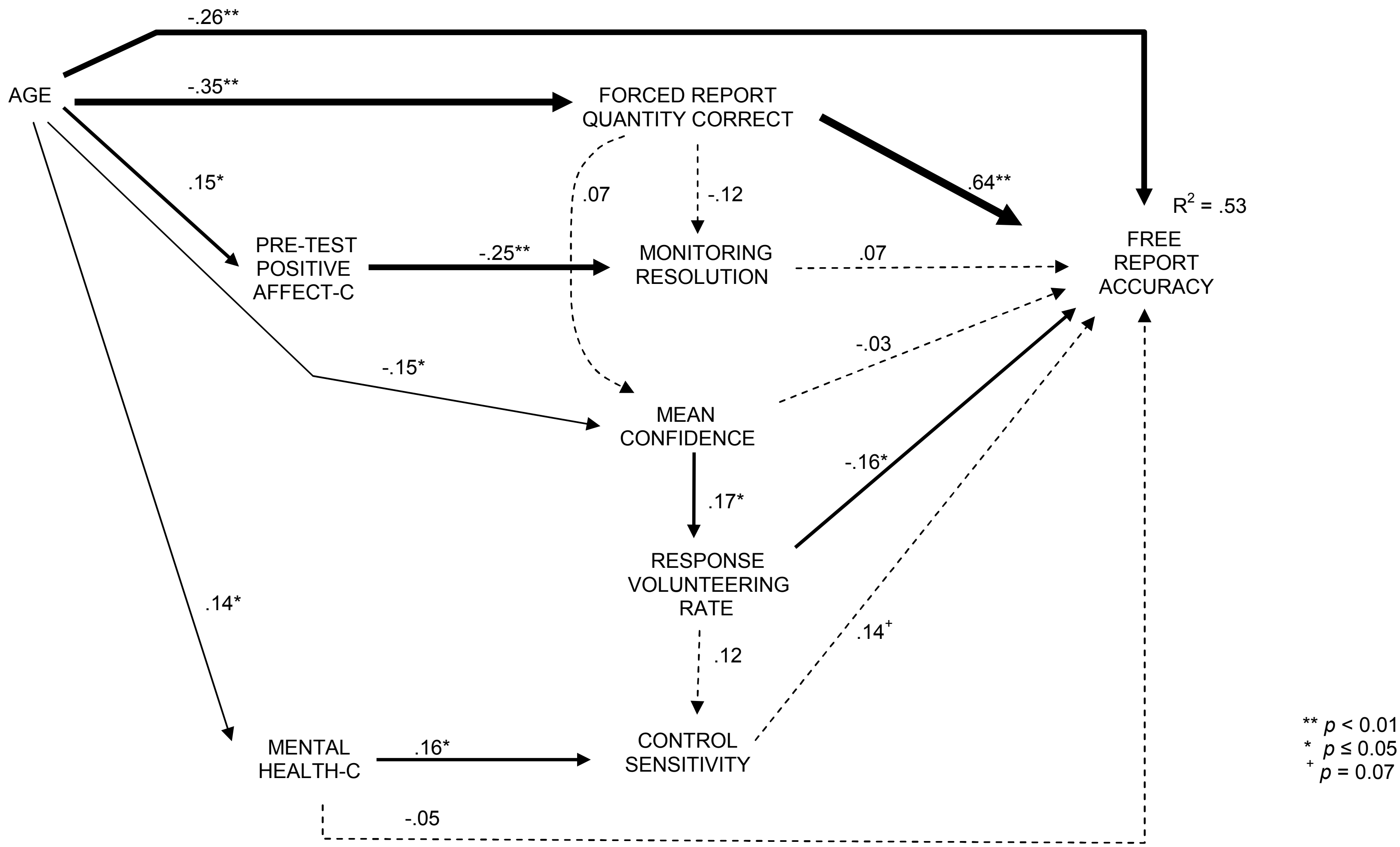


Figure 10: Path models examining the effects of chronological age and individual difference measures for free report accuracy performance in respect of action details and the mediation of these effects by memory and metamemory components. Coefficients for each path represent standardised β weights.

Statistically significant coefficients / paths are indicated by solid arrows.

For the sake of clarity, non-significant paths between the cognitive variables, quantity correct at Forced Report and accuracy at Free Report have been omitted.

Overall, the model outlined at Figure 10 above explains approximately 53% of the variability in memory accuracy when participants reported action details at free report. Consistent with prior results (Kelley & Sahakyan, 2003; Koriat, 1993, 1995; Pansky et al., 2009; Rhodes & Kelley, 2005) retrieval at forced report played a large role in free report accuracy. When reporting action items at forced report, there was a strong direct relationship between the quantity correct, and accuracy at free report. The indirect effect of monitoring on quantity correct at forced report was not significant. Additionally, a strong direct relationship between age and free report accuracy was observed, as well as an indirect relationship that was mediated by quantity correct at forced report. Age also had an indirect relationship to accuracy that was largely mediated by its effects on individual difference measures, which were in turn seen to influence participant monitoring resolution, mean confidence and control sensitivity respectively.

Critical to the investigation of the way in which individual difference measures produce individual differences at free report, pre-test positive affect was seen to influence monitoring resolution when participants reported action details, although monitoring resolution did not influence accuracy. The confidence mean influenced participant response volunteering rate which also influenced accuracy. Self-report measures of mental health were seen to influence participant control sensitivity, which in turn influenced accuracy. Thus an indirect relationship was observed between self-report measures of mental health and participant control policy. The interaction effect of the three individual difference variables, on each of the outcome variables was not significant ($p > 0.05$). This data indicate that memory accuracy when reporting action details at free report was largely affected directly by age and quantity correct at forced report, and through the indirect effects of participant mental health on control policy.

Memory accuracy at free report for descriptive items was examined in a model identical to that used for accuracy on action items (see *Figure 11*).

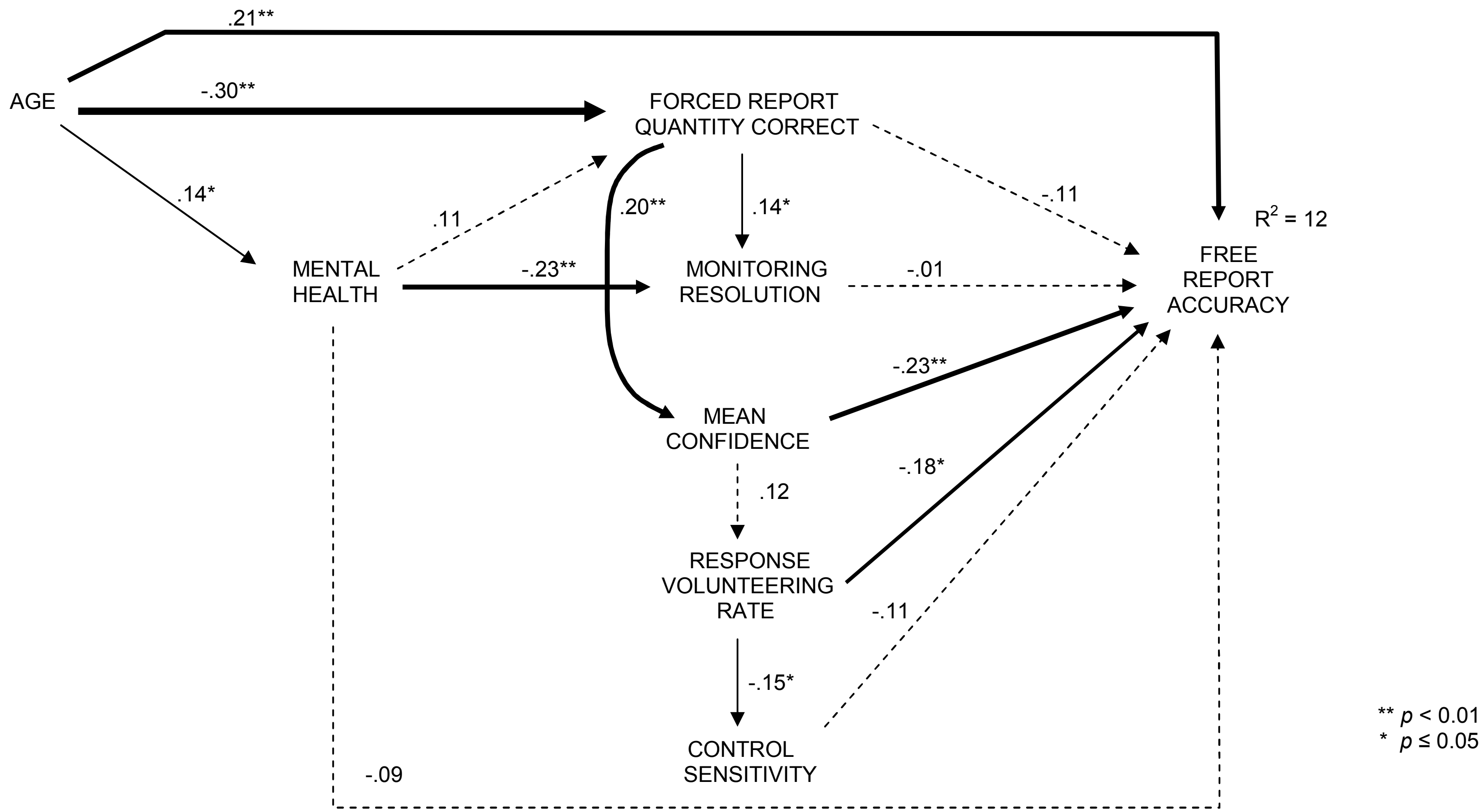


Figure 11: Path models examining the effects of chronological age and Individual Difference measures on Free Report accuracy performance for description details and the mediation of these effects by memory and metamemory components. Coefficients for each path represent standardised β weights. Statistically significant coefficients / paths are indicated by solid arrows.

For the sake of clarity, non-significant paths between several of the cognitive and metacognitive variables have been omitted.

Overall, the model explains approximately 12% of the variability in memory accuracy when participants reported descriptive details at free report. As in the model in respect of action items, there was a direct relationship between age and free report accuracy however the effect was smaller than was the case when reporting action items. Age was seen to influence the quantity correct at forced report. However in contrast to reported action details this did not affect accuracy. This relationship did affect accuracy through the ability to monitor the correctness of these items. Age was also seen to influence self-report measures of mental health, which in turn influenced monitoring resolution. However, monitoring resolution did not affect accuracy. Again unlike accuracy when reporting action details, the confidence mean in respect of descriptive details was directly influenced by the quantity correct at forced report. This data indicates that when reporting descriptive details at free report, memory accuracy was largely affected directly by age and through the indirect effects of control policy and not on quantity correct at forced report. In addition, monitoring did not make a significant independent contribution to accuracy in respect of either action or descriptive details.

3.9.9 Results Summary

Preliminary Analysis

Older adults attained fewer years of education, correctly pronounced more irregular words, and reported poorer mental health, vision and hearing. Significant age effects appeared in the MIA scales of capacity, change, achievement, and MSEQ strength. Younger adults reported significantly more positive beliefs about their memory. Analysis further revealed that correlations between the aforementioned individual difference variables and the outcome performance variables did not reach significance at $p < 0.01$ or better and therefore were not included as covariates in the between-subjects further analysis.

Memory Accuracy at Free Report

Participant recall at free report did not differ by age. However a significant interaction between age group and test format was observed. Younger adult performance was significantly greater in response to recognition compared with cued-recall questions, whereas older adult performance was significantly greater in response to cued-recall questions. A significant interaction was also observed between age group and content. Older adult accuracy was significantly greater when reporting descriptive compared with action details. There was no significant difference in the performance of younger adults when reporting action compared with descriptive details. Overall, accuracy was significantly greater in response to cued-recall questions, and descriptive details.

Memory Accuracy at Forced Report

Memory retrieval at forced report was significantly greater by younger adults, recognition participants and descriptive details. An interaction between age group, test format and content was not obtained.

Performance consequences

At free report, the between-subjects analysis of memory accuracy revealed no main effect of age. Within-subjects analysis of content revealed a significant age x content interaction. Older adult memory accuracy was significantly greater when reporting descriptive compared with action details. No significant differences were observed in the younger adult content performance. At forced report between-subjects analysis revealed a main effect of age. Older adults retrieved significantly fewer correct items of information when compared with younger adults. No significant age x content interactions were observed. Within-subjects analysis of the gains in memory accuracy performance from forced to free report revealed that the older group gained significantly more than younger adults from the opportunity to screen out potential incorrect responses. Significantly greater gains were realised in respect of descriptive details (see *Figure 2* above for the interaction). Gains in older adult accuracy came at the expense of greater losses in quantity correct. Older adults lost a greater number of correct action details compared with descriptive details. Younger adults lost a greater number of correct descriptive compared with action details (see *Figure 3* above for the interaction).

Analysis of Monitoring Effectiveness

Analysis of absolute correspondence (calibration) revealed a significant age x content interaction. Younger adults' overconfidence was significantly greater in respect of descriptive compared with action details. Whereas older adults' overconfidence was significantly greater in respect of action compared with descriptive items. Relative correspondence (resolution) was significantly poorer in response to action details, although this did not vary with age or test format. This did not appear to be as a result of polarization. The categories of 0 and 100% were used when recalling action and descriptive details 36% and 50% of the time respectively.

Analysis of Control

The assessed probability of volunteered items varied as a condition of test format and type of content. Cued-recall participants were significantly more confident in their volunteered answers compared with recognition participants. A higher response criterion was set in response to descriptive compared with action details. These main effects did not vary with age. The assessed probability of withheld items did not vary as a condition of age, test format or content. The confidence-volunteering relationships of cued-recall participants and in response to descriptive details were most sensitive to the assessed probabilites. The proportion of responses volunteered at free report did not differ significantly by age, test format or content. A trend only was observed in respect of age group in that younger adults chose to volunteer a greater proportion of responses compared with older adults. Thus in exercising the option of free report, participants seem to have relied upon their subjective confidence and not on monitoring resolution.

Mediational Analysis

Path analysis indicate that memory accuracy in free report when reporting action details was largely affected directly by age and indirectly via retrieval at forced report, and through the indirect effects of participant mental health on control policy. Memory accuracy in free report when reporting descriptive details was largely affected directly by age and through the indirect effects of control policy and not by retrieval at forced report. In addition, monitoring did not make a significant independent contribution to accuracy in the reporting of either action or descriptive details.

3.10 Discussion

The current study assessed the memory monitoring and control processes when participants responded to action and description items contained within a to-be-remembered neutral event. These issues were examined in younger and older adults, two groups known to differ in memory retrieval and also possibly in terms of monitoring and control processes. Test format was also manipulated as age-related decrements have been found to be larger in recall than in recognition (Craik & McDowd, 1987; Rabinowitz 1984, 1986). The experiment assessed the way in which these factors interact within the general theoretical model developed by Koriat and Goldsmith (1996), as detailed above, which specifies the mediating role of metacognitive processes. In light of previous findings, also detailed above, it was predicted that there would be three main effects of age, test format, and type of content. Memory accuracy at free report was predicted to be significantly greater by younger adults, greater in response to cued-recall questions, and greater when reporting description compared with action details. Interactive patterns of results may be observed. The current study also examined the direct and indirect effect of individual difference measures on free report accuracy performance beyond age, and memory monitoring and control processes.

3.10.1 Memory Accuracy at Free Report

At free report between-subjects analysis observed no significant differences in the accuracy performance of young and older adults. An interaction between age and test format revealed that older adult accuracy was significantly greater in response to cued-recall questions. In contrast, younger adult accuracy was significantly greater in response to recognition questions. Accuracy was greater when participants reported descriptive compared with action details. This was particularly the case in the older adult performance. The interaction between age group, test format and content was not significant. These results partially support the first experimental hypothesis.

That no significant differences were observed in the overall accuracy performance of young and older adults at free report, was contrary to expectations and existing memory research (Kelley & Sahakyen, 2003; Pansky et al., 2009; Rhodes & Kelley, 2005; Wilcock et al., 2008). The null hypothesis cannot therefore be rejected. The option of free report allowed older adults to regulate their memory accuracy to the extent that their performance was equivalent to that of younger adults. This could perhaps imply greater monitoring effectiveness. This point will be addressed later in the discussion section.

Memory accuracy at free report was significantly greater in response to cued-recall questions compared with recognition questions. This finding was as predicted and supports the stated experimental hypothesis and prior results (see Fisher & Patterson, 2004; Hilgard & Loftus, 1979; Ibabe & Sporer, 2004; Koriat & Goldsmith, 1996 –

Experiment 1; Lipton, 1977; Neisser, 1988). When open-ended questions are asked in a non-leading manner, they involve a smaller risk of misleading participants because the questions provide a lower amount of new information than multiple choice questions. The information contained within the question serves as a recall cue for retrieving the correct answer. In contrast, an Alternative-Forced-Choice (AFC) test introduces information which may be chosen on the basis of familiarity, or on a subjective selection of the most likely alternative. The number of alternatives in AFC recognition tests also influences the percentage of correct responses; the probability of producing a correct answer by guessing decreases as the number of alternatives increase. In open-ended questions, the number of possible answers is unlimited compared to recognition tests, thus reducing the chance of correctly guessing an answer.

In the current study, an interaction between age and test format revealed that at free report older adult accuracy was significantly greater in response to cued-recall questions. However younger adult accuracy was significantly greater in response to recognition questions. This interactive pattern was not in the expected direction, and is contrary to existing research which shows greater age-related decrements in response to recall compared with recognition questions (Craik & McDowd, 1987; Nyberg et al., 2003; Rabinowitz 1984, 1986) and greater in response to recall compared with cued-recall questions (Ceci & Tabor 1981; Craik et al., 1987).

The environmental support hypothesis (Craik, 1986, 1994) proposes that age-related reductions in attentional capacity make older adults less able to engage in self-initiated remembering processes (e.g., retrieval searches). According to this theory, older adults' memories are greatly influenced by the amount of contextual cues available from the environment (Hasher et al., 2001). When extensive contextual cues from the original event are present at retrieval, age-related recall deficits are reduced (Fernandez & Alonso, 2001; Sharps & Antonelli, 1997; Yarmey & Yarmey, 1997). Recognition memory is an example of environmental support (Craik & Jennings, 1992), because a target item is re-presented. This is in contrast to cued-recall which involves more self-initiated activity and which is a more explicit source identification task.

In the present study, it would appear that older adults were able to employ sufficient internal processing resources to consciously retrieve the correct answer, by using the information contained within the open-ended questions as a recall cue. This result is contrary to existing research (Ceci & Tabor 1981; Craik et al., 1987; Craik & McDowd, 1987; Nyberg et al., 2003; Rabinowitz 1984, 1986) which suggests that older adults respond on the basis of familiarity or plausibility, and may reflect individual differences in the older adult population sample used. This point will be discussed in greater detail later in this section. In contrast, younger adult accuracy at free report was significantly greater in response to recognition questions which may reflect responses chosen on the basis of familiarity or plausibility. However, this result may simply reflect a spurious result by younger adults.

Accuracy was significantly greater at free report when participants reported descriptive compared with action details. This pattern of results was more pronounced in the performance of older adults. This finding was as predicted and supports the stated experimental hypothesis and prior research (see Aizpurua et al., 2009; Garcia-Bajos & Migueles 2003; Migueles & Garcia-Bajos 2004 – but see Ibabe & Sporer, 2004).

Actions of an event best capture the argument and constitute the gist of an event (Aizpurua et al., 2009). One possible explanation for the results reported here is that when reporting action details, an individual may recall information congruent with the prior knowledge of the event held in the form of situational schemata or scripts (Garcia-Bajos & Migueles, 2003; Greenberg et al., 1998; List, 1986). In contrast, he or she is less likely to possess schemata in respect of descriptive details (Heuer & Reisberg, 1990). Memory bias resulting from these schemata is particularly significant in the recognition for the actions of an event. Access to gist information is relatively automatic, and is preserved with aging (Cohen & Faulkner, 1989; Koutstaal & Schacter, 1997). It is known that older adults are more prone than younger adults to depend on prior knowledge (Mather et al., 1999). If older adults are overly dependent on gist-based processing they are more likely to display memory bias (Craik & Simon, 1980; Hasher & Zacks, 1979; Hashtroudi et al., 1990; Puglisi et al., 1988; Rabinowitz & Ackerman, 1982). Poorer retention will increase the generation of incorrect responses, particularly when reporting action details.

Alternatively, an increase in errors when identifying actor-action links may be due to a reliance on familiarity of the components, that is, the people and the actions rather than recollection of the association between the two. This interpretation is consistent with the associate deficit hypothesis (Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2003; Old & Naveh-Benjamin, 2008), detailed in Chapter 1, which attributes age-related memory deficits to the inability to encode and bind together items or features of an event.

The findings of the present study would appear to extend memory deficits for action details, particularly by older adults, to a neutral scene. However this interpretation should be treated with caution. The present study replicated the research of Ibabe and Sporer (2004), and defined action details as character non verbal behaviours. Descriptive details were defined as physical characteristics of scenes, persons and objects. In contrast Aizpurua et al., (2009) defined actions as verbal or non-verbal behaviours (e.g., pointed a gun at someone's face). The people related content included general physical characteristics of the people involved, as well as descriptions of their clothing and accessories. Details referred to primary object characteristics and circumstantial information surrounding the event. Ibabe and Sporer, (2004) found that young adult accuracy was significantly greater when reporting action compared with descriptive details. In contrast, Aizpurua et al., (2009), found that participants accepted more false actions, thus achieving greater recognition accuracy when reporting people and details. This pattern of results was more pronounced in older adults. In both studies the filmed event consisted of a violent robbery. The differing conceptualisation may have contributed to the differing results.

The differing conceptualisation regarding memory for each component (i.e. individual actions and people) is important to future research within the eyewitness domain. An agreed definition of *action* and *description* details of an event would help eyewitness researchers define and focus the applications of their research. Additionally, memory performance should be investigated within different staged events (neutral scenes and emotional crime events) to assess whether the results reported here are replicated across differing experimental stimuli.

The reported memory performance results may have implications for eyewitness testimony. During an initial investigation, detailed descriptions of what happened, who the perpetrators were and leads to further assist police with their investigation may be obtained. Forensic investigators should be aware that older adults may remember having seen a particular person before, as well as the actions involved in a crime, but may be unable to bind the culprit to the committed offense. This interpretation would seem to be consistent with the knowledge that, as introduced in Chapter 1, older adults tend to commit more errors in identifying criminals, and these errors are most often related to false identifications. This is especially the case following relatively long delays (Havard & Memon, 2009; Memon & Bartlett, 2002; Memon & Gabbert, 2003a, 2003b; Rose et al., 2005; Searcy et al., 2000; Wilcock et al., 2005, 2007; Yarmey, 1993).

Memory Retrieval at Forced Report

Memory retention measured by forced-report memory performance revealed that older adults retrieved significantly fewer action and descriptive items compared with younger adults. This finding was as expected and is in line with prior research which suggests that older adults encode events in a less elaborate, more general way (Craik & Simon, 1980; Hasher & Zacks, 1979; Hashtroudi et al., 1990; Kelley & Sahakyan, 2003; Koutstaal, et al., 1999; Koutstaal et al., 2003; Pansky et al., 2009; Puglisi, et al.,1988; Rabinowitz & Ackerman, 1982; Rhodes & Kelley, 2005).

Retrieval at forced report was significantly greater in response to recognition compared with cued-recall questions, which was as predicted and supports existing research (see Clifford & Scott, 1978; Yuille & Cutshall, 1989). A significantly greater number of descriptive compared with action details were retrieved, which again was as predicted and supports existing research (Garcia-Bajos & Migueles, 2003; Greenberg et al., List, 1986; Migueles & Garcia-Bajos, 2004).

3.10.2 Memory Monitoring and Control

Monitoring Effectiveness

Two indices of monitoring effectiveness were calculated. In line with the Koriat and Goldsmith (1996) framework, both absolute correspondence (calibration) and relative correspondence (resolution) were evaluated. Calibration reflects the overall relation between accuracy and the level of the confidence ratings. In a calibration graph, confidence (*x* axis) is calculated separately for different confidence levels and is then plotted against accuracy (*y* axis). The calibration graph indicates over / under confidence (i.e. the extent to which the confidence judgments are higher or lower than the actual proportions correct). Monitoring resolution is the degree to which assessed probabilities of correctness successfully differentiate correct and incorrect candidate answers, and is measured by within-subject Kruskal-Goodman gamma correlations (Nelson, 1984).

Calibration

As discussed above, memory bias resulting from situational schemata particularly affects memory accuracy when reporting action details (Garcia-Bajos & Migueles, 2003; Migueles & Garcia-Bajos, 2004), that may none-the-less be held in high confidence

(Aizpurua et al., 2009; Ibabe & Sporer, 2004). The second experimental hypothesis predicted that participants' would show greater over-confidence in response to action items. This pattern of results was expected to be more pronounced in older adult performance (Aizpurua et al., 2009). The results partially support the second experimental hypothesis.

The results revealed an interaction between age and content. Younger adults' over-confidence was significantly greater in response to descriptive compared with action items. As predicted older adults' over-confidence was significantly greater in response to action compared with descriptive items. This result supports the findings of (e.g., Aizpurua et al., 2009; Garcia-Bajos & Migueles, 2003; Greenberg et al., 1998; List, 1986; Migueles & Garcia-Bajos, 2004). In this respect, the results follow those for accuracy.

As discussed above, action details best capture the argument and constitute the gist of an event (Aizpurua et al., 2009). Memory bias resulting from existing schemata is particularly significant in the recognition of actions (Garcia-Bajos & Migueles, 2003; Greenberg et al., 1998; List, 1986). Additionally, research suggests that older adults encode events in a less elaborate, more general way (Craik & Simon, 1980; Hasher & Zacks, 1979; Hashtroudi et al., 1990; Kelley & Sahakyan, 2003; Koutstaal, et al., 1999; Koutstaal et al., 2003; Pansky et al., 2009; Puglisi, et al., 1988; Rabinowitz & Ackerman, 1982; Rhodes & Kelley, 2005). Consequently, if older adults are overly dependent on gist-based processing they are more likely to display memory bias (Craik & Simon, 1980; Hasher & Zacks, 1979). Such bias may be used as a metacognitive cue for confidence. Over-confidence in respect of action details resulting from gist-based processing may reflect differences in control sensitivity, i.e. the setting of a more liberal response criterion, which will be addressed later in this discussion.

Monitoring Resolution

When participants are given the option to volunteer or withhold a response, they can enhance memory accuracy only to the extent that they are able to discriminate correct from incorrect answers. Successful monitoring of the validity of responses is dependent on participants being able to adjust confidence accordingly. As predicted, monitoring resolution was significantly lower in response to action compared with descriptive details. This pattern of results reflects differences in the efficiency of encoding processes as reflected in the number of action and descriptive items retrieved at forced report. Contrary to expectations and prior research (Kelley & Sahakyen, 2003; Rhodes & Kelley, 2005; Pansky et al., 2009) this finding did not vary with age. In the current study, older adults made large gains in memory accuracy between forced and free report, particularly when reporting description details. However, this came at a higher cost in quantity correct compared with younger adults.

The analysis reported here draws upon within-subjects comparisons of confidence, which is more sensitive to participants' ability to monitor the correctness of their responses than the between-subjects accuracy-confidence relationship, as described in Chapter 1, reported for identification decisions (Sporer, 2008; Sporer et al., 1995). The traditional measurement of the AC relation is to calculate a point-biserial correlation coefficient resting on two data points per participant. In the current study 24 items were entered into the calculation. The results would seem to confirm Sporer's (2008) argument that the AC relationship for event memory details may be higher. The present findings (if replicated with different stimulus material and questions) would also have important practical implications. Researchers and practitioners should clearly separate confidence in respect of individual questions about the event from confidence relating to an identification decision. Additionally, while the AC relation for identification decisions may be weak (Bothwell et al., 1987; Penrod et al., 1982; Sporer et al., 1995; Wells & Murray, 1984) and malleable (Granhag, 1997; Granhag et al., 2000; Hastie et al., 1978; Odinot et al., 2009; Shaw, 1996; Shaw & McClure, 1996; Turtle & Yuille, 1994), witnesses, particularly older adults, may be better able to calibrate their confidence when reporting descriptive compared with action details contained within an event.

Control Sensitivity - Subjective confidence and the decision to respond

Control processes influence accuracy by determining whether a response is volunteered or withheld. According to the Koriat and Goldsmith (1996) model, this decision is based on (a) the output of the monitoring mechanism and (b) the incentive for accuracy. If participants are distrustful of their memory monitoring as well as their memory *per se*, this might produce a weaker relationship between confidence assessments and the decision to respond or withhold a potential response. It was predicted that this pattern of findings might be more pronounced in older adults resulting in reduced control sensitivity (Kelley & Sahakyen, 2003; Rhodes & Kelley, 2005; Pansky et al., 2009). Contrary to expectations and prior research, no significant differences were observed between the conditions. Older adults were able to selectively withhold incorrect responses from free report such that their accuracy was equivalent to that of younger adults.

Response Criterion Setting

According to Koriat and Goldsmith's (1996) framework, the response criterion may vary based on the level of accuracy incentive. This can be adjusted upward where large losses are associated with a commission error (wrong answer) or downward if there is no penalty for a commission error and the emphasis is placed on the quantity of correct answers. Under high incentives for accuracy compared with low or moderate incentives, people will be able to increase accuracy if they have effective monitoring of the probability of correct candidate answers, good control sensitivity, and effective response criterion setting.

In the present study, accuracy incentive was manipulated by rewarding correct candidate responses. Prior research has manipulated accuracy incentive by imposing a monetary penalty for each incorrect candidate response (Koriat & Goldsmith, 1994, 1996; Kelley & Sahakyan, 2003; Pansky et al., 2009). Younger adults respond to such an incentive structure, however older adults do not (Kelley & Sahakyan, 2003; Pansky et al., 2009). Older adults have been shown to respond to various pay-off matrixes regarding bias on a recognition test, although not to the same degree as younger adults (Baron & Surdy, 1990). The final experimental hypothesis suggested that participants might be expected to set a more conservative response criterion in an effort to improve their accuracy performance. Consequently, participants were also expected to volunteer significantly fewer correct (and incorrect) responses at free report. This pattern of results might be more pronounced in older adults.

The results partially support this hypothesis. Participants were significantly more confident in their volunteered compared with withheld answers. A trend only was observed in that the interaction effect between age group, test format and incentive approached significance (see *Figures 8* and 9 above). Participants did not respond to the high accuracy incentive by setting their response criterion higher. The high accuracy incentive was manipulated by offering 5 points for each correct candidate answer, with a deduction of 1 point for each incorrect candidate answer. Such an incentive structure may have encouraged participants to guess, resulting in the setting of a lower response criterion rather than the desired higher response criterion. Cued-recall, older adults only responded to the moderate incentive condition by setting their response criterion higher. Older adult recognition participants set a higher response criterion compared with younger adults. This result is contrary to previous results (Aizpurua et al., 2009; da Silva & Sunderland, 2010; Kelley & Sahakyan, 2003; Koutstaal & Schacter, 1997; Pansky et al., 2009), whereby older adults were observed to adopt a more lenient response criterion than young adults. This may reflect individual differences in the older adult population sample used in the experiment reported here, in that they believe that they have a strong memory for particular stimuli, expect to be able to remember those stimuli, and therefore have more confidence in their own memory. No other main effects or interactions were observed.

Participants did not respond to the higher incentive condition by volunteering fewer answers. A trend only was observed for age group and test format. Younger adults volunteered proportionately more responses in free report compared with older adults, which relates to an overall lower confidence in candidate responses. Recognition participants volunteered proportionately more responses in free report compared with cued-recall participants, which also relates to an overall lower confidence in candidate responses. This result differs from prior research where confidence has been observed to be higher for recognition than recall tests (e.g. Robinson, Johnson, & Herndon, 1997).

Although the older group gained significantly more than younger adults from the opportunity to screen out incorrect potential responses, they did so at the expense of greater losses in quantity correct. Koriat and Goldsmith's (1996) simulations of accuracy and quantity trade-off patterns as a function of response criterion setting found that accuracy increases as a linear function of response criterion, but that quantity performance decreases as a positively accelerated function of response criterion. They assumed perfect calibration and a uniform distribution of response probabilities. In the simulation, there was no cost of increased accuracy for relatively low response criteria, but as response criterion was raised, there were increasing costs.

People who believe that they have a strong memory for particular stimuli expect to be able to remember those stimuli and set a more demanding criterion for reporting their occurrence (Stretch & Wixted, 1998; Morrell et al., 2002). Conversely, people who believe the opposite (i.e. a weak memory) relax their criterion (Stretch & Wixted, 1998). Several results indicate age-related differences in measures of control (Aizpurua et al., 2009; da Silva & Sunderland, 2010; Koutstaal & Schacter, 1997; Pansky et al., 2009).Younger adults respond to a high incentive for accuracy by setting their response criterion higher (Koriat & Goldsmith, 1994, 1996; Kelley & Sahakyan, 2003; Pansky et al., 2009). Older adults adopt a more lenient response criteria than young adults (Aizpurua et al., 2009; da Silva & Sunderland, 2010; Kelley & Sahakyan, 2003; Koutstaal & Schacter, 1997; Pansky et al., 2009), and are less likely than younger adults to withhold answers when given the option of free report, even though this leads them to higher rates of false memory (Jacoby et al., 2005; Kelley & Sahakyan, 2003; Pansky et al., 2009). Therefore, a liberal report criterion is likely to result in a greater number of volunteered incorrect (and correct) answers.

In the present study, it would appear that the two groups chose different strategies in trading quantity for accuracy, with the older group placing greater emphasis on accuracy and the younger group greater emphasis on quantity of event information. Thus, partially in line with predictions, the older adult participants tended to be more conservative in their control policy. However, this result is also contrary to existing literature whereby older adults are less likely compared with younger adults to withhold answers when given the option of free report, even though this leads to higher rates of false memory (Jacoby et al., 2005; Kelley & Sahakyan, 2003; Pansky et al., 2009).

The results of the present study add to the literature regarding older adults' responsiveness to incentives for accuracy. However this warrants further study. Future research should explore incentives for accuracy in a within-subjects design.

Mediational analysis (Baron & Kenny, 1986) investigated whether individual differences in pre-test positive affect, and mental health predicted directly or indirectly to accuracy performance at free report beyond age, and memory monitoring and control processes. Regression analysis revealed that memory accuracy in free report in respect of action details was largely affected directly by age and through the indirect effects of quantity correct at forced report, and through individual differences in participant mental health on control policy. Memory accuracy in free report in respect of descriptive details was largely affected directly by age and through the indirect effects of control policy. Individual difference measures were not observed to directly predict accuracy over and above age.

3.10.3 Summary

Experiment 1 provides an extension of Koriat and Goldsmith's (1996) framework from answering general knowledge questions to answering episodic memory queries for details of an event. The framework and method produced parallel results in the current study. Accuracy in free report was much lower when participants reported action items compared with descriptive items (Koriat & Goldsmith, 1996, Experiment 2; Kelley & Sahakyan, 2003, Experiment 1). An age effect was observed as this pattern was more pronounced in older adult performance. Examination of the underlying metacognitive processes revealed that lower memory accuracy when reporting action items can be traced back to forced report, where retrieval in respect of action compared with descriptive items was lower, partly as a result of poorer encoding. Additionally, older adults volunteered fewer (correct and incorrect) responses which relates to the setting of an overall higher response criterion. Consequently the setting of a higher response criterion by older adults compounded the problem created by poor retrieval and led to lower memory accuracy in respect of action compared with descriptive items. From an applied perspective, this finding would imply that evaluators of witness testimony should be aware of such tendencies, and should be mindful of the relative high confidence placed by older adults in particular when reporting action details.

3.10.4 Applied Implications

Studies of free recall typically treat performance as a direct measure of retention without acknowledging the role of monitoring and control over responding that must occur to produce performance. Studies which compare performance across different retrieval monitoring conditions (e.g. Koutstaal, Schacter, Galluccio, & Stofer 1999; Multhaup, 1995) point to the dynamic quality of monitoring and control and its consequences for memory accuracy. The results of this study would appear to have implications for real-world forensic interviews. Witness information that can be relied upon is vital to criminal investigations. Under free report conditions, people tend to provide only information they believe to be correct so that their performance is mediated by a decision process used to avoid incorrect answers (Klatsky & Erdelyi, 1985; Koriat & Goldsmith, 1994). The results may inform forensic investigators in understanding why eyewitnesses report what they do (Brewer & Weber, 2008), as well as the potential contribution of witness monitoring and control processes to witness accuracy in free report situations. By using monitoring and control procedures witnesses actively *sift* the information provided in an initial free recall.

That older adults achieved comparable memory accuracy at free report as younger adults, but did so at the expense of quantity correct has potential implications for interviewing protocols. The Cognitive Interview (CI) and Enhanced Cognitive Interview (ECI), described in Chapter 1, are among the most rigorously tested and widely accepted methods for improving the accuracy and completeness of eyewitness reports. However, research that has assessed a CI's effectiveness to facilitate the event memory of older adults has resulted in inconsistent findings. As introduced in Chapter 1, this may be explained by the differing methodologies employed including different control interviews, different versions of the CI, small sample sizes and an older adult sample who may not be representative of the older population as a whole (Mello & Fisher, 1996). Additionally, because interviewers were aware of the experimental hypotheses, their expectations may have resulted in systematic differences in the length of interviews, or in the number and quality of questions asked in the interviews (Hayes & Delamothe, 1997).

In the current study, the setting of a more conservative response criterion by older adults resulted in the volunteering of fewer correct (and incorrect) responses. This suggests that the older adult sample believed that they had a strong memory for the to-be-remembered event and expected to be able to remember such details. It is for the investigators to decide beforehand on the purpose of the interview (i.e. whether accuracy or quantity of information is more important) and to advise older adult participants about their control policies accordingly.

The distinction between action and descriptive details is an important one for forensic investigations. Ibabe and Sporer (2004) note that decision makers are usually concerned with central details that are of legal relevance i.e. who did what to whom and can include the actions and descriptions of the main perpetrators. The results of the present study would appear to suggest that participants and particularly older adults may have difficulty in providing an accurate account of what happened within an event. An older adult may remember having seen a particular person before, as well as the actions involved in a crime, but may be unable to bind the culprit to the committed offense. As introduced in Chapter 1, older adults tend to commit more errors in identifying criminals, and these errors are most often related to false identifications. This is especially the case following relatively long delays (Havard & Memon, 2009; Memon & Bartlett, 2002; Memon & Gabbert, 2003a, 2003b; Rose et al., Searcy et al., 2000; Wilcock et al., 2005, 2007; Yarmey, 1993).

3.10.5 Methodology Limitations

There are several limitations in the present study that deserve some comment. Although longitudinal research suggests that memory deficits are particularly marked after age 60 (Rice, 1986) assigning older adults to a single group (e.g., Coxon & Valentine, 1997, List, 1986; Mello & Fisher, 1996; McMahon 2000) may be inappropriate as they are less sensitive to age-related trends. There is great variability among older adults, to the extent that some older adults are even more capable than some younger adults. The older adult participant sample (over 60 years) used in the experiment reported here may not have been representative of the general population. The older adult participant sample was recruited mainly from a University Lifelong Learning Centre, and as a result may have been more practised in recall and use of mnemonic devices. Additionally, they may not experience the same rate of memory decline and have more confidence in their memory abilities. To address this limitation, future research should recruit participants from the wider community rather than a university course.

Recall by older adults is affected by task meaningfulness, so that when experimental tests are perceived as irrelevant to daily life, motivation and performance may be impaired (Cockburn & Smith, 1991). In comparison, when they are highly motivated by tasks that are personally meaningful, older adults show a relatively small decrement (Stokes & Pankowski, 1988). In the current study participants were aware that they would need to remember the film, therefore in all probability made a deliberate effort to encode information (i.e. intentional learning occurred). Older adults' recall is optimal when they have sufficient time to process cognitive tasks, when their attention is undivided and when activities are uncomplicated (Anderson, 1999). In the present study participants were given as much time as they needed to complete each task. The researcher proceeded only when she was sure that all participants understood the task. Such conditions may have sufficiently enhanced the older adult performance relative to that of young adults and may contribute to the findings.

Comparing accuracy as a condition of question form depends highly on the way individual questions are constructed, as well as on the difficulty level of specific questions. Although attempting to formulate questions that were parallel across question forms, and foils in the Alternative–Forced-Choice test which were as plausible as possible, the external validity of this study cannot be guaranteed. Significant results did emerge on a number of the dependent variables in the expected direction however, which lends some support to the validity of the tests used and the populations evaluated.

Questionnaires and tasks were given in a fixed order, and this may have introduced carry-over effects. Koriat and Goldsmith (1996 – Experiment 2) counterbalanced the order of the two phases across participants. The results revealed that, in general, phase order had little or no effect, and the same pattern of results obtained in Experiment 1 were replicated. However, in the present study, order effects cannot be discounted.

As introduced in Chapter 1, the malleable nature of eyewitness memory during the interviewing process has been extensively demonstrated (cf. Loftus, 1975; 1989; Malpass, 1996). A person who is asked repeatedly to recall a witnessed event tends to gain, lose, and change details over time (Turtle & Yuille, 1994). A feature of the present study revealed that 1% of responses were changed from Phase 1 to Phase 2 in the absence of any feedback. Inconsistencies between first and second answers may occur simply as a result of memory failure (Gudjonsson, 2003). Alternatively, the repetition of questions without explanation may have communicated to participants that their previous responses were in some way incorrect or inappropriate and should be changed (Register & Kihlstrom, 1988). This hypothesis is examined further in Experiment 2 when participant response is assessed under different levels of performance feedback.

Chapter 4: Experiment 2

4.1 Introduction

As introduced in Chapter 2, judges, litigators, and legal scholars deem witness consistency to be one of the most important features of witness credibility (Brewer et al., 1999; Fisher et al., 2009; Potter & Brewer, 1999). However, the fallibility of human memory has been well documented in the literature (e.g., Bartlett, 1932; Clifford & Bull, 1978; Crombag et al., 1996; Cutler & Penrod, 1995; Jelicic et al., 2006; Johnson et al., 1993; Loftus & Palmer, 1974; Smith, 1930). Another issue is the reliability of evidence given during police questioning by individuals who may be particularly susceptible to suggestion (e.g., Gudjonsson & Gunn, 1982). Psychological vulnerabilities may place witnesses at a disadvantage in terms of coping with the demand characteristics of a police interview (and subsequent Court process) such that the credibility and reliability of their testimony is called into question.

Interrogative suggestibility as defined in Chapter 2, can be a serious psychological vulnerability during police investigative interviews (Gudjonsson, 2003; Gudjonsson, Sigurdsson, Einarsson, Bragason, & Newton, 2008; Gudjonsson, Young, & Bramham, 2007). Both situational factors (the nature of the interaction between circumstances / contextual factors, custodial factors, vulnerability factors, and support factors for example the presence of a lawyer or appropriate adults) and dispositional factors (vulnerability), may affect the capacity of an interviewee to cope with police interviews.

The Current Study

The experimental study described in this chapter follows those cited above in seeking to further develop the Gudjonsson and Clark (1986) theoretical model, by investigating the relationship between subjective memory beliefs, response change, compliance and objective memory performance across young, middle-aged and older adults.

Overview of Experiment

As previously described 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. This is in contrast to episodic memory required to recall past events. Of particular relevance to this research semantic memory is less sensitive than episodic memory to impairment by age (e.g., Nyberg et al., 2003). Some aspects of age-related cognitive decline are apparent throughout the adult lifespan, beginning in healthy educated adults when they are in their 20s and 30s (Salthouse, 2009), therefore the participant sample for this study was aged from 18 to 85 years. The present study used a mock witness paradigm. The stimulus material and questioning procedure were those used in McGroarty and Baxter (2007, 2009). The questions asked were not overtly misleading.

Participants completed The Metamemory in Adulthood Questionnaire (MIA; Dixon, et al., 1988) and four scales from the Memory Self – Efficacy Questionnaire (MSEQ; Berry et al., 1989) prior to attending the experiment. Participants completed the Positive and Negative Affect Scale (PANAS; Watson et al., 1988) then viewed the videotaped event of a non-violent robbery. Delayed testing took place after 10 minutes during which time all participants completed the Wechsler Test of Adult Reading (WTAR; Wechsler, 2001) and Participant Details Questionnaire. After watching the video the participants provided free recall. They were then questioned about the video-taped event, following which participants received either negative or neutral feedback. The questions were then repeated. Participants completed the Interview Rating Form, the Gudjonsson Compliance Scale (Gudjonsson, 1997) and the second PANAS **(**Watson, et al., 1988).

Experimental Aims

There were three main study aims. Firstly, the present study aimed to investigate the relationship between subjective memory beliefs, compliance and response change across young, middle-aged and older adults, irrespective of the type of feedback delivered. The results of Experiment 1 of this thesis revealed that 1% of responses were changed from Phase 1 to Phase 2 in the absence of any feedback. Therefore a second aim of the present study was to examine the effects of investigative pressure associated with neutral and negative feedback. The effects of age on response change scores following feedback were also investigated. As described in the opening paragraphs of this chapter, psychological vulnerabilities should not be interpreted in isolation to other factors such as context. Therefore, a third aim was to assess the relationship between the individual difference measures of years of education, crystallised intelligence, health status, sensory deficits, affective states, compliance and response change. As introduced in Chapter 1 of this thesis, significant amounts of age-related variance in memory performance are accounted for by health status (Albert et al., 1995; Nilsson et al., 1997), sensory deficits (McCoy et al., 2005; Tun et al., 2006; Wingfield et al., 2005), and affective states (Oaksford et al., 1996). In the experiment reported here participants provided self-report measures of physical and mental health, hearing and eyesight on a 5-point Likert-type scale, prior to experimental manipulation. Self-report measures of positive and negative affect were obtained on the PANAS (Watson et al., 1988) prior to and following experimental manipulation. Crystallised intelligence was measured by scores on the Wechsler Test of Adult Reading (Wechsler, 2001).

Experimental Hypothesis

This study had three main hypotheses. In light of previous findings detailed above, the first experimental hypothesis predicted that irrespective of type of feedback delivered, participants with poorer memory beliefs would exhibit higher levels of compliance and response change compared with participants who reported to have excellent memory capabilities. Following on from McGroarty and Baxter (2007, 2009), the second experimental hypothesis predicted that there would be a main effect of feedback. It was further predicted that negative feedback would be associated with more response change compared with neutral feedback, and with higher ratings of interview difficulty.

Older adults generally perceive more change in their memory, report lower levels of memory capacity, and perceive that they have less control over their memory than younger adults, therefore it could be hypothesised that compliance and response change might be more pronounced in older adult performance. However, subjective ideas about memory do not always correspond to objective memory performance (Ponds & Jolles, 1996; Ponds et al., 2000). For example, healthy older people (more than 55 years old) who rate their memory as very poor often exhibit normal performance on standard memory tasks. Polczyk et al., (2004) compared younger (mean 22.3 years) and older adults (mean 64.1 years) on interrogative suggestibility as measured by the GSS2 (Gudjonsson, 1987b, 1997; Polish version: Polczyk, 2000). Negative feedback was administered. Compared to younger adults, older adults scored higher on measures of Yield 1 and 2, but not on Shift. Regression analysis revealed that memory performance and memory self-assessment were independent predictors of yielding to suggestive questioning in the group of older adults, when controlling for the effects of age. The authors suggested that older adults may be less susceptible to negative feedback than younger adults (mean 22.3 years) because they are more self-confident and less dependent on authority. These conflicting results will be investigated in the present study.

4.2 Methodology

4.2.1 Design

The relationship between subjective memory beliefs, response change, compliance and objective memory performance was considered in terms of individual differences. Participants received either negative or neutral feedback. Participant self report ratings on the PANAS (Watson et al., 1988) were within-subject measures. Age was a continuous variable.

4.2.2 Participants

The participant sample consisted of 102 adults, mean age of 48.35 (SD = 20.92, range = 18 – 85 years). Participants aged less than 60 years (N = 61) were staff and students of the University of Strathclyde. Older adult participants (N = 41) were recruited from the University of Strathclyde Centre for Lifelong Learning, community organisations and social clubs within the Glasgow Area. Participants received £5.00 for participating in the study. From an original recruitment sample of 102, one older adult participant aged 85 did not complete the experiment in full. Data analysis was conducted on a sample size of 101. The research was approved by the standing Ethics Committee of the Department of Psychology, University of Strathclyde.

4.2.3 Materials

4.2.3.1 Participant details questionnaire

Participants completed a demographics questionnaire providing age, gender and years of education attained. Self-ratings were also provided in respect of everyday vision, hearing, physical and mental health on a 5-point Likert-type scale ranging from (1) *excellent* to (5) *very poor*. The participant details questionnaire is reproduced in Appendix D.

4.2.3.2 Self-report scales

Memory beliefs

See Experiment 1 – Methodology.

Participants completed the Metamemory in Adulthood Questionnaire (MIA; Dixon et al., 1988) and four scales from the Memory Self-Efficacy Questionnaire (MSEQ; Berry et al., 1989). The MIA questionnaire has been shown to be internally consistent with Cronbach's α ranging from .71 to .93 (Hultsch et al., 1987). The alpha reliabilities for each sub scale for the current sample were: Strategy (α = .80), Task (α = .80), Capacity (α = .77), Change (α = .90), Anxiety (α = .82), Achievement (α = .77), and Locus (α = .70). The MSEQ factors as a single scale (Berry et al., 1996), and shows similar age relationships across scales and moderate to high intercorrelations among scales (Berry et al., 1989; West & Berry, 1994). The alpha reliability for the current sample was (α = .91) when each individual question (N = 20) was entered into the analysis.

Positive and Negative Affect Scale (PANAS*;* Watson, et al., 1988)

See Experiment 1 – Methodology

The values of PA and NA ranged from a minimum of 10 to a maximum of 50. PANAS questionnaires were completed by all participants both at the outset of the test and on conclusion of the experimental manipulation. The reliabilities of the present data were good. The Cronbach's α value for PA at outset and conclusion was .85 and .89 respectively. The Cronbach's α value for NA at outset and conclusion was .84 and .81 respectively.

Wechsler Test of Adult Reading (WTAR; Wechsler, 2001)

See Experiment 1 – Methodology

Mini Mental State Exam (MMSE; Folstein et al., 1975)

See Experiment 1 – Methodology

4.2.4 Apparatus / Materials

4.2.4.1 Videotaped Event

The stimulus material consisted of a 77 second video-recording of a staged incident, depicting the theft of a briefcase (previously used by McGroarty & Baxter, 2007; 2009). There were three main characters; two men and one woman who conversed, and two male passers-by one of whom took a briefcase without the owner noticing. Using a technique based on the work of Allwood et al., (2005), the film stimulus was written-up in template form by the researcher.

The template consisted of 106 units of information which were further sub-divided into 74 units of person information, 6 units of object information, 4 units of surrounding information, and 22 units of dialogue information (The film summary template is included in Appendix H).

4.2.4.2 Event Memory Questionnaire (EMQ)

Normative data was obtained in a pilot study from 20 participants. The pilot study participants did not participate in the main experiment. Twenty nine questions were asked about the videotaped event. The questions 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?”). 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. (The questions are shown in Appendix I)

4.2.4.3 Interview Rating Form

This 8-item questionnaire required participants to rate their ease of recall, distraction, influence, comfort, concentration, reasonableness of questions, experience of pressure, and overall task difficulty – experienced during the questioning phase of the experiment – on a 5-point Likert-type scale, ranging from (1) *not at all* to (5) *very much so*. A total score varying from 8- 40 could be obtained with higher scores indicating greater interview difficulty. The Cronbach's α for the current sample was 0.79. (The interview rating form is shown in Appendix J).

4.2.4.4 The Gudjonsson Compliance Scale (GCS: Gudjonsson, 1997)

The GCS is a self-report inventory and consists of 20 statements which are answered *True* or *False*. Examples of the items are as follows: "I give in easily when I am pressured", "People in authority make me feel uncomfortable and uneasy", "I generally believe in doing as I am told" and "I try to please others". Each reply gives either a compliant or a non-compliant score. After re-coding items 17 to 19, a total GCS score ranging from 0-20 could be obtained by summing the number of *true* responses, with higher scores indicating more compliant behaviour. The Cronbach's α for the current sample was 0.54.

4.3 Procedure

Prior to coming into the laboratory, participants received an envelope in the mail that contained (a) The Metamemory in Adulthood Questionnaire (MIA; Dixon, et al., 1988) and (b) four scales from the Memory Self-Efficacy Questionnaire (MSEQ; Berry et al., 1989). Participants were asked to bring the completed questionnaires to the laboratory. Participants were tested individually. Prior to taking part in the experiment, all participants gave verbal consent. Participants then read and signed a consent form. The researcher (a female in her late 40's) conducted all experiments. Participants were given verbal and written instructions for all tasks. Older adults required just slightly more explanation of the tasks, and the interviewer proceeded only when she was sure that participants understood the tasks.

Participants completed the PANAS questionnaire (Watson et al., 1988) then viewed the videotaped event. Participants were informed that they would view a non violent scene containing people, actions, objects and dialogue. They were instructed to watch carefully and pay attention to everything that happened as if they were a real eyewitness (Campos & Alonso-Quecuty, 1999). The video was operated via a Panasonic VCR and shown on a 51cm screen Panasonic T.V. Participants were seated approximately four feet from the T.V. screen. The researcher was present in the testing room throughout the showing of the video. Delayed testing took place after 10 minutes during which time all participants completed the WTAR (Weschler, 2001) and Participant Details Questionnaire.

Throughout the testing phases of the experiment, the participant sat at a desk, facing the interviewer across from them. Time limits were not imposed; instead participants were advised to take as long as they needed to complete each task.

After playing the video the interviewer said to participants: “Tell me everything you can remember about the scene you witnessed on the video”. Following this free recall, twenty nine questions were available to be asked about the videotaped event. Participants were asked only those questions relevant to their accounts. Most participants (N = 60) reported seeing two passers-by in the background and so were asked the questions relating to them. The remaining participants (N = 41) stated that they saw only one passer-by or none. In these instances, the interviewer proceeded to the next appropriate question in the question set.

The interviewer’s 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 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”. The questions were then repeated.

In the neutral feedback condition, the interviewer said, “Thank you for answering these questions. To ensure that we have your answers recorded correctly, we’ll run through the questions once more”. The questions were then repeated.

Participants then completed the Interview Rating Form, the GCS (Gudjonsson, 1997) and the second PANAS (Watson et al., 1988). On completion of the experiment participants were de-briefed, thanked for their participation and paid £5.00. It was explained to participants in the conditions involving negative feedback that this did not accurately reflect their performance and was merely a feature of the experiment.

4.4 Scoring

Recall was a secondary measure, intended only to allow a basic comparison between the event-memory available to each group prior to experimental manipulation. Analysis revealed no significant differences between conditions. For simplicity of scoring, recall was based on answers given during the first round of questions, rather than on participant free recall. For the majority of questions, participants received one point for each correct answer. Five of the twenty nine questions; Q1, Q4, Q11, Q18, and Q21 asked for an estimate of either height or age. Scores on these questions ranged from 0 to 10 according to the accuracy of the estimate, with 10 being completely accurate. Where participants gave a rough estimate of age (and height) e.g., *Early 20s* or *Mid to late 30s* rather than stating a precise figure, it was not possible to calculate the difference between actual and estimated age.

Question 5 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 awarded 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 101 interviews. The rater was provided with the coding frame. Otherwise he 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*. Because the number of questions asked varied between participants, response change scores were computed as a percentage of the number of questions answered.

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 were: *Dark* to *Light* (where the correct answer was *Dark*), *Right* to *Left* (where the correct answer was *Right*), and *One of the passers-by* to *Don't know* (where the correct answer was *One of the passers-by*).

The number of changed responses toward accuracy was also scored. Participants received one point each time this occurred. Examples of responses increasing in accuracy were *Left* to *Right* (where the correct answer was *Right*) *Long* to *Short* (where the correct answer is *Short*). Further examples are provided in Appendix I. The number of responses increasing in accuracy and toward inaccuracy were computed both as a percentage of the number of questions answered and as a percentage of the number of changed responses.

4.5 Results

Analyses for the current study can be broadly categorised into analyses of individual difference measures, memory recall, and response change following feedback, which will be considered in turn. An individual differences analysis was used to assess the relationship between participant, education, crystallised intelligence, subjective memory beliefs, health, sensory deficits, affective states, compliance and response change, within an adult participant sample aged 18 to 85 years.

Individual difference measures are summarised in *Table 13* below, as well as *t*-tests concerning differences between the samples. Pre- and post-test measures of positive and negative affect were analysed by Analysis of Variance (ANOVA). The means and standard deviations (SD) of overall memory performance measures are provided in *Table 14*. Correlations were carried out between individual difference measures, memory performance, compliance and response change and are provided in *Tables 15* and *16*. The outcome of regression analysis is provided in *Tables* 17 and 18. The effects of neutral and negative feedback on memory performance were also investigated. The dependent variable of interest was the proportion of response changes following feedback, and the data are summarised in *Table* 20 below. The outcome of regression analysis in respect of neutral and negative feedback are provided at *Tables 25* and 26 below.

Kolmogorov-Statistical tests were computed on the distribution of scores on all performance measures prior to analysis. Where the assumptions to allow statistical analysis with parametric statistics, ANOVA were not met, non parametric equivalent analysis was carried out. Unless otherwise noted, the value of α for all statistical tests was set at (0.05), (Sheskin, 2007).

4.5.1 Individual Difference Measures

Table 13: Means and (SD) of individual difference measures by feedback condition (neutral; negative)

Variable	Neutral Feedback	Negative Feedback	*p*
N	55	46	
MMSE[a]	29.30 (.82)	29.41 (.80)	*ns*
WTAR[b]	42.40 (6.98)	41.83 (6.67)	*ns*
Education	15.78 (3.08)	16.02 (2.53)	*ns*
Eyesight	2.25 (1.06)	2.48 (1.01)	*ns*
Hearing	2.07 (1.03)	2.17 (.83)	*ns*
Physical Health	2.00 (.77)	1.78 (.79)	*ns*
Mental Health	1.65 (.89)	1.74 (.83)	*ns*
Positive Affect (1)	31.76 (6.70)	31.04 (6.50)	
Positive Affect (2)	31.53 (7.74)	30.50 (7.41)	
Negative Affect (1)	12.85 (4.40)	12.28 (3.67)	
Negative Affect (2)	12.07 (3.69)	11.54 (2.47)	
MSEQ Sel[c]	17.91 (3.06)	17.98 (2.48)	*ns*
MSEQ Sest[d]	66.71 (14.97)	66.28 (12.71)	*ns*
MIA[e]			
Strategy	61.04 (9.08)	62.48 (8.06)	*ns*
Task	61.15 (5.82)	59.43 (7.08)	*ns*
Capacity	54.76 (8.29)	52.43 (8.18)	*ns*
Change	52.96 (10.34)	50.00 (12.20)	*ns*
Anxiety	39.56 (8.94)	41.67 (6.83)	*ns*
Achievement	56.75 (7.15)	57.52 (7.87)	*ns*
Locus	29.31 (4.58)	29.07 (4.28)	*ns*
GCS[g]	9.05 (3.54)	9.39 (2.93)	*ns*
Interview Difficulty	15.60 (4.79)	18.13 (5.19)	$p < 0.05$

[a] Mini Mental State Examination (Folstein et al., 1975) score represents the number of points earned out of a possible total of 30.
[b] Wechsler Test of Adult Reading (Wechsler, 2001) score represents the total number of correctly pronounced irregular words out of a total of 50 words.
[c] Memory Self Efficacy Questionnaire, Self-efficacy Level (MSEQ; Berry et al., 1989).
[d] Memory Self Efficacy Questionnaire, Self-efficacy strength (MSEQ; Berry et al., 1989).
[e] Metamemory in Adulthood Questionnaire (MIA, Dixon et al., 1988), Scales: Strategy, Task, Capacity, Change, Anxiety, Achievement, Locus.
[g] Gudjonsson Compliance Scale (Gudjonsson, 1997)

Do self-report measures of affect change with feedback?

Pre- and post-test measures were taken of positive and negative affect. The data are summarised in *Table 13* above. A feedback (neutral, negative) x positive affect mixed model ANOVA was carried out on the pre- and post-test positive affect scores. The results revealed no significant main effects of feedback, positive affect, or interaction ($p > 0.05$).

The same analysis was carried out on the pre- and post-test negative affect scores. Box's Test of Equality of Covariances was significant ($p < 0.001$). Levene's Test of Equality of Error Variances was significant for both pre- and post-test negative affect scores ($p < 0.01$). The value of α for this analysis was therefore set at (0.001), (Stevens, 1996). The results revealed no significant main effects of feedback, negative affect, or interaction ($p > 0.001$).

Gudjonsson Compliance Scale (GCS; Gudjonsson, 1997)

Do measures of compliance vary with type of feedback?

The item ratings were summed to give a measure of compliance. Scores ranged from 3 to 18. The means and standard deviations (SD) of measures of compliance are summarised in *Table 13* above. Tests of normality were carried out on the GCS scores. The Kolmogorov-Statistic was not significant ($p > 0.05$), indicating a normal distribution. An independent-samples *t*- test carried out revealed no significant differences in compliance between feedback conditions ($p > 0.05$).

Do ratings of interview difficulty vary with type of feedback?

The item ratings were summed to give an overall measure of interview difficulty. Scores ranged from 9 to 28. The means and standard deviations (SD) of interview difficulty scores are summarised in *Table 13* above. Tests of normality were carried out on the interview difficulty scores. The Kolmogorov-Statistic was significant ($p < 0.05$), suggesting a violation of the ANOVA assumption that observations are normally distributed on the dependent variable. No observations with scores in excess of 3.29 standard deviations from the mean ($p < 0.001$) were recorded. As discussed on pages 115 above the distribution of interview difficulty scores were considered normal and no further action was necessary (Tabachnick & Fidell, 2001).

An independent-samples ANOVA was carried out on the ratings of interview difficulty. The main between-subjects effect of feedback was significant (F (df, 1, 99) = 9.47, MS Error = 22.77, $p < 0.02$, $\eta^2 = .06$). Participants receiving negative feedback reported significantly higher levels of interview difficulty (M = 18.13) compared with those receiving neutral feedback (M = 15.60). When the analysis was computed with self-report measures of vision included as a covariate (see p. 206), the result remained significant (F (df, 1, 99) = 5.60, MS Error = 24.25, $p = 0.02$, $\eta^2 = .05$), although the value of p and the effect size was reduced.

4.5.2 Memory Recall

Recall was a secondary measure, intended only to allow a basic comparison between the event memory available to each group prior to experimental manipulation. The mean recall scores in response to open-ended and closed questions by feedback condition are summarised in *Table 14* below. Tests of normality were carried out on measures of total memory recall, open-ended and closed questions. The Kolmogorov-Statistics were not significant ($p > 0.05$), indicating a normal distribution.

Table 14: Means and (SD) of memory recall, by feedback condition (neutral; negative)

Variable	Neutral Feedback ($N = 55$)	Negative Feedback ($N = 46$)	p
Recall (Closed Items)	12.42 (4.39)	11.98 (4.13)	*ns*
Recall (Open-ended Items)	20.96 (7.57)	19.93 (7.52)	*ns*
Total Recall	33.38 (10.69)	31.87 (10.11)	*ns*

Separate independent-samples *t*- tests carried out on recall measures revealed no significant differences between feedback conditions ($p > 0.05$).

Of 2558 responses recorded, a total of 306 responses (approx. 12%) were changed across the two feedback conditions. Of these 114 were changed from accuracy to inaccuracy and 148 changed towards accuracy. The remaining 44 responses changed neither towards nor away from accuracy. As the number of questions answered varied between participants (from 15 to 29) the response change scores were computed as a percentage of the number of questions answered.

To assess the relationship between the individual difference measures and measures of objective memory performance, compliance and overall response change scores, partial correlations were computed. The correlation matrix is included in *Tables 15 and 16* below. The correlations were computed after partialling out age.

Table 15: Correlation matrix between individual difference measures, memory recall, compliance and response change

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1.	Strategy	1.00	.15	.05	-.18	**.24**	**.20**	.06	.05	.07	**.30**	-.08	.06	-.10	-.05	-.08
2.	Task		1.00	-.04	-.03	.03	**.21**	.01	.03	.11	.02	-.18	-.03	.14	.09	**-.31**
3.	Capacity			1.00	**.48**	**-.39**	.15	.08	**.29**	**.31**	.00	-.14	-.08	.00	-.03	.03
4.	Change				1.00	**-.39**	-.06	**.23**	.09	.15	**-.23**	-.14	.03	.13	.11	.04
5.	Anxiety					1.00	**.22**	.18	**-.22**	**-.24**	**.34**	.14	-.01	-.17	-.12	.03
6.	Achievement						1.00	.15	-.14	.02	.18	-.02	-.05	-.05	-.05	-.05
7.	Locus							1.00	-.01	.04	.09	-.06	-.01	-.07	-.06	-.03
8.	MSEQ Sel								1.00	**.77**	**.03**	-.01	.19	.18	**.21**	-.11
9.	MSEQ Sest									1.00	-.01	-.18	.10	**.21**	.19	-.14
10.	GCS										1.00	.03	.03	-.16	-.10	-.08
11.	Interview Difficulty											1.00	-.05	-.05	-.05	.16
12.	Rec. Closed												1.00	**.54**	**.79**	-.17
13.	Rec. Open													1.00	**.94**	-.15
14.	Total Rec.														1.00	-.18
15.	Response Change															1.00

All correlations in bold type-face were significant at *p* < 0.05 or better.
1. Metamemory in Adulthood Questionnaire (MIA; Dixon et al., 1988) Scales: Strategy, 2. Task, 3. Capacity, 4. Change, 5. Anxiety, 6. Achievement, 7. Locus .
8. Memory Self Efficacy Questionnaire (MSEQ; Berry et al., 1989) - Self-efficacy level, 9. Self-efficacy strength; Gudjonsson Compliance Scale (GCS; Gudjonsson, 1997);
11. Self-report ratings of Interview Difficulty; 12. Closed Qs – Recall; 13. Open Qs – Recall; 14. Total Recall; 15. Response Change

Table 16: Correlation matrix between individual difference measures, memory recall, compliance and response change

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1.	WTAR	1.0	**.37**	-.06	.02	.07	**.22**	**-.36**	-.07	**-.32**	-.12	.03	-.06	.07	.06	.07	.05
2.	Education		1.0	-.15	.02	-.06	.00	**-.27**	.03	-.14	.04	.07	-.09	.05	.04	.06	-.12
3.	Vision			1.0	.10	.15	.12	-.07	-.11	-.03	-.00	.01	**.23**	-.09	.03	-.01	.05
4.	Hearing				1.0	.13	.19	-.03	.16	-.07	**.23**	.11	.07	-.15	-.18	-.19	.01
5.	Physical Health					1.0	**.50**	-.01	.06	.05	.04	.01	.09	-.09	-.06	-.08	-.07
6.	Mental Health						1.0	**.24**	.12	-.19	.09	.08	.16	-.02	-.09	-.07	.07
7.	Positive Affect (1)							1.0	.07	**.83**	.12	-.07	-.03	.06	.08	.08	-.18
8.	Negative Affect (1)								1.0	.14	**.76**	.12	.09	.06	-.15	-.08	-.13
9.	Positive Affect (2)									1.0	**.22**	.01	-.08	.05	.10	.09	**-.22**
10.	Negative Affect (2)										1.0	.17	.09	.06	-.08	-.03	-.08
11.	GCS											1.0	.03	.03	-.16	-.10	-.08
12.	Interview Difficulty												1.0	-.05	-.05	-.05	.16
13.	Rec. Closed													1.0	**.54**	**.79**	-.17
14.	Recall Open														1.0	**.94**	-.15
15.	Total Recall															1.0	-.18
16.	Response Change																1.0

All correlations in bold type-face were significant at *p* < 0.05 or better.

1. Wechsler Test of Adult Reading (Wechsler, 2001); 2. Years of education attained; 3. Vision - Self-report measure on 5-point scale; 4. Hearing – self-report measure on 5-point scale; 5. Physical Health – self-report measure on 5 point scale; 6. Mental Health – self-report measure on 5-point scale; 7 – 10 Self-report measures of Positive and Negative Affect (PANAS; Watson et al., 1988); 11. Gudjonsson Compliance Scale (GCS, Gudjonsson, 1997); 12. Self-report ratings of Interview Difficulty; 13. Closed Qs – Recall; 14. Open Qs – Recall; 15. Total Recall; 16. Response Change.

Scales of the Metamemory in Adulthood (MIA) Questionnaire correlated with different variables. Of particular interest, scores on the strategy (r = .30, p < 0.005), and anxiety (r = .34, p = 0.001) subscales were significantly positively correlated with compliance scores. While scores on the the change subscale (r = -.23, p < 0.05) were significantly negatively correlated with compliance scores. In other words, the more confident participants were in their own memory the lower their compliance scores. Scores on the task subscale were significantly negatively correlated with response change scores (r = -.31, p < 0.005), showing the more knowledge participants had of memory processes the lower their response change scores. Memory self-efficacy level was significantly positively correlated with total memory recall (r = .21, p < 0.05), while memory self-efficacy strength was significantly positively correlated with memory recall in response to open-ended questions (r = .21, p < 0.05). In other words the greater the belief in their memory ability, the greater the participant's memory recall performance. Self-report measures of vision were positively correlated with participant ratings of interview difficulty (r = .23, p < .05), showing the poorer the participants' eyesight the more difficult they found the interview. Pre-test positive affect was positively correlated with post-test positive affect (r = .83, p < 0.001). Post-test positive affect was positively correlated with post-test negative affect (r = .22, p < 0.05), and negatively correlated with response change scores (r = -.22, p < 0.05). Pre-test negative affect was positively correlated with post-test negative affect (r = .76, p < 0.001). The number of closed questions recalled was positively correlated with the number of open questions recall (r = .54, p < 0.001) and the total number of items recalled (r = .79, p < 0.001). The number of open questions recalled was positively correlated with the total number of items recalled ($r = .94$, $p < 0.001$). No significant correlation was observed between recall, compliance and response change scores ($p > 0.05$).

4.5.3 Predictors of Compliance and Response Change Scores

To assess the unique contribution of participant scores on the MIA subscales of strategy, anxiety and change on compliance scores, multiple regression analysis was performed. Age and participant scores on the strategy, anxiety and change subscales were included as predictors. *Table 17* below gives the outcome of the regression analysis.

Table 17: Regression Analysis – Predictors of compliance scores

		B	*SE B*	*β*	*p*
Model					
	Constant	2.68	3.92		
	Age	-.02	.02	-.15	$p > 0.05$
	Strategy	.09	.04	.25	$p < 0.05$
	Anxiety	.10	.04	.25	$p < 0.05$
	Change	-.03	.03	-.11	$p > 0.05$

$\Delta R^2 = .14$

To explain compliance scores participant scores on the MIA subscales of strategy and anxiety only were identified as unique predictors.

To assess the unique contribution of age, participant scores on the MIA subscale of task, and post-test measures of positive affect on the overall response change scores, multiple regression analysis was performed. Age, scores on the task subscale and post-test measures of positive affect were included as predictors. *Table 18* below gives the outcome of the regression analysis.

Table 18: Regression Analysis – Predictors of Overall Response Change Scores

		B	*SE B*	β	α
Model					
	Constant	57.90	12.53		
	Age	-.10	.06	-.16	$p > 0.05$
	Task	-.55	.20	-.28	$p < 0.05$
	Positive Affect (2)	-.26	.16	-.15	$p > 0.05$

$\Delta R^2 = .11$

To explain overall response change scores, participant scores on the MIA subscale of task only were identified as unique predictors.

4.5.3.1 Summary of Preliminary Analysis

Preliminary analysis revealed a significant difference in participant ratings of interview difficulty by feedback condition. Participants perceived the interview to be significantly more difficult following negative feedback**.** No other significant differences were observed in the individual difference variables by feedback condition. Participant scores on the MIA subscales correlated with different variables. The scales of strategy ($r = .30$, $p < 0.005$) and anxiety ($r = .34$, $p = 0.01$) were positively correlated with GCS scores. Scores on the change scale were negatively correlated with GCS scores ($r = -.23$, $p < 0.005$). Scores on the task subscale were negatively correlated with response change ($r = -.31$, $p < 0.01$). Self-report measures of post-test positive affect ($r = -.22$, $p < .05$) were negatively correlated with response change. Self-report measures of vision were positively correlated with interview ratings ($r = .23$, $p < 0.05$), showing that the poorer the participant believed their vision to be, the more difficult the interview was perceived. Regression analysis revealed that participants' scores on the MIA subscales of strategy and anxiety were unique predictors of compliance scores. While participants' scores on the task scale (knowledge of basic memory processes) were unique predictors of overall response change scores irrespective of the type of feedback.

4.5.4 What is the relationship between subjective memory beliefs, compliance and response change scores?

The first aim of the present study was to investigate the relationship between subjective memory beliefs, compliance and response change across young, middle-aged and older adults, irrespective of the type of feedback delivered. The results detailed above show that beliefs about one's own memory were negatively related to compliance and response change scores. This shows that the more confidence participants had in their own memory the lower the compliance and response change scores. Age was not a predictor of compliance or overall response change scores. This finding was contrary to expectations. The means and standard deviations (SD) of participant scores on the MIA subscales of strategy, anxiety and task by age group are summarised in *Table 19* below.

Table 19: Means and (SD) of participant scores on the MIA subscales of strategy, anxiety and task by agegroup				
Age Group	(N)	Strategy	Anxiety	Task
18 - 24	19	57.00 (9.38)	42.11 (8.39)	62.47 (5.44)
25 - 34	15	63.93 (6.43)	41.53 (6.29)	61.73 (9.93)
35 – 44	13	58.08 (11.90)	38.92 (8.18)	59.69 (5.17)
45 – 54	12	64.83 (5.94)	41.25 (9.37)	59.92 (8.48)
55 – 64	12	62.50 (6.46)	40.42 (7.19)	58.58 (6.69)
65 – 74	19	63.68 (8.06)	38.95 (7.35)	59.11 (4.37)
75+	11	63.27 (8.08)	40.36 (11.14)	60.27 (3.32)

Separate independent samples ANOVAs were carried out on self-report measures of strategy, anxiety and task. The results revealed no significant differences between agegroups on any of the measures ($p > 0.05$).

4.5.5 Feedback and Response Change

Does the total number of response changes vary with type of feedback?

Of 2558 responses recorded, a total of 306 responses (approx. 12%) were changed. The means and standard deviations (SD) and direction of response change (actual and as a percentage of the number of questions answered) relating to the type of feedback are summarised in *Table 20* below.

Table 20: Means and (SD) of response change (RC), and as a percentage of the number of questions answered away from and toward accuracy scores, by type of feedback (neutral; negative)

Variable	Neutral Feedback (N = 55)	Negative Feedback (N = 46)
RC	2.29 (1.99)	3.91 (3.27)
RC – Away from Accuracy	0.82 (0.91)	1.50 (1.41)
RC – Toward Accuracy	1.13 (1.20)	1.87 (2.02)
RC as % of No. of Questions Answered	8.73 (7.01)	16.15 (16.54)
RC – Away as % of No. of Questions Answered	3.16 (3.36)	5.93 (5.64)
RC – Toward as % of No. of Questions Answered	4.24 (4.44)	7.92 (10.38)

Tests of normality were carried out on the total number of changed responses. The Kolmogorov-Statistic was significant ($p < 0.05$), suggesting a violation of the ANOVA assumption that observations are normally distributed on the dependent variable. No observations with scores in excess of 3.29 standard deviations from the mean ($p < 0.001$) were recorded therefore as described on page 115 above, distributions were considered normal and no further action was necessary (Tabachnick & Fidell, 2001).

An independent-samples ANOVA was carried out on the total response change scores. Levene's Test of Equality of Error Variances was significant ($p = 0.001$). The main between-subjects effect of feedback was significant (F (df, 1, 99) = 9.12, MS Error = 151.16, $p = 0.003$, $\eta^2 = .08$). Regardless of whether the response change was away from or towards accuracy, participants receiving negative feedback changed significantly proportionately more responses (M = 16.15, SE 1.81) compared with those receiving neutral feedback (M = 8.73, SE 1.66).

As previously discussed, participant scores on the MIA task subscale were negatively correlated with response change scores ($r = -.30$, $p < 0.01$). Also, self-report measures of post-test positive affect $r = -.24$, $p < .05$) were negatively correlated with response change. In other words, the higher the participants' level of post-test positive affect the lower the response change scores.

When the independent-samples ANOVA, reported above, was computed with self-report measures of post-test positive affect and the MIA subscale of task included as covariates, the result remained significant (F (df, 1, 99) = 7.32, MS Error = 139.69, $p = 0.008$, $\eta^2 = .07$), although the value of p was reduced.

Does the response change toward inaccuracy scores vary with type of feedback?

The means and standard deviations (SD) of response change toward inaccuracy scores are described in *Table 20* above. Scores ranged from 0 to 20%.

Tests of normality were carried out on the changed responses away from accuracy scores as a percentage of the number of questions answered. The Kolmogorov-Statistics were significant ($p < 0.05$), suggesting a violation of the ANOVA assumption that observations are normally distributed on the dependent variable. One observation was recorded with a score in excess of 3.29 standard deviations from the mean ($p < 0.001$). As discussed previously, statistical analysis was conducted both with the outlying score included and excluded (Stevens, 1996).

An independent-samples *t*-test was carried out. Participants receiving negative feedback, changed significantly proportionately more responses away from accuracy (M = 5.93) than those receiving neutral feedback (M = 3.16), (t (99) = -3.05, $p < 0.001$). A separate independent-samples *t*-test was carried out with the extreme score deleted. The result remained the same (t (98) = -2.82, $p = 0.001$).

Does the response change toward accuracy scores vary with type of Feedback?

The means and standard deviations (SD) of response change toward accuracy scores are described in *Table 20* above. Scores ranged from 0 to 46.7%.

Tests of normality were carried out on the number of response change towards accuracy scores as a percentage of the number of questions answered. The Kolmogorov-Statistics were significant ($p < 0.05$), suggesting a violation of the Analysis of Variance assumption that observations are normally distributed on the dependent variable. Two observations were recorded with scores in excess of 3.29 standard deviations from the mean ($p < 0.001$). As previously discussed analysis was conducted both with the extreme score included and excluded (Stevens, 1996).

An independent-samples *t*-test was carried out. Participants receiving negative feedback, changed significantly proportionately more responses toward accuracy (M = 7.92) compared with those receiving neutral feedback (M = 4.24), (t (99) = -2.39, $p < 0.005$). A separate independent-samples *t*-test was carried out with the two extreme scores deleted. The result remained the same (t (97) = -1.78, $p < 0.05$).

Response change as a percentage of the number of changed responses.

The response change away from and towards accuracy scores were also computed as a percentage of the response change scores The means and standard deviations (SD) of response change scores away from and towards accuracy as a percentage of the total response change scores, as it relates to the type of feedback are summarised in *Table 21* below. At times, scores were incalculable (e.g., when no changed responses away from or toward accuracy were recorded) and the data was not included in the analysis. This occurred in the recall of 10 neutral feedback and 6 negative feedback participants.

Table 21: Means and (SD) of response change (RC), away from accuracy scores by type of feedback (neutral; negative), as % of total response change scores

Variable	Neutral Feedback ($N = 45$)	Negative Feedback ($N = 40$)
RC – Away as % of total Response Change Scores	39.04 (37.41)	42.10 (31.96)
RC – Toward Accuracy as % of total Response Change Scores	46.28 (37.32)	45.59 (30.31)

Tests of normality were carried out on the number of response change scores away from accuracy as a percentage of the total number of changed responses. The Kolmogorov-Statistics were significant ($p < 0.05$), suggesting a violation of the Analysis of Variance assumption that observations are normally distributed on the dependent variable. No observations with scores in excess of 3.29 standard deviations from the mean ($p < 0.001$) were recorded, therefore as discussed on page 115 above the distribution of the response change scores away from accuracy as a percentage of the total response change scores were considered as normal and no further action was necessary (Tabachnick & Fidell, 2001).

An independent-samples *t*-test revealed no significant differences between the feedback conditions ($p > 0.05$). The response change away from accuracy scores as a percentage of the response change scores did not vary as a condition of feedback.

Tests of normality were carried out on the response change towards accuracy scores as a percentage of the total response change scores. The Kolmogorov-Statistics were significant ($p < 0.05$), suggesting a violation of the ANOVA assumption that observations are normally distributed on the dependent variable. No observations with scores in excess of 3.29 standard deviations from the mean ($p < 0.001$) were recorded, therefore as discussed on page 115 above the distribution of the number of response change toward accuracy scores as a percentage of the total number of response change scores were considered normal and no further action was necessary (Tabachnick & Fidell, 2001).

An independent-samples *t*-test revealed no significant differences between the feedback conditions ($p > 0.05$). The response change toward accuracy scores as a percentage of the total response change did not vary as a condition of feedback.

4.5.6 Age, Feedback and Response Change

Do age and type of feedback affect the response change scores?

A further aim of the present study was to examine the effect of age on response change scores. Age was a continuous variable and the participants were aged from 18 to 85. Descriptive data of the number of response changes and as a percentage of the number of questions asked by age group and feedback condition are provided in *Tables 22 and 23* below

Table 22: Means and (SD) of response change, by age group and type of feedback (neutral; negative)

Age Group	(*N*)	Neutral Feedback	(*N*)	Negative Feedback
18 – 24	9	1.44 (1.42)	10	6.00 (4.16)
25 – 34	9	1.89 (1.17)	6	4.67 (2.94)
35 – 44	7	3.86 (1.35)	6	2.00 (1.41)
45 – 54	7	1.86 (1.77)	5	4.40 (3.44)
55 – 64	3	3.33 (4.16)	9	4.00 (3.35)
65 – 74	12	2.67 (2.77)	7	2.14 (2.27)
75 +	8	1.75 (0.71)	3	2.33 (1.53)

Table 23. Means and (SD) of response change (%), by age group and type of feedback (neutral; negative)

Age Group	(*N*)	Neutral Feedback	(*N*)	Negative Feedback
18 – 24	9	5.10 (4.88)	10	25.76 (22.91)
25 – 34	9	7.49 (4.02)	6	17.37 (11.73)
35 – 44	7	14.08 (3.93)	6	7.26 (4.96)
45 – 54	7	6.87 (6.57)	5	16.05 (11.31)
55 – 64	3	11.49 (14.35)	9	19.77 (21.40)
65 – 74	12	10.81 (9.89)	7	7.70 (7.65)
75 +	8	6.99 (2.33)	3	8.41 (4.80)

Between-subjects parametric statistical analysis by ANOVA is not appropriate on the above data. Age is a continuous variable. Separate correlations were computed to assess the impact of age on response change scores (%) by type of feedback. Age was positively correlated with response change ($r = .13$, $p > 0.05$, *ns*), following neutral feedback, showing the older the participant the higher the response change score. In contrast age was negatively correlated with response change ($r = -.26$, $p = 0.08$, *ns*), following negative feedback. This shows that the older the participant the lower the response change score.

The means and standard deviations (SD) of interview difficulty ratings by age group following negative feedback are summarised in *Table 24* below.

Table 24: Means and (SD) of interview difficulty ratings by age group following negative feedback (N = 46)

Age group	(N)	Int. Difficulty
18 – 24	10	21.20 (5.33)
25 – 34	6	18.00 (3.69)
35 – 44	6	17.67 (4.46)
45 – 54	5	17.40 (5.73)
55 – 64	9	16.56 (4.55)
65 – 74	7	18.57 (6.08)
75+	3	14.00 (7.00)

Younger adults aged 18 – 24 perceived the interview to be more difficult following negative feedback. Parametric statistical analysis was not appropriate on the above data.

4.5.6.1 Summary of response change by feedback condition

Negative feedback was associated with a significantly greater number of response changes, and a greater tendency of participants to change their answers away from or towards accuracy (see *Table 20* above). These results did not vary significantly with age. However, a non significant interactive pattern was observed (see *Tables 22* and *23* above). Age was positively correlated with response change following neutral feedback, showing the older the participant, the greater the tendency to change responses. In contrast age was negatively correlated with response change following negative feedback, showing the younger the participant the higher response change scores. Younger adults aged 18 – 24 rated the interview as more difficult following negative feedback.

4.5.7 Mediational Analysis

Irrespective of the type of feedback, negative correlations were observed between scores on the MIA task subscale (r = -.31, p < 0.01) and measures of post-test positive affect (r = -.22, p < .05) and overall response change. To explain their respective contribution to response change scores by type of feedback, separate multiple regression analyses were carried out. Scores on the MIA subscale of task and measures of post-test positive affect were entered as predictors. The outcome of the regression analyses is provided in *Tables 25* and *26* below.

Table 25: Regression analysis – predictors of response change following neutral feedback

		B	*SE B*	β	*p*
Model					
	Constant	27.85	9.81		
	Task	-.31	.16	-.26	$p \leq 0.05$
Model					
	Constant	32.40	9.63		
	Task	-.25	.16	-.21	$p > 0.05$
	Positive Affect (2)	-.27	.12	-.30	$p < 0.05$

$R^2 = .07$; $\Delta R^2 = .05$; $\Delta R^2 = .12$

Participant scores on the subscale of task and post-test measures of positive affect were identified as unique predictors of response change following neutral feedback.

Table 26: Regression analysis – predictors of response change following negative feedback

		B	*SE B*	β	*p*
Model					
	Constant	53.52	20.29		
	Task	-.63	.34	-.27	$p > 0.05$
Model					
	Constant	58.71	21.10		
	Task	-.56	.35	-.24	$p > 0.05$
	Positive Affect (2)	-.30	.33	-.14	$p > 0.05$

$R^2 = .07$; $\Delta R^2 = .05$; $\Delta R^2 = .05$

Neither scores on the subscale of task nor post-test measures of positive affect were identified as unique predictors of response change following negative feedback.

4.6 Results Summary

Preliminary analysis revealed a significant difference in participant ratings of interview difficulty by feedback condition. Participants perceived the interview to be significantly more difficult following negative feedback**.** No other significant differences were observed in the individual difference variables by feedback condition. Participant scores on the MIA subscales correlated with different variables. The scales of strategy ($r = .30$, $p < 0.005$) and anxiety ($r = .34$, $p = 0.01$) were positively correlated with GCS scores. Scores on the change scale were negatively correlated with GCS scores ($r = -.23$, $p < 0.005$). Scores on the task scale were negatively correlated with response change ($r = -.31$, $p < 0.01$). Self-report measures of post-test positive affect ($r = -.22$, $p < .05$) were negatively correlated with response change. Self-report measures of vision were positively correlated with interview ratings ($r = .23$, $p < 0.05$), showing that the poorer the participant believed their vision to be, the more difficult the interview was perceived. Regression analysis revealed that participants' scores on the MIA subscales of strategy and anxiety were unique predictors of compliance scores. While participants' scores on the task scale (knowledge of basic memory processes) were unique predictors of overall response change scores irrespective of the type of feedback.

Negative feedback was associated with a significantly greater number of response changes, away from and towards accuracy (see *Table 20* above). A non significant interactive pattern was observed as to the effect of age following feedback (see *Tables 22* and *23* above). Age was positively correlated with response change following neutral feedback, showing the older the participant, the greater the tendency to change responses. In contrast age was negatively correlated with response change following negative feedback, showing the younger the participant the higher response change scores. Younger adults aged 18 – 24 rated the interview as more difficult following negative feedback (see *Table 24*).

Regression analysis revealed that participant scores on the MIA subscale of task and measures of post-test positive affect were significant predictors of response change following neutral feedback, accounting for approximately 12% of the variance. No significant predictors of response change following negative feedback were identified (see *Tables 25* and *26* above).

4.7 Discussion

There were three main study aims. First, the present study aimed to investigate the relationship between subjective memory beliefs, compliance and response change across young, middle-aged and older adults, irrespective of the type of feedback delivered. A second aim was to examine the effects of investigative pressure associated with neutral and negative feedback. The effects of age on response change scores following feedback were also investigated. As described in the opening paragraphs of this chapter, psychological vulnerabilities should not be interpreted in isolation to other factors such as context. Therefore, a third aim was to assess the relationship between the individual difference measures of education, crystallised intelligence, health status, sensory deficits, affective states and response change.

The first experimental hypothesis was that irrespective of the type of feedback, participants with poor memory beliefs would exhibit higher levels of response change and compliance compared with those who reported having excellent memory capabilities. Furthermore, as older adults generally perceive more changes in their memory, report lower levels of memory capacity, and perceive that they have less control over their memory than younger adults, compliance and response change would be more pronounced in older adults. The results obtained partly support this experimental hypothesis.

First, beliefs about one's own memory were related to compliance, supporting the stated hypothesis and the findings of Van Bergen et al., (2009). Scores on the subscales of the Metamemory in Adulthood (MIA) Questionnaire correlated with different variables (see *Table 15*). The MIA has 108 items distributed among 7 factor derived scales that are answered on a 5-point scale. Overall, the scale is a measure of knowledge, beliefs, and affect about memory. The seven scales are strategy, task, capacity, change, anxiety, achievement and locus. A significant positive correlation was obtained between the subscale of strategy and compliance, showing that the greater the use of external memory aids, the higher participants' compliance scores. A significant negative correlation was obtained between the subscale of change and compliance showing that the greater the perceived stability in remembering capacities, the lower participants' compliance scores. A significant positive relationship was obtained between the subscale of anxiety and compliance, showing that the more memory anxious participants were the higher their compliance scores. Regression analysis revealed that scores on the subscales of anxiety and strategy only were identified as unique predictors of compliance. The model accounted for approximately 14% of the variance in compliance scores (see *Table 17*). Contrary to expectations, the results reported here did not vary with age.

As described in Chapter 2, compliance is a tendency of the individual to go along with propositions, requests or instructions for some immediate instrumental gain (Gudjonsson, 2003). Gudjonsson (1990) found a significant correlation between compliance and suggestibility, and Gudjonsson (2003) argues that the two constructs overlap. Compliance may be viewed as an ineffective coping mechanism during tasks perceived as stressful (Costa & McCrae, 1992). Gudjonsson, (2003) argued that people with subjective memory problems tend to rely on external cues and suggestions, making them vulnerable to false memories.

The results of this study support this line of reasoning. Participants who rated themselves as memory anxious were significantly more likely to report compliance. One explanation is that people suffering from memory distrust (i.e. who evaluate their memory as poor) are more susceptible to the authority of others. Within the context of a forensic interview, witnesses may perceive the opinion of the interrogator as more important than their own. However, because this relationship was based only on correlations, the causal relationship between these two concepts cannot be established. It could also be argued that people start to distrust their memory because they are more easily intimidated by authorities.

Furthermore, following Gudjonsson's reasoning, it was predicted that participants with negative memory beliefs would display greater response change following feedback. The results would also appear to support the stated hypothesis. A significant negative relationship was obtained between the MIA sub-scale of task and response change scores, showing the higher the participant's level of knowledge of basic memory processes and functions, the lower the response change scores. This pattern of results is contrary to the findings of Van Bergen et al., (2009) who found no significant correlations between the Squire Subjective Memory Questionnaire (SSMQ; Squire et al., Van Bergen et al., 2009) and suggestibility (Yield or Shift) as measured by the Dutch translation of the Gudjonsson Suggestibility Scale (Gudjonsson, 1983, 1984; Merckelbach et al., 1998).

The use of differing experimental stimuli to measure participant memory beliefs may account for the different pattern of results. The SSMQ is a self-report questionnaire, consisting of 18 items that are answered on a 9-point scale (- 4 = disastrous, + 4 = perfect). Sample items are "My ability to recall things when I really try is". Scores are summed to obtain a total SSMQ score (varying from -72 to 72), with a negative score indicating a negative subjective evaluation of one's own memory. Trait memory distrust manifests itself as a personality trait. The SSMQ taps a quality that is more trait than state. In contrast, the present study assessed memory beliefs with the Metamemory in Adulthood Questionnaire (MIA; Dixon et al., 1988). As discussed above, the MIA has 108 items distributed among 7 factor derived scales that are answered on a 5-point scale. Of particular relevance is the subscale of task. Sample items are "For most people, facts that are interesting are easier to remember than facts that are not". Scores represent knowledge of basic memory processes and functions. The MIA may perhaps be more sensitive to individual variations in subjective knowledge, beliefs and affect about memory. It may also distinguish between state and trait memory distrust.

As older adults generally perceive more changes in their memory, report lower levels of memory capacity, and perceive that they have less control over their memory than younger adults (Dixon et al., 1988), it was predicted that compliance and response change scores would be more pronounced in older adults. Contrary to expectations, there was no indication that compliance or overall response change scores varied with age. Additionally, no significant differences between young, middle-aged or older adults were observed on measures of subjective memory beliefs (see *Table 19*.). One explanation for the result may be that the older adult participant sample (participants aged over 60) may not be representative of the general population. In the present study the older adult participant sample were recruited mainly from a University Lifelong Learning Centre. As such they may not experience the same rate of memory decline and have more confidence in their memory abilities.

In a forensic setting, the current study results might have far-reaching consequences. Given the weight that triers of fact attach to confidence (Lieppe et al., 1992), people with pessimistic ideas may erroneously be treated as less credible witnesses (Van Bergen et al., 2009). People who judge their memory to be very poor because they suffer from the *memory distrust syndrome* are thought to be especially prone to memory distortions (Gudjonsson & MacKeith, 1982). Memory distrust is “a condition where people develop profound distrust of their memory recollections, as a result of which they are particularly susceptible to relying on external cues and suggestions" A number of court cases in which defendants suffering from memory distrust developed false memories, eventually resulted in false confessions (Gudjonsson, 2003, p. 196; Gudjonsson et al., 1999).

The second experimental hypothesis that negative feedback would be associated with greater response change scores compared with neutral feedback was supported and in line with the findings of Tata and Gudjonsson (1990), Baxter (under review); Baxter et al., (2006), and McGroarty and Baxter (2007, 2009). 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 their 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 ratings of interview difficulty (see *Table 13*). Participants receiving negative feedback reported higher overall difficulty with the interview than those receiving neutral feedback, also supporting the findings of McGroarty and Baxter, (2007, 2009).

Consistent with prior research (e.g., Baxter et al., 2006; McGroarty & Baxter, 2007, 2009), the questions asked in the present study were not overtly misleading. The finding that negative feedback continued to be associated with response change scores in the absence of leading questions supports previous findings (Baxter et al., 2006; McGroarty & Baxter, 2007, 2009) as well as Gudjonsson's argument that the interrogative pressure and suggestive questioning aspects of interrogative suggestibility are relatively independent (Gudjonsson, 1984, 1991, 2003).

Leading questions are one type of suggestive stimulus. As previously defined, these are questions phrased in such a way as to communicate expectations and premises to interviewees and hence to suggest a particular answer. The use of leading questions in police investigations is widely recognised as problematic (Fisher, 2010), as their use has been shown to produce distorted responses. The acceptance or rejection of misleading information contained in leading questions is mediated by cognitive mechanisms such as discrepancy detection (Schooler & Loftus, 1986; Tousignant, Hall, & Loftus, 1986). In contrast, reactions to interrogative pressure may depend primarily on social factors (Bain & Baxter, 2000; Bain et al., 2004; Baxter & Boon, 2000; Baxter et al., 2003; Gudjonsson, 2003; Gudjonsson & Lister, 1984). Forensic interviews may reflect an imbalance of social status and power (e.g., a police officer and a child or vulnerable witness). In order to deal with uncertainty and the potential threat of inadequate performance, interviewees may employ inefficient coping mechanisms (e.g., a greater reliance upon the interviewer for guidance as to whether a question has been answered correctly (Gudjonsson, 1988b; Emmett et al., 2003), particularly through facial cues and interviewer demeanour (Baxter & Boon, 2000; Baxter et al., 2003).

The percentage of response change following negative feedback found by Baxter et al., (2006) and McGroarty and Baxter (2007) was 10.75 and 10.81 respectively. As expected the rate of response change was less than the GSS norm since leading questions were absent. In the present study, response change of 3.91 i.e. 16.15% following negative feedback was recorded. This result might be an artefact of an increased participant sample. However, variations in age appear to be related interactively to response change. This aspect will be discussed later in the discussion.

The present study also observed response change in the neutral feedback condition, which was not unexpected. The repetition of questions may, in itself have acted as an implicit form of negative feedback communicating to participants that their previous responses were in some way incorrect or inappropriate and should be changed (Gudjonsson, 2003; Gudjonsson & Clark, 1986; Register & Kihlstrom, 1988; McGroarty & Baxter, 2007). However, memory for some details may have been too weak to be consistent on requestioning.

Participant's tendency to change their answers toward or away from accuracy in response to negative feedback was also measured, replicating the work of McGroarty and Baxter (2007, 2009). Where response change occurred, responses were changed both towards and away from accuracy. Significantly more responses changed away from accuracy resulted from negative feedback than from neutral feedback. 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 correct 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 interrogative pressure without any added recall effort or a belief that their new answers were *better*.

That age had no significant effect on the response change scores following negative feedback would seem to support the findings of Polczyk et al., (2004). When younger adults (mean 23.3 years) were compared with older adults (mean 64.1 years) on interrogative suggestibility as measured by the GSS2 (Gudjonsson, 1987b, 1997; Polish version: Polczyk, 2000), older adults scored higher on measures of Yield 1 and 2 but not on Shift, following negative feedback.

Closer inspection of the present study data (see *Tables 22* and *23*) revealed that older adults aged 65 – 74 and 75+ recorded response changes following negative feedback of on average 2.14 (i.e. 7.7%) and 2.33 (i.e. 8.4%) respectively. In contrast younger adults aged 18 – 24 recorded on average 6 response changes (i.e. 26%). In the present study it would seem that adults aged 65 and over recorded response changes in the absence of leading questions consistent with the findings of Baxter et al., (2006) and McGroarty and Baxter (2007). In contrast, younger adults aged 18 – 24 would appear to be less resistant to interrogative pressure in the form of negative feedback. Additionally, younger adults aged 18 – 24 reported higher ratings of perceived interview difficulty following negative feedback. However, as insufficient power precluded systematic between-groups analysis to determine whether significant differences existed between age groups, the above findings represent a trend only. Future work should ensure a sufficiently large sample to investigate age-related differences in response change follwing negative feedback.

One possible explanation for the observed pattern of results is that younger adults are less self-confident and more dependent on authority such that they attended less to recall and more on managing their interpersonal situation and meeting the demands of the interviewer (Bain & Baxter, 2000; Baxter & Boon, 2000; Gudjonsson & Clark, 1986; Gudjonsson & Lister, 1984). To some extent this interpretation is corroborated by the results of the regression analysis. Neither subjective memory beliefs nor positive affect were statistically significant predictors of proneness to change responses following negative feedback in adults at any age. Such an interpretation is also consistent with the fact that adolescents score higher than adults on Shift but not on Yield (Gudjonsson, 2003).

The higher younger adults' response change scores may also be a reflection of interviewer behaviour, which is an important influence on interrogative suggestibility. Participants interviewed with the GSS (Gudjonsson, 1983, 1984) in an *abrupt* manner scored higher on measures of Shift and Total Suggestibility than those interviewed in a *friendly* manner (Baxter et al., 2003; Bain et al., 2004). While abrupt or stern interviewing styles raise suggestibility scores above the GSS population norms, relaxed and friendly styles lower scores below the GSS norms. Shift rather than Yield is the GSS measure most affected by these variations. In the present study, younger adults may have perceived the interviewer to be abrupt which resulted in higher response change scores. However, these two explanations might not be independent.

It should be noted that in the present study age was a continuous variable. Therefore the above discussion results from a trend in the data only. Future research should ensure a sufficiently large sample to allow between-groups analysis of the respective contributions of participant age, self-confidence and dependence on authority with interviewer behaviour.

Additionally, acceptance or rejection of negative feedback is also dependent on the perceived credibility of the person giving the feedback. Skagerberg and Wright (2009) found that susceptibility to post-identification feedback was only observed when co-witnesses were attributed to a high credibility source. The older adults in the present study may have perceived the demands of the interview as well as the interviewer to be less than credible and so rejected the negative feedback. Future research should ensure a sufficiently large sample to allow between-groups analysis of the perceived credibility of source in relation to response change following negative feedback.

Gudjonsson argued that psychological vulnerabilities should not be interpreted in isolation from other factors. Where they exist, issues relating to the physical and mental health of witnesses have the potential to influence their behaviour within a forensic interview. A third aim was to assess the relationship between the individual difference measures of health status, sensory deficits, affective states and overall response change irrespective of the type of feedback. The results show that measures of post-test positive affect were significantly negatively correlated with overall response change scores, showing that the more positive the participant the lower the response change scores (see *Table 16*). However, again it must be noted that although significant, the relationship between post-test positive affect and response change scores was weak accounting for less than 1 per cent of the variance. Regression analysis revealed that participant scores on the MIA subscale of task only was identified as a statistically significant predictor of an overall tendency to change answers irrespective of feedback (see *Table 18*). Further regression analysis revealed that scores on the MIA subscale of task and measures of post-test positive affect were significant predictors of a tendency to change answers following neutral feedback. No significant predictors were detected in the negative feedback condition (see *Tables 25* and *26*).

There are several limitations in the present study that deserve some comment. Firstly, as discussed above, the older adult participant sample may not have been representative of the general population. Second, although the significant correlation between the MIA subscale of task and overall response change scores was significant the relationship was weak accounting for less than 1 per cent of any variance. One possible explanation is the artificial procedures employed in the present study which limit the generalisability of the results to real-world forensic interviews. Levels of interrogative pressure during simulated forensic interviews are modest. In the forensic cases involving memory distrust (Gudjonsson, 2003; Gudjonsson et al., 1999) all suspects had been exposed to potentially high levels of interrogative stress. Therefore, the results of the present study do not preclude the possibility that when people suffering from memory distrust are exposed to high interrogative pressure, they become suggestible. Furthermore, as discussed above, the reactions to interrogative pressure may depend primarily on social factors. Therefore a test used to measure state memory distrust as opposed to trait memory distrust may generate higher changes in response. Further work might test this possibility.

All interviews were conducted by the same interviewer who was familiar with both the purpose and expected outcomes of the research, and this may have influenced interviewer behaviour (Hayes & Delamothe, 1997). Previous studies have minimised such problems by asking several trained assistants who are unaware of the research hypotheses to conduct interviews (e.g. Milne & Bull, 2002; Searcy, Bartlett, Memon & Swanson, 2001). However, a consistent interview manner is also required. As introduced in Chapter 2, interviewer behaviour is an important influence on interrogative suggestibility. Participants interviewed with the GSS (Gudjonsson, 1983, 1984) in an *abrupt* manner scored higher on measures of Shift and Total Suggestibility than those interviewed in a *friendly* manner (Baxter et al., 2003; Bain et al., 2004). While abrupt or stern interviewing styles raise suggestibility scores above the GSS population norms, relaxed and friendly styles lower scores below the GSS norms. Shift rather than Yield is the GSS measure most affected by these variations. In the present study the same interviewer tried to maintain a consistently formal demeanor during the interview neither abrupt nor friendly. Had a pilot study been carried out a rating could have been obtained of the interviewer's manner to ensure standardization of interviewer behaviour across participants.

Chapter 5: General Discussion

One of the most challenging tasks facing the criminal justice system is the evaluation and determination of the credibility and reliability of information received from victims, witnesses and suspects (Brewer & Weber, 2008; Gudjonsson, 2010). The relationship between witness confidence and accuracy (Brewer & Weber, 2008; Brewer et al., 2007; Odinot et al., 2009; Penrod & Cutler, 1995; Sporer et al., 1995; Wells & Murray, 1984), and witness consistency (Fisher et al., 2009) across younger and older adults are two features of witness credibility considered within this thesis. The two studies presented above used an experimental procedure, a mock witness paradigm, and an individual differences analysis to investigate witness response within a simulated forensic interview.

5.1 Main Results

Experiment 1 provides an extension of Koriat and Goldsmith's (1996) framework from answering general knowledge questions to answering episodic memory queries for details of an event. The framework proposes that memory accuracy is mediated by metacognitive monitoring and control processes, whereby people strategically regulate the accuracy of their memory reports by using confidence judgments to guide what they report. The way in which age (younger *vs.* older adults), test format (cued recall *vs.* four alternative forced-choice (4-AFC) recognition questions) and type of content (action *vs.*

descriptive details) interact and relate to both the quantity of correct information retrieved and the monitoring of this information, was investigated. Several differences were found between older and younger adults both in memory accuracy and in its metacognitive determinants.

Firstly, the option of free report allowed older adults to regulate their memory accuracy to the extent that their performance was equivalent to that of younger adults. However, such gains in accuracy came at a higher cost in quantity retrieved compared with younger adults This result did not support the experimental hypothesis and is contrary to existing research (Kelley & Sahakyan, 2003; Pansky et al., 2009; Rhodes & Kelley, 2005).

That older adults achieved comparable memory accuracy at free report as younger adults, but did so at the expense of quantity correct has potential implications for interviewing protocols. The Cognitive Interview (CI) and Enhanced Cognitive Interview (ECI) are among the most rigorously tested and widely accepted methods for improving the accuracy and completeness of eyewitness reports. However, research that has assessed a CI's effectiveness to facilitate the event memory of older adults has resulted in inconsistent findings. This may be explained by the differing methodologies employed including different control interviews, different versions of the CI, small sample sizes and an older adult sample who may not be representative of the older population as a whole (Mello & Fisher, 1996). Additionally, because interviewers were aware of the experimental hypotheses, their expectations may have resulted in systematic differences in the length of interviews, or in the number and quality of questions asked in the interviews (Hayes & Delamothe, 1997).

In the current study, the setting of a more conservative response criterion by older adults resulted in the volunteering of fewer correct (and incorrect) responses. This suggests that the older adult sample believed that they had a strong memory for the to-be-remembered event and expected to be able to remember such details. It is for the investigators to decide beforehand the purpose of the interview (i.e. whether accuracy or quantity of information is more important) and to advise older adult participants about their control policies accordingly.

Second, at free report memory accuracy was greater in respect of descriptive details. A significant interaction between age group and content resulted in older adults' greater gain in memory accuracy between forced and free report, when reporting description details. However, again this came at a higher cost in quantity retrieved compared with younger adults. This finding supports the experimental hypothesis as well as the findings of Aizpurua, et al., (2009), Garcia-Bajos and Migueles (2003), and Migueles and Garcia-Bajos (2004), (– but see Ibabe & Sporer, 2004).

Actions of an event best capture the argument and constitute the gist of an event (Aizpurua et al., 2009). One possible explanation for the results reported here is that when reporting action details, an individual may recall information congruent with the prior knowledge of the event held in the form of situational schemata or scripts (Garcia-Bajos & Migueles, 2003; Greenberg et al., 1998; List, 1986). In contrast, he or she is less likely to possess schemata in respect of descriptive details (Heuer & Reisberg, 1990). Memory bias resulting from these schemata is particularly significant in the recognition for the actions of an event. Access to gist information is relatively automatic, and is preserved with aging (Cohen & Faulkner, 1989; Koutstaal & Schacter, 1997). It is known that older adults are more prone than younger adults to depend on prior knowledge (Mather et al., 1999). If older adults are overly dependent on gist-based processing they are more likely to display memory bias (Craik & Simon, 1980; Hasher & Zacks, 1979; Hashtroudi et al., 1990; Puglisi et al., 1988; Rabinowitz & Ackerman, 1982). Poorer retention will increase the generation of incorrect responses, particularly when reporting action details.

Alternatively, an increase in errors when identifying actor-action links may be due to a reliance on familiarity of the components, that is, the people and the actions rather than recollection of the association between the two. This interpretation is consistent with the associate deficit hypothesis (Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2003; Old & Naveh-Benjamin, 2008), detailed in Chapter 1, which attributes age-related memory deficits to the inability to encode and bind together items or features of an event. Future research should systematically investigate whether older adults' poorer performance in recalling action details of an event reflects a greater use of gist-based processing or an inability to encode and bind together items or features of an event.

The distinction between action and descriptive details is an important one for forensic investigations. During an initial investigation, witnesses detailed descriptions of an event i.e. what happened who the perpetrators were and leads to further assist police with their investigation, may be obtained. Investigators and triers of fact (e.g. jurors) are more likely to be concerned with central action details considered essential for effective prosecution and resolution of the case (Ibabe & Sporer, 2004). Forensic investigators should be aware that witnesses, particularly older adults may remember having seen a particular person before, as well as the actions involved in a crime, but may be unable to bind the culprit to the committed offense. Older adults tend to commit more errors in identifying criminals, and these errors are most often related to false identifications. This is especially the case following relatively long delays (Havard & Memon, 2009; Memon & Bartlett, 2002; Memon & Gabbert, 2003a, 2003b; Rose et al., 2005; Searcy et al., 2000; Wilcock, Bull, & Vrij, 2005, 2007; Yarmey, 1993).

Third, participants displayed greater over-confidence in response to descriptive compared with action details. However, a significant interaction between age group and content was observed. Younger adults' over-confidence was greater when reporting descriptive compared with action items. In contrast, older adults' over-confidence was significantly greater when reporting action compared with descriptive items. This result supports the experimental hypothesis and the findings of (e.g., Aizpurua et al., 2009; Garcia-Bajos & Migueles, 2003; Greenberg et al., 1998; List, 1986; Migueles & Garcia-Bajos, 2004). In this respect, the results follow those for accuracy.

The analysis reported here draws upon within-subjects comparisons of confidence, which is more sensitive to participants' ability to monitor the correctness of their responses than the between-subjects Accuracy-Confidence (AC) relationship, as described in Chapter 1, reported for identification decisions (Sporer, 2008; Sporer et al., 1995). The traditional measurement of the AC relation is to calculate a point-biserial correlation coefficient resting on two data points per participant. In the current study 24 items entered into the calculation. The results would seem to confirm Sporer's (2008) argument that the AC relationship for event memory details may be higher. The present findings (if replicated with different stimulus material and questions) would also have important practical implications. Researchers and practitioners should clearly separate confidence in respect of individual questions about the event from confidence relating to an identification decision. Additionally, while the AC relation for identification decisions may be weak (Bothwell et al., 1987; Penrod et al., 1982; Sporer et al., 1995; Wells & Murray, 1984) and malleable (Granhag, 1997; Granhag et al., 2000; Hastie et al., 1978; Odinot et al., 2009; Shaw, 1996; Shaw & McClure, 1996; Turtle & Yuille, 1994), witnesses, particularly older adults, may be better able to calibrate their confidence when reporting descriptive compared with action details contained within an event.

Fourth, older adults chose to volunteer proportionately fewer responses which indicate that the older adults employed a more conservative control policy compared with younger adults. This result, together with the finding of a significantly greater reduction in memory quantity performance stemming from the exercise of a free-report option by the older adults, could be taken to reflect a general difference between younger and older adults in their control policy in trading accuracy for quantity. It would appear that the older group placed greater emphasis on accuracy and the younger group greater emphasis on quantity of event information. This result is contrary to previous results (Aizpurua et al., 2009; da Silva & Sunderland, 2010; Kelley & Sahakyan, 2003;

Koutstaal & Schacter, 1997; Pansky et al., 2009), whereby older adults were observed to adopt a more lenient response. Additionally previous research has demonstrated that older adults are less likely compared with younger adults to withhold answers when given the option of free report, even though this leads them to higher error rates (see Jacoby, et al; Kelley & Sahakyan, 2003). The result reported here may reflect individual differences in the older adult population sample used in that they have more confidence in their own memory and set a more demanding criterion for reporting (Stretch & Wixted, 1998; Morrell, Gaitan, & Wixted, 2002). The proportionately higher volunteering rate by younger adults related to an overall lower confidence in responses. This might imply that younger adults used a different basis for memory responses i.e. plausibility or familiarity rather than recollection (Kelley & Sahakyan, 2003).

Path analysis indicate that memory accuracy in free report when reporting action details was largely affected directly by age and indirectly via retrieval at forced report, and through the indirect effects of participant mental health on control policy. Memory accuracy in free report when reporting descriptive details was largely affected directly by age and through the indirect effects of control policy and not by retrieval at forced report.

Finally, memory accuracy was significantly greater in response to cued-recall questions. This result supports the stated hypothesis and established eyewitness research (see Fisher & Patterson, 2004; Hilgard & Loftus, 1979; Ibabe & Sporer, 2004; Koriat & Goldsmith, 1996- Experimwnt 1; Lipton, 1977; Neisser, 1988). An interaction was observed between age group and test format. Older adult accuracy at free report was significantly greater in response to cued-recall questions. However younger adult accuracy was significantly greater in response to recognition questions. This interactive pattern was not in the expected direction, and is contrary to existing research which shows greater age-related decrements in response to recall compared with recognition questions (Craik & McDowd, 1987; Nyberg et al., 2003; Rabinowitz 1984, 1986), and greater in response to recall compared with cued-recall questions (Ceci & Tabor, 1981; Craik et al., 2001). In the present study, it would appear that older adults were able to employ sufficient internal processing resources to consciously retrieve the correct answer, by using the information contained within the open-ended questions as a recall cue. In contrast, greater younger adult accuracy in response to recognition questions may reflect responses chosen on the basis of familiarity or plausibility. Alternatively, this result may simply reflect a spurious result by younger adults.

Existing eyewitness research holds that testing procedures involving recognition or directed questioning can have contaminating effects on memory (e.g., see Brown et al., 1977; Gorenstein & Ellsworth, 1980; Hilgard & Loftus, 1979; Loftus, 1979, 1982; Loftus & Hoffman, 1989). Thus the general recommendation is to elicit information in an initial free narrative format before moving on to directed questioning. (see Fisher et al., 1987; Hilgard & Loftus, 1979). The results reported here would seem to provide additional support for this established position.

Experiment 2 had three main aims in seeking to develop the Gudjonsson and Clark (1986) theoretical model of Interrogative Suggestibility. Firstly, the study aimed to investigate the relationship between subjective memory beliefs, compliance and response change across young, middle-aged and older adults, irrespective of the type of feedback delivered. A second aim was to examine the effects of investigative pressure associated with neutral and negative feedback. The effects of age on response change scores following feedback were also investigated. A third aim was to assess the extent to which individual differences in measures of education, crystallised intelligence, health status, sensory deficits, and affective states moderated response change scores.

The findings of Experiment 2 can be summarised as follows:

Firstly, participant subjective memory beliefs were significantly associated with compliance, supporting the stated hypothesis and the findings of Van Bergen et al., (2009). Compliance is a tendency of the individual to go along with propositions, requests or instructions for some immediate instrumental gain (Gudjonsson, 2003). Compliance may be viewed as an ineffective coping mechanism during tasks perceived as stressful (Costa & McCrae, 1992). Gudjonsson, (2003) argued that people with subjective memory problems tend to rely on external cues and suggestions, making them vulnerable to false memories. The results of this study support this line of reasoning. Participants who rated themselves as memory anxious were significantly more likely to report higher compliance scores. One explanation is that people suffering from memory distrust (i.e. who evaluate their memory as poor) are more susceptible to the authority of others. Within the context of a forensic interview, witnesses may perceive the opinion of the interrogator as more important than their own. However, because this relationship was based only on correlations, the causal relationship between these two concepts cannot be established. It could also be argued that people start to distrust their memory because they are more easily intimidated by authorities.

Furthermore, following Gudjonsson's reasoning, it was predicted that participants with negative memory beliefs would display greater response change irrespective of the type of feedback delivered. The results would also appear to support the stated hypothesis. A significant negative relationship was obtained between the Metamemory in Adulthood (MIA) subscale of task and response change scores, showing the higher the participant's level of knowledge of basic memory processes and functions, the lower the response change scores. This pattern of results is contrary to the findings of Van Bergen et al., (2009) who found no significant correlations between the Squire Subjective Memory Questionnaire (SSMQ; Squire et al., 1979; Van Bergen et al., 2009) and suggestibility (Yield or Shift) as measured by the Dutch translation of the Gudjonsson Suggestibility Scale (Gudjonsson, 1983, 1984; Merckelbach et al., 1998). The use of differing experimental stimuli may account for the different pattern of results. The MIA may be more sensitive to individual variations in subjective knowledge, beliefs and affect about memory. It may also distinguish between state and trait memory distrust. Contrary to expectations, there was no indication that compliance or overall response change scores varied with age. Additionally, no significant differences between young, middle-aged or older adults were observed on measures of subjective memory beliefs. One explanation for the results may be that the older adult participant sample (participants aged over 60) may not be representative of the general population. In the present study the older adult participant sample were recruited mainly from a University

Lifelong Learning Centre. As such they may not hold such extreme opinions about their memory decline.

In a forensic setting, the current study results (if replicated with differing samples and questions) might have far-reaching consequences. Given the weight that triers of fact attach to confidence (Lieppe et al., 1992), people with pessimistic ideas may erroneously be treated as less credible witnesses (Van Bergen et al., 2009). People who judge their memory to be very poor because they suffer from the *memory distrust syndrome* are thought to be especially prone to memory distortions (Gudjonsson & MacKeith, 1982). Memory distrust is "a condition where people develop profound distrust of their memory recollections, as a result of which they are particularly susceptible to relying on external cues and suggestions" A number of court cases in which defendants suffering from memory distrust developed false memories, eventually resulted in false confessions (Gudjonsson, 2003, p. 196; Gudjonsson et al., 1999).

Second, in the absence of leading questions, negative feedback was associated with more response change compared with neutral feedback. This finding was as expected and confirms existing research (Tata & Gudjonsson ,1990; Baxter, under review; Baxter et al., 2006, and McGroarty & Baxter, 2007, 2009). This result did not vary significantly with age.

That age had no significant effect on the response change scores following negative feedback would seem to support the findings of Polczyk et al., (2004). When younger adults (mean 23.3 years) were compared with older adults (mean 64.1 years) on interrogative suggestibility as measured by the GSS2 (Gudjonsson, 1987a, 1997; Polish version: Polczyk, 2000), older adults scored higher on measures of Yield 1 and 2 but not on Shift, following negative feedback.

Closer inspection of the present study data (see *Tables 22* and *23*) revealed that older adults aged 65 – 74 and 75+ recorded response changes following negative feedback of on average 2.14 (i.e. 7.7%) and 2.33 (i.e. 8.4%) respectively. In contrast younger adults aged 18 – 24 recorded on average 6 response changes (i.e. 26%). In the present study it would seem that adults aged 65 and over recorded response changes in the absence of leading questions consistent with the findings of Baxter et al., (2006) and McGroarty and Baxter (2007). In contrast, younger adults aged 18 – 24 would appear to be less resistant to interrogative pressure in the form of negative feedback. Additionally, younger adults aged 18 – 24 reported higher ratings of perceived interview difficulty following negative feedback. However, as insufficient power precluded systematic between-groups analysis to determine whether significant differences existed between age groups, the above findings represent a trend only. Future work should ensure a sufficiently large sample to investigate age-related differences in response change follwing negative feedback.

One possible explanation is that younger adults are not so self-confident and more dependent on authority such that they attended less to recall and more to managing their interpersonal situation and meeting the demands of the interviewer (Bain & Baxter, 2000; Baxter & Boon, 2000; Gudjonsson & Clark, 1986; Gudjonsson & Lister, 1984). To some extent this interpretation is corroborated by the results of the regression analysis. Although significant correlations were observed between participant subjective memory beliefs and measures of post-test positive affect, neither subjective memory beliefs nor positive affect were statistically significant predictors of response change scores following negative feedback in adults at any age. Such an interpretation is also consistent with the fact that adolescents score higher than adults on Shift but not on Yield (Gudjonsson, 2003).

The results would suggest that negative feedback increased participants', particularly young adults' 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 et al., 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. 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). In the present study, negative feedback may also have affected participants' expectations, such that they attended less to recall and more to managing their 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 ratings of interview difficulty (see *Tables 13 and 24*). Participants, particularly younger adults, receiving negative feedback reported higher perceived difficulty with the interview than those receiving neutral feedback, also supporting the findings of McGroarty and Baxter, (2007, 2009).

However, the higher younger adult response change scores may be a reflection of interviewer behaviour, which is an important influence on interrogative suggestibility. Participants interviewed with the GSS (Gudjonsson, 1983, 1984) in an *abrupt* manner scored higher on measures of Shift and Total Suggestibility than those interviewed in a *friendly* manner (Baxter et al., 2003; Bain et al., 2004). While abrupt or stern interviewing styles raise suggestibility scores above the GSS population norms, relaxed and friendly styles lower scores below the GSS norms. Shift rather than Yield is the GSS measure most affected by these variations. In the present study, younger adults may have perceived the interviewer to be abrupt which resulted in higher response change scores. However, these two explanations might not be independent. It should be noted that in the present study age was a continuous variable. Therefore the above discussion results from a trend in the data only. Future research should ensure a sufficiently large sample to allow between-groups analysis of the respective contributions of participant age, self-confidence and dependence on authority with interviewer behaviour.

Additionally, acceptance or rejection of negative feedback is also dependent on the perceived credibility of the person giving the feedback. Skagerberg and Wright (2009) found that susceptibility to post-identification feedback was only observed when co-witnesses were attributed to a high credibility source. The older adults in the present study may have perceived the demands of the interview as well as the interviewer to be less than credible and so rejected the negative feedback. Future research should ensure a sufficiently large sample to allow between-groups analysis of the respective contribution of perceived credibility of source following negative feedback.

The findings reported here may have some implications for forensic interviewing practise. Judges, litigators, and legal scolars deem witness consisteny to be one of the most important features of witness credibility (Brewer, et al., 1999; Fisher et al., 2009; Potter & Brewer, 1999). It is often inferred that inconsistent testimony is as a result of poor memory or deception. However, the results reported here show that in the absence of overtly leading questions, some interviewees may alter their answers in response to perceived disapproval communicated by the interviewer, whether this is verbal, non-verbal or both (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.

A final aim was to assess the relationship between the individual difference measures of health status, sensory deficits, affective states and response change. Regression analysis revealed that participant scores on the MIA subscale of task only were identified as a statistically significant predictor of an overall tendency to change answers irrespective of feedback (see *Table 18*). Further regression analysis revealed that scores on the MIA subscale of task and measures of post-test positive affect were significant predictors of a tendency to change answers following neutral feedback. No significant predictors were detected in the negative feedback condition (see *Tables 25* and *26*).

5.2 Methodological issues

There are several limitations in the present studies that may limit the generalisability of the findings to real-world forensic settings. In field settings, the amount of emotional involvement may be stronger and and the delays between an incident and the interviewing of witnesses longer than can be replicated within a laboratory setting. Due to the reconstructive nature of memory, memories may be prone to decay and information presented after an event can change or add to aspects of a memory report (Hyman et al., 1995; Loftus & Palmer, 1974; Loftus et al., 1978; Loftus & Pickrell, 1995; Wright & Stroud, 1998). In the current research participants were asked to recall a staged video-taped event almost immediately after witnessing it. It remains a possibility that their memories for the events were artificially strong. Given these differences between laboratory conditions and what may occur in the field, it can reasonably be argued that the results reported here offer only conservative estimates of witness response. Thus the theoretical principles investigated here should also be tested under more ecologically valid conditions.

In both experiments questionnaires and tasks were given in a fixed order, and this may have introduced order effects. Koriat and Goldsmith (1996 – Experiment 2) counterbalanced the order of the two phases across participants. The results revealed that, in general, phase order had little or no effect, and the same pattern of results obtained in their Experiment 1 were replicated. However, in the current research order effects cannot be discounted.

Comparing accuracy as a condition of question form depends highly on the way individual questions are constructed, as well as on the difficulty level of specific questions. Although attempting to formulate questions that were parallel across question forms, and foils in the Alternative–Forced-Choice test (Experiment 1) which were as plausible as possible, the external validity of the questions used in the current research cannot be guaranteed. Significant results did emerge on a number of the dependent variables in the expected direction however, which lends some support to the validity of the tests used and the populations evaluated.

Levels of interrogative pressure during simulated forensic interviews are modest. In the forensic cases involving memory distrust, as detailed above, (Gudjonsson, 2003; Gudjonsson et al., 1999) all suspects had been exposed to potentially high levels of interrogative stress. Therefore, given these differences between experimental conditions and what may occur in the field, it can reasonably be argued that the present studies offer only conservative estimates of witness response within a forensic interview.

All interviews were conducted by the same interviewer who was familiar with both the purpose and expected outcomes of the research. This may have influenced the perceptions of the participants which in turn may have affected their behaviour. It is usual for police interviewers not to have witnessed an incident and to the extent that there is little or no evidence or other reliable witness accounts, the interviewer may be totally reliant on the account of the interviewee to further the investigation. It may be evident to interviewees that the interviewer has much less or perhaps more knowledge of an incident than they do. To minimise such an effect, in Experiment 1, the researcher announced that she had not seen the film, explaining that this was an experimental control designed to prevent her from *leaking* information about the video during the subsequent testing (Wright & Holliday, 2007). The researcher stood outside the testing room with the door closed while the video was showing. However, in Experiment 2 the videotaped event was presented to the participants by the interviewer herself. A suggested improvement for future studies is to ensure that the interviewer is blind to the contents of the video presentation and that this is clearly communicated to participants.

All interviews were conducted by the same interviewer who was familiar with both the purpose and expected outcomes of the research, and this may have influenced interviewer behaviour (Hayes & Delamothe, 1997). Previous studies have minimised such problems by asking several trained assistants who are unaware of the research hypotheses to conduct interviews (e.g. Milne & Bull, 2002; Searcy et al., 2001). However, a consistent interview manner is also required. As introduced in Chapter 2, interviewer behaviour is an important influence on interrogative suggestibility. Participants interviewed with the GSS (Gudjonsson, 1983, 1984) in an *abrupt* manner scored higher on measures of Shift and Total Suggestibility than those interviewed in a *friendly* manner (Baxter et al., 2003; Bain et al., 2004). While abrupt or stern interviewing styles raise suggestibility scores above the GSS population norms, relaxed and friendly styles lower scores below the GSS norms. Shift rather than Yield is the GSS measure most affected by these variations. In the present study the same interviewer tried to maintain a consistently formal demeanor during the interview neither abrupt nor friendly. A suggested improvement for future work is to conduct a manipulation check of the interviewer's manner to ensure standardization of interviewer behaviour across participants.

5.3 Future Work

Experiment 1 provides an extension of Koriat and Goldsmith's (1996) framework from answering general knowledge questions to answering episodic memory queries for details of an event. The framework proposes that memory accuracy is mediated by metacognitive monitoring and control processes, whereby people strategically regulate the accuracy of their memory reports by using confidence judgments to guide what they report. The results reported here suggest that the framework might usefully be applied to real-world forensic interview settings. Future work might usefully extend and assess the utility of the framework by testing upon recall of a real-life event.

The findings of Experiment 1 would also appear to extend memory deficits in response to action details, particularly by older adults, to a neutral scene. However this interpretation should be treated with caution. The present study replicated the research of Ibabe and Sporer (2004), and defined action details as character non verbal behaviours. Descriptive details were defined as physical characteristics of scenes, persons and objects. In contrast Aizpurua et al., (2009) defined actions as verbal or non-verbal behaviours (e.g., pointed a gun at someone's face). The people related content included general physical characteristics of the people involved, as well as descriptions of their clothing and accessories. Details referred to primary object characteristics and circumstantial information surrounding the event. Ibabe and Sporer, (2004) found that young adult accuracy was significantly greater when reporting action compared with descriptive details. In contrast, Aizpurua et al., (2009), found that participants accepted more false actions, thus achieving greater recognition accuracy when reporting people and details. This pattern of results was more pronounced in older adults. In both studies the filmed event consisted of a violent robbery. The differing conceptualisation may have contributed to the differing results.

The differing conceptualisation regarding memory for each component (i.e. individual actions and people) is important to future research within the eyewitness domain. An agreed definition of *action* and *description* details of an event would help eyewitness researchers define and focus the applications of their research. Additionally, memory performance should be investigated within different staged events (neutral scenes and emotional crime events) to assess whether the results reported here are replicated across differing experimental stimuli.

The use of certain stimulus materials (e.g., picture slides or the spoken narrative of the GSS 1 and 2) may limit the generalisability of findings to the dynamics of real police interviews concerned with events. 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 (Beail, 2002; Cardone & Dent, 1996; Scullin & Ceci, 2001). Age-related deficits are more pronounced with tests of episodic memory (Park & Minear, 2004). The results of

Experiment 2 relating to interrogative pressure in the form of negative feedback support the findings of McGroarty and Baxter (2007, 2009). It could be argued that such results are not restricted to spoken narrative stimulus material. Future work might usefully extend and assess this viewpoint by testing the effects of interrogative pressure upon recall of a real-life event. If it can be demonstrated that recall of a real-life event is similarly prone to the influence of interrogative pressure, then it can be more confidently assumed that effects observed in GSS studies are not associated solely to the scales stimulus material and that they have relevance beyond the laboratory.

5.4 Conclusion

Memory processing and witness testimony are dynamic and complex. Participant metacognitive judgments as well as interviewer demeanour potentially have consequences for memory accuracy. Professional forensic interviewers should be aware of the potential influence of these forces in order to facilitate accurate witness recall, while simultaneously minimising the possibility of interfering with accurate witness recall via explicit or implicit negative feedback.

References

Adams-Price, C. (1992), Eyewitness memory and aging: predictors of accuracy in recall and person recognition. *Psychology and Aging*, *7*, 602-608.

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

Aizpurua, A., Garcia-Bajos, E., & Migueles, M. (2009). False Memories for a Robbery in Young and Older Adults. *Applied Cognitive Psychology, 23*, 174-187.

Akehurst, L., Milne, R., & Köhnken, G. (2003). The effects of children's age and delay on recall in a cognitive and structured interview. *Psychology, Crime & Law, 9*, 97-107. In R. Bull, T. Valentine, & T. Williamson (Eds.). *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions* (pp. 137-160), Wiley-Blackwell: Chichester.

Albert, M.S., Jones, K., Savage, C.R., Berkman, L., Seeman, T., Blazer, D., & Rowe, J.W. (1995). Predictors of cognitive change in older persons: McArthur Studies of Successful Aging. *Psychology and Aging, 10,* 578-589.

Alley, D., Suthers, K. & Crimmins, E. (2007). Education and Cognitive Decline in Older Americans, Results from the AHEAD Sample. *Research on Aging, 29(1)*, 73-94.

Allwood, C.M., Ask, K., & Granhag, P.A. (2005). The Cognitive Interview: Effects on the realism in witnesses' confidence in their free recall. *Psychology, Crime & Law*, *11(2)*, 183–198.

Anderson, N.D. (1999). The attentional demands of encoding and retrieval in younger and older adults: II. Evidence from secondary task reaction time distributions. *Psychology and Aging*, *14 (4),* 645-655.

Axmacher, N., Gossen, A., Elger, C.E., & Fell, J. (2010). Graded Effects of Social Conformity on Recognition Memory. *Plos One, 5(2): Art. No. e9270*.

Bäckman, L. (1991), Recognition memory across the adult life span: The role of prior knowledge. *Memory & Cognition*, *19*, 63-71. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Bäckman, L., Small, B.J., & Larsson, M. (2000). Memory and ageing. In J.G.Evans, T.F. Williams, B.L. Beattie, J-P Michel, & G.K. Wilcock (Eds.), *Oxford textbook of geriatric medicine* (2nd ed., pp. 906-916). New York: Oxford University 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.

Balcerak, L.J., Rebok, G.W. (1986). *Mnemonic training, self-efficacy, and memory performance in young and old adults.* Paper presented at the 94th annual convention of the American Psychological Association, Washington DC.

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

Balota, D.A., Dolan, P.O., & Duchek, J.M. (2000). Memory changes in healthy older adults. In E. Tulving & F.I.M. Craik (2nd Ed.), *The Oxford handbook of memory* (pp. 395-410). Oxford, England: Oxford University Press.

Baltes, P.B. (1993). "The Aging Mind: Potential and Limits". *The Gerontologist, 33*, 580-594.

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

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (1986). *Social foundation of thought & action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall. In J. Metcalfe and P. Shimamura (Eds.). *Metacognition: Knowing about Knowing*. (pp. 227-251), Cambridge: MIT Press.

Bandura, A. (1988). Self-efficacy conception of anxiety. *Anxiety Research, 1*, 77-98. In J. Metcalfe and P. Shimamura (Eds). *Metacognition: Knowing about Knowing*. (pp. 227-251), Cambridge: MIT Press.

Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology, 50*, 492-501.

Banks, W.P. (1970). Signal detection theory and human memory. *Psychological Bulletin, 74*, 81-99.

Barnes, A.E., Nelson, T.O., Dunlosky, J., Mazzoni, G., & Narens, L. (1995). An integrative system of metamemory components involved in retrieval. In D. Gopher & A. Koriat (Eds.), *Attention and Performance XVII: Cognitive Regulation of*

Performance: Interaction of theory and application (pp. 287-313), Cambridge, MA: MIT Press.

Baron, R.M., & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Baron, A., & Surdy, T.M. (1990). Recognition memory in older adults: Adjustment to changing contingencies. *Journal of the Experimental Analysis of Behaviour, 54*, 201-212.

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

Baxter, J.S. (under review). The Relative Influence of Leading Questions and Negative Feedback on Response Change on the Gudjonsson Suggestibility Scale (2): Implications for Forensic Interviewing. *Psychology, Crime & Law.*

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., & 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 diasability. *Journal of Applied Research in Intellectual Disabilities, 15*, 129-137.

Benton, A. L. & Hamsher, K. (1976). *Multilingual Aphasia Examination*. Iowa City: Aja Associates.

Bernbach, H.A. (1967). Decision processes in memory. *Psychological Review, 74*, 462-480.

Berry, J.M. (1999). Memory self-efficacy in its social cognitive context. In T.M. Hess & F. Blanchard-Fields (Eds), *Social Cognition and Aging* (pp. 69-96). San Diego, CA: Academic Press.

Berry, J.M., & West, R.L. (1993). Cognitive self-efficacy in relation to personal mastery and goal setting across the life span. *International Journal of Behavioural Development, 16,* 351–379

Berry, J.M., West, R.L., & Dennehy, D.M. (1989). Reliability and validity of the self – efficacy questionnaire. *Developmental Psychology, 25*, 701–713.

Berry, J.M., West, R.L., & Scogin, F.R. (1983). *Predicting everyday and laboratory memory skill.* Paper presented at the meeting of the Gerontological Society of America, San Francisco.

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

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

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.

Bock, R.D. (1975). *Multivariate Statistical Methods in behavioural research*. New York: McGraw-Hill, cited in Stevens, J.P. (2009), (5th Ed.). *Applied Multivariate Statistics for the Social Sciences.* Routledge, Hove.

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.

Bothwell, R.K., Deffenbacher, K.A., & Brigham, J.C. (1987). Correlations of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology, 72*, 691-695.

Bornstein, B.H., Witt, C.J., Cherry, K.E., & Greene, E. (2000), The suggestibility of older witnesses. In M.B. Rothman, B.D. Dunlop & P. Entzel (Eds.), *Elders, crime and the criminal justice system: Myth perception and reality in the 21st century* (pp. 149-161), New York: Springer.

Botwinick, J. (1978). *Aging and behaviour.* New York: Springer.

Bower, G.H. (1967). A multicomponent theory of a memory trace. In *The Psychology of Learning and Motivation* (Vol. 1), Spence KW, Spence JT (eds.), Academic Press: New York.

Brewer, N., & Day, K. (2005). The confidence-accuracy and decision latency-accuracy relationships in children's eyewitness identification. *Psychiatry, Psychology and Law, 12*, 119-128.

Brewer, N., & Weber, N. (2008). Eyewitness Confidence and Latency: Indices of Memory Processes Not Just Markers of Accuracy. *Applied Cognitive Psychology, 22*, 827-840.

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

Brewer, N., Keast, A., & Rishworth, A. (2002). The confidence-accuracy relationship in eyewitness identification: The effects of reflection and disconfirmation on correlation and calibration. *Journal of Experimental Psychology: Applied, 8*, 46-58.

Brewer, N., Weber, N., & Semmler, C. (2007). A role for theory in eyewitness identification research. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M. Toglia (Eds), *Handbook of eyewitness psychology: Memory for people* (Vol 2, pp. 201-218). Mahwah, NJ: Lawrence Erlbaum.

Brewer, N., Potter, R., Fisher, R.P., Bond, N., & Luszcz, M. (1999). Beliefs and data on the relationship between consistency and accuracy of eyewitness testimony.

Applied Cognitive Psychology, 13, 297-313. In R. Bull, T.Valentine, & T. Williamson (Eds.), *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions,* (pp. 121-134), Wiley-Blackwell: Chichester.

Brigham, J.C. & Malpass, R.S. (1985). The role of experience and contact in the recognition of faces of own-race and other-race persons. *Journal of Social Issues*, *41*, 139-155.

Brigham, J.C. & Ready, D.J. (1985). Own-race bias in line-up construction. *Law and Human Behaviour, 9*, 415-424.

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

Brown, J. (Ed.). (1976). *Recall and Recognition*. London: Wiley.

Brown, A.L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional Psychology* (pp. 367-406). Hillsdale, NJ: Erlbaum.

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

Bruce, P.R., Coyne, A.C., & Botwinick, J. (1982). Adult age differences in metamemory. *Journal of Gerontology, 37*, 354-357. In J. Metcalfe & A.P. Shimamura (Eds.), *Metacognition: Knowing about Knowing*, (pp. 227-251).

Bull, R. (2010). The investigative interviewing of children and other vulnerable witnesses: Psychological research and working/professional practice. *Legal and Criminological Psychology, 15,* 5-23.

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

Burgess, P.W., & Shallice, T. (1996). Confabulation and the control of recollection. *Memory, 4*, 359-411.

Burke, A., Heuer, F., & Reisberg, D. (1992). Remembering emotional events. *Memory & Cognition, 20*, 277-290.

Burke, D.M., & Light, L.L. (1981). Memory and aging: The role of retrieval processes. *Psychological bulletin, 90*, 513-546.

Burtt, H.E. (1948). *Applied Psychology*. New York: Prentice-Hall.

Campos, L., & Alonso-Quecuty M. L. (1999). The Cognitive Interview: Much More than Simply “Try Again”. *Psychology, Crime & Law*, *5*, 47-59.

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.

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

Cavanaugh, J.C., Feldman, J.M., & Hertzog, C. (1998). Memory beliefs as social cognition: A reconceptualization of what memory questionnaires assess. *Review of General Psychology, 2*, 48-65.

Cavanaugh, J.C., & Green, E.E. (1990). I believe, therefore I can: Self-efficacy beliefs in memory aging. In E.A. Lovelace (Ed.), *Aging and Cognition: Mental processes, self-awareness, and interventions* (pp. 189-230). Amsterdam: Elsevier.

Cavanaugh, J.C., & Poon, L.W. (1989). Metamemorial predictors of memory performance in young and older adults. *Psychology and Aging, 4*, 365-368.

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

Ceci, S.J., & Tabor, L. (1981), Flexibility and Memory: Are the elderly really less flexible? *Experimental Aging Research*, *7*, 147-158, in F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Chalfonte, B.L., & Johnson, M.K. (1996). Feature memory and binding in young and older adults. *Memory and Cognition, 24*, 403-416.

Chasteen, A.L., Bhattacharyya, S., Horhota, M., Tam, R., & Hasher, L. (2005). How feelings of stereotype threat influence older adults' memory performance. Experimental Aging Research, 31, 235-260. In F.I.M. Craik & T.Salthouse (3rd Ed.), *Handbook of aging and cognition* (251-310). Psychology Press, New York, NY.

Choynowski, M. (1959). Skala pamieci - podrecznik tymczasowy. [Memory Scale – Preliminary Manual]. Warsawa: PAN.

Christensen, H., Mackinnon, A.J., Korten, A.E., Jorm, A.F., Henderson, A.S., Jacomb, P., & Rodgers, B. (1999). An analysis of diversity in the cognitive performance of elderly community dwellers: Individual differences in change scores as a function of age. *Psychology and Aging, 14*, 365-379. In Chamorro – Premuzic, T. Furnham, A. & von Stumm, S. (Editors), *Handbook of Individual Differences* (in press).

Christianson, S.A., & Loftus, E.F. (1987). Memory for traumatic events. *Applied Cognitive Psychology, 1*, 225-239.

Christianson, S.A., & Loftus, E.F. (1991). Remembering emotional events: the fate of detailed information. *Cognition and Emotion, 5*, 81-108.

Clare, IC.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.

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

Clifford, B.R. & Scott, J. (1978), Individual and situational factors in eyewitness testimony. *Journal of Applied Psychology*, *63*, 352 – 359.

Cockburn, J. & Smith, P.T. (1991). The relative influence of intelligence and age on everyday memory. *Journal of Gerontology*, *46*, 31-36.

Cohen, G., & Faulkner, D. (1989), Age differences in source forgetting: Effects on reality monitoring and on eyewitness testimony. *Psychology and Aging*, *4*, 10 – 17.

Cohen, R.L., & Harnick, M.A. (1980). The susceptibility of child witnesses to suggestion. *Law and Human Behaviour, 4,* 201-210.

Committee on Pattern Jury Instructions of the District Judges Association of the Sixth Circuit (2005). *Sixth Circuit Criminal Pattern Jury Instructions*. Cincinnati, OH: Judicial Council for the U.S. Court of Appeals for the Sixth Circuit. In R. Bull, T.Valentine, & T. Williamson (Eds,), *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions,* (pp. 121-134), Wiley-Blackwell: Chichester.

Costa, P.T. Jr. & McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory.* Odessa, Florida. FL: Psychological Assessment Resources.

Coxon, P. & Valentine, T. (1997), The Effects of Age of Eyewitnesses on the Accuracy and Suggestibility of their Testimony, *Applied Cognitive Psychology*, *11*, 415 – 430.

Coyne, A.C. (1985). Adult age, presentation time, and memory performance. *Experimental Aging Research, 11,* 147-149. In J. Metcalfe & A.P. Shimamura (Eds.), *Metacognition: Knowing about Knowing*, (pp. 227-251)

Craik, F.I.M. (1977). Age differences in human memory. In J.E. Birren and K.W. Schaie (Eds.), *Handbook of the Psychology of Aging*. New York: Van Nostrand Reinhold, (pp. 384-420).

Craik, F.I.M. (1986). A functional account of age differences in memory. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Craik, F.I.M. (1994). Memory changes in normal aging. *Current directions in Psychological Science, 3 (5),* 155-158.

Craik, F.I.M., Byrd, M., & Swanson, J.M. (1987). Patterns of memory loss in three elderly samples. *Psychology and Aging*, *2*, 79-86. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Craik, F.I.M. & Jennings, J. (1992), Human Memory. In F.I.M. Craik & T.Salthouse (2nd Ed), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Craik, F.I.M., & McDowd, J.M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *13*, 474-479. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Craik, F. I.M., Morris, R.G., & Gick, M.L. (1990). Adult age differences in working memory. In R.A. Dixon, Lars Bäckman & Lars-Göran Nilsson, *New Frontiers in Cognitive Aging* (19-40), Oxford University Press.

Craik, F.I.M., & Simon, E. (1980), Age differences in memory: The roles of attention and depth of processing. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Crombag, H., Merckelbach, H., & Elffers, H. (2000). Other people's memory. *Psychology, Crime & Law, 6*, 251-265.

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.

Crook, T.H., & Larrabee, G.J. (1990). A self rating scale for evaluating memory in everyday life. *Psychology and Aging, 5*, 48-57.

Cutler, B., & Penrod, S.D. (1995). *Mistaken identification: The eyewitness, psychology and the law*. New York: Cambridge University Press

Cutler, B., Penrod, S.D., & Dexter, H.R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behaviour*, *14*, 185-191.

Cutler, B., Penrod, S,D., & Stuve, T.E. (1988). Juror decision making in eyewitness identification cases. *Law and Human Behaviour*, *12*, 41-55.

Da Silva, L. & Sunderland, A. (2010). Effects if immediate feedback and errorless learning on recognition memory processing in young and older adults. *Neuropsychological Rehabilitation, 20(1)*, 42-58.

Dahl, M., Allwood, C.M., Hagberg, B. (2009). The Realism in Older People's Confidence Judgments of Answers to General Knowledge Questions. *Psychology And Aging, 24, 1*, 234-238.

Dando, C. J., & Milne, R. (2009). The Cognitive Interview. In R. N. Kocsis (Ed.), *Applied Criminal Psychology: A Guide to Forensic behavioural Science.*

Danielsdottir, G., Sigurgeirsdottir, S., Einarsdottir, H. R., & Haraldsson, E. (1993). Interrogative suggestibility in children and its relationship with memory and vocabulary. *Personality and Individual Differences, 14*, 499-502.

Danion, J.M., Gokalsing, E., Robert, P., Massin-Krauss, M., & Bacon, E. (2001). Defective relationship between subjective experience and behaviour in schizophrenia. *American Journal of Psychiatry, 158*, 2064-2066.

Davies, G., Flin, R., & Baxter, J. (1986). The child witness. *The Howard Journal of Criminal Justice, 25*, 81-99.

Deffenbacher, K.A. (1991). A maturing of research on the behaviour of eyewitnesses. *Applied Cognitive Psychology, 5*, 377-402.

Deffenbacher, K.A. (2008). Estimating the Impact of Estimator Variables on Eyewitness Identification: A Fruitful Marriage of Practical Problem Solving and Psychological Theorizing. *Applied Cognitive Psychology, 22*, 815-826.

Dent, H. (1986). An experimental study of the effectiveness of different techniques of questioning mentally handicapped child witnesses. *British Journal of Clinical Psychology*, *25*, 13-17.

Devolder, P.A., Brigham, M.C., & Pressley, M. (1990). Memory performance awareness in younger and older adults. *Psychology and Aging, 5*, 291-303.

Dixon, R.A., Hertzog, C., & Hultsch, D.F. (1986). The multiple relationships among Metamemory in Adulthood (MIA) scales and cognitive abilities in adulthood. *Hum. Learn. 5*, 165-177.

Dixon, R.A., & Hultsch, D. F. (1983a). Metamemory and memory for text relationships in adulthood: A cross-validation study. *Journal of Gerontology. 38*: 689–694. In J. Metcalfe & A.P. Shimamura (Eds.), *Metacognition: Knowing about Knowing*, (pp. 227-251)

Dixon, R.A., & Hultsch, D. F. (1983b). Structure and development of metamemory in adulthood. *Journal of Gerontology, 38*, 682-688. In J. Metcalfe & A.P.

Shimamura (Eds.), *Metacognition: Knowing about Knowing*, (pp. 227-251), Cambridge:MIT Press.

Dixon, R.A., Hultsch, D. F., & Hertzog, C. (1988). The metamemory in adulthood (MIA) questionnaire. *Psychopharmacology Bulletin*, *24*, 671-688.

Dodson, C.S., Bawa, S., & Kruger, L.E. (2007). Aging, metamemory, and high confidence errors: A misrecollection Account. *Psychology and Aging, 22*, *No. 1*, 122-133.

Dodson, C.S., Bawa, S., & Slotnick, S.D. (2007). Aging, Source Memory, and Misrecollections. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33, No. 1*, 169-181.

Dodson, C.S., & Kruger, L.E. (2006). I misremember it well: Why older adults are unreliable witnesses. *Psychonomic Bulletin & Review, 13, No. 5*, 770-775.

Dornburg, C.C. & McDaniel, M.A. (2006). The Cognitive Interview Enhances Long-Term Free Recall of Older Adults. *Psychology and Aging, 21(1)*, 196-200.

Drake, K. E., Bull, R., & Boon, J.C.W. (2008). Interrogative suggestibility, self-esteem, and the influence of negative life events. *Legal and Criminological Psychology, 13*, 299-307.

Dux, M.C., Woodard, J.L., Calamari, J.E., Messina, M., Arora, S., Chik, H., Pontarelli, N. (2008). The moderating role of negative affect on objective verbal memory performance and subjective memory complaints in healthy older adults. *Journal of the International Neuropsychological Society, 14(2)*, 327-336.

Emmett, D., Clifford, B.R., & Gwyer, P. (20003). An investigation of the interaction between cognitive style and context reinstatement on the memory performance of eyewitnesses. *Personality and Individual Differences, 34(8),* 1495-1508.

Erber, J.T. (1974). Age differences in recognition memory. *Journal of Gerontology, 29*, 177-181.

Erber, J.T. & Rothberg, S.T. (1991). Here's looking at you: The relative effect of age and attractiveness on judgments about memory failure. *Journal of Gerontology: Psychological Sciences, 46*, 116-123.

Erber, J.T., Szuchman, L.T., & Prager, I.G. (1997). Forgetful but forgiven: How age and lifestyle affect perceptions of memory failure. *Journal of Gerontology: Psychological Sciences, 52B*, 303-307.

Erber, J.T., Szuchman, L.T., & Rothberg, S.T. (1990). Everyday memory failure: Age differences in appraisal and attribution. *Psychology and Aging, 5*, 236-241.

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

Fernandez, A. & Alonso, M.A. (2001). The relative value of environmental context reinstatement in free recall. *Psicologica, 22 (2)*, 253-266.

Fisher, R. (1998). Personal Communication.

Fisher, R.P. (2010). Interviewing cooperative witnesses. *Legal and Criminological Psychology, 15*, 25-38.

Fisher, R.P., & Castano, N. (2008). Cognitive interview. In B. Cutler (Ed.), *Encyclopedia of psychology and law*. Thousand Oaks, CA: Sage.

Fisher, R.P., & Patterson, T. (2004). *The relationship between consistency and accuracy of eyewitness memory*. Paper presented at 45th annual meeting of the Psychonomic Society: Minneapolis. In R. Bull, T.Valentine, & T. Williamson (Eds). *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions,* (pp. 121-134), Wiley-Blackwell: Chichester.

Fisher, R.P., Brewer, N., & Mitchell, G. (2009). The Relation between Consistency and Accuracy of Eyewitness Testimony: Legal versus Cognitive Explanations. In R. Bull, T.Valentine, & T. Williamson (Eds). *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions,* (pp. 121-134), Wiley-Blackwell: Chichester.

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

Fisher, R.P., Amador, M., & Geiselman, R.E. (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 interviewing techniques. *Journal of Police Science and Administration*, *15*, 177-185.

Fisher, R.P., Geiselman, R.E., Raymond, D.S., Jurkevich, L.M., & Warhaftig, M.L. (1987). Enhancing eyewitness memory: refining the cognitive interview. *Journal of Police Science and Administration*, *15*, 291-7.

Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new era of cognitive-development inquiry. *American Psychologist, 34,* 906-911. In J. Metcalfe & A.P. Shimamura (Eds.), *Metacognition: Knowing about Knowing,* (pp. 227-251).

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.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). 'Mini-Mental State': A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

Gabbert, F., Memon, A., Allan, K., & Wright, D.B. (2004), Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology*, *9*, 215 – 227.

Garcia-Bajos, E., & Migueles, M. (2003). False memories for script actions in a mugging account. *European Journal of Cognitive Psychology, 15*, 195-208.

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., Sullivan, S.J., Avetissian, I.V., & Prosk, A.L. (1984). Enhancement of eyewitness memory: An empirical evaluation of the cognitive interview. *Journal of Police and Science 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*, 401-12.

George, R.C. (1991). *A field evaluation of the cognitive interview*. Unpublished masters thesis, Polytechnic of East London.

George, R.C., & Clifford, B.R. (1996). Making the most of witnesses. *Policing, 8*, 185-198.

Gilewski, M.J., & Zelinski, E.M. (1986). Questionnaire assessment of memory compaints. In L.W. Poon (Ed.), *Handbook for clinical memory assessment of older adults* (pp. 93-107). Washington, DC: American Psychological Association.

Gilewski, M.J., Zelinski, E.M., & Schaie, K.W. (1990). The Memory Functioning Questionnaire for assessments of memory complaints in adulthood and old age. *Psychology and Aging, 5*, 482-490.

Gilewski, M.J., Zelinski, E.M., Schaie, K.W., & Thompson, L.W. (1983). *Abbreviating the Metamemory questionnaire: Factor structures and norms for adults.* Paper presented at the 91st Annual Meeting of the American Psychological Association, Anaheim, CA.

Glass, G., Peckham, P., & Saunders, J. (1972). Consequences of failure to meet assumptions underlying the fixed effects of variance and covariance. *Review of Educational Research, 42*, 237-288, cited in Stevens, J.P. (2009), (5th Ed.). *Applied Multivariate Statistics for the Social Sciences.* Routledge, Hove.

Goldsmith, M., & Koriat, A. (1999). The strategic regulation of memory reporting: Mechanisms and performance consequences. In D. Gopher & A. Koriat (Eds.), *Attention and Performance XVII: Cognitive Regulation of Performance: Interaction of theory and application* (pp. 287-313), Cambridge, MA: MIT Press

Goldsmith, M., & Koriat, A. (2008). The strategic regulation of regulation of memory accuracy and informativeness. In A. Benjamin & B. Ross (Eds.), *The psychology of learning and motivation* (*Vol. 48*, pp. 1-60). London Academic Press.

Goldsmith, M., Koriat, A., & Pansky, A. (2005). Strategic regulation of grain size in memory reporting over time. *Journal of Memory and Language, 52*, 505-525.

Goldsmith, M., Koriat, A., & Weinberg-Eliezer, A. (2002). Strategic regulation of grain size memory reporting. *Journal of Experimental Psychology: General, 131*, 73-95.

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, S.K. & Clark, W.C. (1974). Application of signal detection theory to prose recall and recognition in elderly and young adults. *Journal of Gerontology, 29*, 64-72.

Gorenstein, G.W., & Ellsworth, P.C. (1980). Effect of choosing an incorrect photograph on a later identification by an eyewitness. *Journal of Applied Psychology, 65*, 616-622.

Granhag, P.A. (1997). Realism in eyewitness confidence as a function of type of event witnessed and repeated recall. *Journal of Applied Psychology*, *82*, 599-613.

Granhag, P.A., Stromwall, L.A., & Allwood, C.M. (2000). Effects of Reiteration, Hindsight Bias, and Memory on Realism in Eyewitness Confidence, *Applied Cognitive Psychology*, *14***,** 397 – 420.

Green, D.M., & Swets, J.A. (1966). *Signal detection theory and psychophysics*. New York: Wiley

Greenberg, M.S., Westcott, D.R., & Bailey, S.E. (1998). When believing is seeing: The effect of scripts on eyewitness memory. *Law and Human Behaviour, 22*, 685-694.

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

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. (1995). The Effects of Interrogative Pressure on Strategic Coping. *Psychology, Crime & Law, 1*, 309-318.

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 John Wiley and Sons.

Gudjonsson, G.H. (2006). The psychological vulnerabilities of witnesses and the risk of false accusations and false confessions. In A. Heaton-Armstrong, E. Shepherd, G.

Gudjonsson, & D. Wolchover (Eds.), *Witness Testimony. Psychological, investigative and evidential perspectives* (pp. 61-75). Oxford: Oxford University Press.

Gudjonsson, G.H. (2010). Psychological vulnerabilities during police interviews. Why are they important? *Legal and Criminological Psychology, 15*, 161-175.

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., Hannesdottir, K., Petursson, H., & Bjornsson, G. (2002). The effects of alcohol withdrawal on mental state, interrogative suggestibility and compliance: An experimental study. *Journal of Forensic Psychiatry, 13*, 53-67.

Gudjonsson, G.H., & Henry, L. (2003). Child and adult witnesses with intellectual disability: the importance of suggestibility. *Legal and Criminological Psychology*, *8,* 241-252.

Gudjonsson, G.H., Kopelman, M.D., & MacKeith, J.A.C. (1999). Unreliable admissions to homicide: A case of mis-diagnosis of amnesia and misuse of abreaction technique. *British Journal of Psychiatry, 174*, 455-459.

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

Gudjonsson, G.H., & MacKeith, J. (1997). *Disputed confessions and the criminal justice system. Maudsley Discussion Paper No 2*. London: Institute of Psychiatry.

Gudjonsson, G.H., & MacKeith, J. (1982). False confessions: Psychological effects of interrogation. A discussion paper. In A. Trankell (Ed.), *Reconstructing the past: The role of psychologists in criminal trials* (pp. 253-269. Deventer, The Netherlands: Kluwer.

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. (2003). The Relationship of Compliance with Coping Strategies and Self-Esteem. *European Journal of Psychological Assessment, 19(2)*, 117-123.

Gudjonsson, G.H., Sigurdsson, J.F., Einarsson, E., Bragason, O.O., & Newton, A.K. (2008). Interrogative suggestibility, compliance and false confessions among prisoners and their relationship with attention deficit hyperactivity disorder (ADHD) symptoms. *Psychological Medicine, 38,* 1037-1044.

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.

Gudjonsson, G.H., Young, S., & Bramham, J. (2007). Interrogative suggestibility in adults diagnosed with attention-deficit hyperactivity disorder (ADHD). A potential vulnerability during police questioning. *Personality and Individual Differences, 43*, 737-745.

Hansson, P., Rönnlund, M., Juslin, P., & Nilsson, L.G. (2008). Adult age differences in the realism of confidence judgments: Overconfidence, format dependence, and cognitive predictors. *Psychology And Aging, 23, 3,* 531-544.

Hart, R.P., Kwentus, J.A., Wade, J.B., & Taylor, J.R. (1988). Modified Wisconsin Sorting Test in elderly normal, depressed and demented patients. *The Clinical Neuropsychologist, 2*, 49-56.

Hasher, L., Tonev, T., Lustig, C., & Zacks, R.T. (2001). Inhibitory control, environmental support, and self-initiated processing in aging. In M. Naveh – Benjamin, M. Moscovitch, & H.L. Roediger, III (Eds.), *Perspectives on Human Memory and Cognitive Aging: Essays in Honour of Fergus Craik*, (pp. 148-160). New York: Psychology Press.

Hasher, L. & Zacks, R.T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*, 356-388. In G.H. Bower (Ed), *The psychology of learning and motivation, 22*, (193-225). New York: Academic Press.

Hasher, L. & Zacks, R.T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist*, *39*, 1372-1388. In G.H. Bower (Ed.), *The psychology of learning and motivation, 22*, (193-225). New York: Academic Press.

Hasher, L. & Zacks, R.T. (1988). Working memory, comprehension and aging: A review of a new view. In G.H. Bower (Ed.), *The psychology of learning and motivation, 22*, (193-225). New York: Academic Press.

Hashtroudi, S., Johnson, M.K., & Chrosniak, L.D. (1990). Aging and qualitative characteristics of memories for perceived and imagined complex events. *Psychology and Aging*, *4*, 106-112. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Hastie, R., Landsman, R., & Loftus, E. (1978). Eyewitness Testimony: The Dangers of Guessing. *Jurimetrics Journal*, *19*, 1 – 8.

Havard C, & Memon A. (2009). The influence of face age on identification from a video line-up: A comparison between older and younger adults. *MEMORY, 17 (8)*, 847-859.

Hayes, B.K., & Delamothe, K. (1997). Cognitive Interviewing procedures and suggestibility in children's recall. *Journal of Applied Psychology*, *82*, 562-577.

Heath, W.P., & Erickson, J.R. (1998). Memory for central and peripheral actions and props after varied post-event presentation. *Legal and Criminological Psychology, 3*, 321-346.

Heckhausen, J., & Baltes, P.B. (1991). Perceived controllability of expected psychological change across adulthood and old age. *Journal of Gerontology: Psychological Sciences, 46*, 165-173.

Heckhausen, J. Dixon, R.A., & Baltes, P.B. (1989). Gains and losses in development throughout adulthood as perceived by different age gtroups. *Developmental Psychology, 25*, 109-121.

Henkel, L.A., Johnson, M.K., & De Leonardis, D.M. (1998). Aging and source monitoring: cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General, 127*: 371-376.

Hernández-Fernaud, E.H., & Alonso-Quecuty, M.A. (1997). The cognitive interview and lie detection: A new magnifying glass for Sherlock Holmes? *Applied and Cognitive Psychology, 11,* 55-68.

Herrmann, D.J., & Neisser, U. (1978). An inventory of everyday memory experiences. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical Aspects of Memory* (pp. 35-51). New York: Academic Press.

Hertzog, C., Dixon, R.A., & Hultsch, D.F. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging, 5*, 215-227.

Hertzog, C. & Hultsch, D.F. (2000). Metacognition in adulthood and old age. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Hertzog, C., McGuire, C.L., & Lineweaver, T.T. (1998). Aging, attributions, perceived control, and strategy use in a free recall task. Aging, *Neuropsychology, and Cognition, 5*, 85-106.

Hess, T.M. (2005). Memory and Aging in Context. *Psychological Bulletin*, *131*(3), 383-406.

Hess, T.M., Auman, C., Colcombe, S.J., & Rahhal, T.A. (2003). The impact of stereotype threat on age differences in memory performance. *Journal of Gerontology: Psychological Sciences, 58B*, 3-11.

Hess, T.M., Donley, J., & Vandermaas, M.O. (1989). Aging-related changes in the processing and retention of script information. *Experimental Aging Research*, *15*, 89-96. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Hess, T.M., & Hinson, J.T. (2006). Age-related variation in the Influences of Aging Stereotypes on Memory in Adulthood. *Psychology and Aging, 21(3)*, 621-625.

Hess, T.M., Hinson, J.T., & Statham, J.A. (2004). Implicit and explicit stereotype activation effects on memory: Do age and awareness moderate the impact of priming? *Psychology and Aging, 19*, 495-505.

Heuer, F., & Resiberg, D. (1990). Vivid memories of emotional events: the accuracy of remembered minutiae. *Memory & Cognition, 18*, 496-506.

Higham, P.A. (2002). Strong cues are not necessarily weak: Thomson & Tulving (1970) and the encoding specificity principle revisited. *Memory and Cognition, 30*, 67-80.

Higham, P.A. (2007). No special K! A signal detection framework for the strategic regulation of memory accuracy. *Journal of Experimental Psychology: General, 136*, 1-22.

Higham, P.A., & Tam, H. (2005). Generation failure: Estimating metacognition in cued recall. *Journal of Memory and Language, 52*, 595-617.

Hill, R.D., van Boxtel, M.P.J., Ponds, R., Houx, P.J., & Jolles, J. (2005). Positive affect and its relationship to free recall memory performance in a sample of older Dutch adults from the Maastricht Aging Study. *International Journal of Geriatric Psychiatry, 20*, 429-435.

Hilgard, E.R., & Loftus, E.F. (1979). Effective interrogation of the eyewitness. *The International Journal of Clinical and Experimental Hypnosis, 27*, 342-357.

Holliday, R.E. (2003a). The effect of a prior Cognitive Interview on children's acceptance of misinformation. *Applied Cognitive Psychology*, *17 (4)***,** 443-457.

Holliday, R.E. (2003b). Reducing misinformation effects in children with Cognitive Interviews: dissociating recollection and familiarity. *Child Development, 74*, 728-751.

Holliday, R.E., & Albon, A.J. (2004). Minimising misinformation effects in young children with cognitive interview mnemonics. Applied Cognitive Psychology, 18, 263-281. In R. Bull, T.Valentine, & T. Williamson (Eds). *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions,* (pp. 137-160). Wiley-Blackwell: Chichester.

Holliday, R.E., Brainerd, C.J., Reyna, V.F., & Humphries, J.E. (2009). The Cognitive Interview: Research and Practise across the Lifespan. In R. Bull, T.Valentine, & T. Williamson (Eds.). *Handbook of Psychology of Investigative Interviewing:*

Current Developments and Future Directions, (pp. 137-160). Wiley-Blackwell: Chichester.

Holliday, R.E., Humphries, J.E., Milne, R., Memon, A., Houlder, L., Lyons, A., & Bull, R. (in press). Reducing Misinformation in Older Adults. Psychology and Aging.

Home Office (2008). Crime reduction – distraction burglary. Retrieved August, 2010, from http://www.crimereduction.homeoffice.gov.uk/boguscaller/boguscaller8.htm

Home Office (2008). *Police and Criminal Evidence Act 1984, Codes of Practice – Code C detention, treatment and questioning of persons by police*. London:Author.

Horn, J.L., & Cattell, R.B. (1966). Refinement and Test of the Theory of Fluid and Crystallised Intelligence. *Journal of Educational Psychology, 57*, 253-270.

Horton, S., Baker, J., Pearce, W., & Deakin, J.M. (2010). Immunity to Popular Stereotypes of Aging? Seniors and Stereotype Threat. *Educational Gerontology, 36*, 353-371.

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

Howitt, D. (2002). *Forensic and Criminal Psychology*, Pearson: Prentice Hall.

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

Hultsch, D.F., Hertzog, C., & Dixon, R.A., (1987). Age differences in Metamemory: Resolving the inconsistencies. *Canadian Journal of Psychology, 41*:193-208.

Hultsch, D.F., Hertzog, C., Dixon, R.A., & Davidson, H. (1988). Memory self – knowledge and self-efficacy in the aged. In M.L. Howe & C.J. Brainerd (Eds.), *Cognitive Development in adulthood: Progress in cognitive development research* (pp. 65–92). New York: Springer-Verlag.

Hummert, M.L., Garstka, T.A., & Shaner, J.L. (1997). Stereotyping of older adults: The role of target facial cues and perceiver characteristics. *Psychology and Aging, 12(1)*, 107-114.

Hummert, M.L., Garstka, T.A., Shaner, J.L., & Strahm, S. (1994). Stereotypes of the elderly held by young, middle-aged, and elderly adults. *Journal of Gerontology: Psychological Sciences, 49*, 240-249.

Hyman, I.E., Husband, T.H., & Billings, F.J. (1995). False memories of childhood experiences. *Applied Cognitive Psychology*, *9*, 181-197.

Ibabe, I., & Sporer, S.L. (2004). How You Ask Is What You Get: On the Influence of Question Form on Accuracy and Confidence. *Applied Cognitive Psychology*, *18,* 711-726.

Ingram, R.E., Lumry, A.E., Cruet, D., & Sieber, W. (1987). Attentional Processes In Depressive-Disorders. *Cognitive Therapy And Research, 11(3)*, 351-360.

Innocence Project. (2010). Innocence Project. Retrieved June, 2010, from http://www.innocenceproject.org

Isingrini, M., Hauer, K., & Fontaine, R. (1996). Effet du vieillissement sur les réponses basées sur la familiarité and sur la recherché en situation de reconnaissance. *L'Année Psychologique, 96*, 263-281.

Jacoby, L.L. (1999). Deceiving the elderly: Effects of accessibility bias in cued recall performance. *Cognitive Neuropsychology, 16*, 417-436.

Jacoby, L.L., Debner, J.A., & Hay, J.F. (2001). Proactive interference, accessibility bias, and process dissociations: Valid subjective reports of memory. *Journal of Experimental Psychology: Learning, Memory & Cognition, 27*, 686-700.

Jacoby, L.L., Bishara, A.J., Hessels, S., & Toth, J.P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General*, *134*, 131-148.

Jacoby, L.L. & Hollingshead, A. (1990). Toward a generate/ recognise model of performance on direct and indirect tests of memory. *Journal of Memory and Language, 29*, 433-454.

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-existant film. *Applied Cognitive Psychology, 20*, 591-596.

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993), Source monitoring *Psychological Bulletin*, *144*, 3-28.

Jorm, A.F., Christensen, H., Henderson, A.S., Korten, A.E., Mackinnon, A., & Scott, R. (1994). Complaints of cognitive decline in the elderly: A comparison of reports by subjects and informants in a community survey. *Psychological Medicine, 24(2),* 365-374.

Jorm, A.F., Christensen, H., Korten, A.E., Henderson, A.S., Jacomb, P.A., & Mackinnon, A. (1997). Do cognitive complaints either predict future cognitive decline or reflect past cognitive decline? A longitudinal study of an elderly community sample. *Psychological Medicine, 27(1),* 91-98.

Juslin, P., Olsson, H., & Winman, A. (1995). Calibration and diagnosticity of confidence in eyewitness identification. Comments on what cannot be inferred from a low confidence – accuracy correlation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *22*, 1304-1316.

Karpel, M.E., Hoyer, W.J., & Toglia, M.P. (2001), Accuracy and Qualities of Real and Suggested Memories: Nonspecific Age Differences. *Journal of Gerontology: PSYCHOLOGICAL SCIENCES*, *56B*, *No. 2*, 103-110.

Kato, T. (1985). Semantic-memory sources of episodic retrieval failure. *Memory & Cognition, 13*, 442-452.

Kausler, D.H. (1991). *Experimental psychology, cognition, and human aging* (2nd ed.). New York: Springer-Verlag.

Keast, A., Brewer, N., & Wells, G.L. (2007). Children's metacognitive judgments in an eyewitness identification task. *Journal of Experimental Child Psychology, 97*, 286-314.

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.

Keith, T.Z. (2006). *Multiple Regression and Beyond*. Pearson Education, Inc.

Kebbell, M., & Milne, R. (1998). Police officers perceptions of eyewitness factors in forensic investigations. *Journal of Social Psychology*, *138*, 323-330.

Kelley, C.M., & Jacoby, L.L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language, 35*, 157-175.

Kelley, C.M., & Lindsay, D.S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32*, 1-24.

Kelley, C.M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language, 48*, 704-721.

Kemper, S. (1994). "Elderspeak": Speech accommodation to older adults. *Aging and Cognition, 1*, 17-28.

Kintsch, W. (1967). Memory and decision aspects of recognition learning. *Psychological Review, 74*, 496-504.

Klatzky, R.l., & Erdelyi, M.H. (1985). The Response Criterion Problem In Tests Of Hypnosis And Memory. *International Journal Of Clinical And Experimental Hypnosis, 33, 3*, 246-257.

Köhnken, G., Milne, R., Memon, A., & Bull, R. (1999). The cognitive interview: A meta-analysis. *Psychology, Crime, & Law*, *5*, 3-27.

Köhnken, G., Schimossek, E., Aschermann, E., & Hofer, E. (1995). The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Psychology, 80*, 671-684.

Köhnken, G.,Thurer, C., & Zoberbier, D. (1994). The cognitive interview: Are the investigators' memories enhanced too? *Applied Cognitive Psychology, 8*, 13-24.

Koren, D., Poyurovsky, M., Seidman, L.J., Goldsmith, M., Wenger, S., & Klein, E.M. (2005). The neuropsychological basis of competence to consent in first-episode schizophrenia: A pilot metacognitive study. *Biological Psychiatry, 57*, 609-616.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*, 609-639.

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123*, 297-316.

Koriat, A., & Goldsmith, M. (1996). Monitoring and Control Processes in the Strategic Regulation of Memory Accuracy. *Psychological Review, 103(3)*, 490-517.

Koriat, A., & Goldsmith, M., & Halamish, V. (2008). Control processes in voluntary remembering. In I.H.L. Roediger (Ed.), *Cognitive psychology of memory* (*Vol. 2*, pp. 307-324). Oxford, UK: Elsevier.

Koriat, A., Lichtenstein, S., & Fischoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 107-118.

Koutstaal, W. (2003). Older adults encode-but do not always use- perceptual details: Intentional vs. unintentional effects of detail on memory judgment. *Psychological Science*, *14*, 189-193.

Koutstaal, W. (2006). Flexible remembering. *Psychonomic Bulletin and Review, 13*, 84-91.

Koutstaal, W., Reddy, C., Jackson, E.M., Prince, S., Cendan, D.L., & Schacter, D.L. (2003). False recognition of abstract versus common objects in older and younger adults: Testing the semantic categorization account. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 499-510.

Koutstaal, W., & Schacter, D.L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language, 37*, 555-583.

Koutstaal, W., Schacter, D.L., & Brenner, C. (2001). Dual task demands and gist-based false recognition of pictures in younger and older adults. *Journal of Memory and Language, 44*, 399-426.

Koutstaal, W., Schacter, D. L., Galluccio, L., & Stofer K.A. (1999). Reducing Gist-Based False Recognition in Older Adults: Encoding and Retrieval Manipulations. *Psychology and Aging, 14(2)*, 220-237.

Krech, D., & Crutchfield, R.S. (1948). *Theory and problems of social psychology*. New York: McGraw-Hill

Kroll, N.E.A., Knight, R.T., Metcalfe, J., Wolf, E.S., & Tulving, E. (1996). Cohesion failure as a source of memory illusions. *Journal of Memory and Language, 35*, 176-196.

Lachman, J.L., & Lachman, R. (1980). Age and actualization of world knowledge. In L.W. Poon, J.L. Fozard, L.S. Cermak, D. Arenberg, & L.W. Thompson (Eds.), *New directions in memory and aging* (pp. 285-311). Hillsdale, NJ: Erlbaum.

Lachman, M.E. (1986). Locus of control in aging research: A case for multidimensional and domain-specific assessment. *Psychology and Aging, 1*, 34-40.

Lachman, M.E., Bandura, M., Weaver, S.L., & Elliott, E. (1995). Assessing control beliefs: The Memory Controllability Inventory. *Aging and Cognition, 2*, 67-84.

Lachman, M.E., & McArthur, L.Z. (1986). Adult age differences in causal attributions for cognitive, physical, and social performance. *Psychology and Aging, 1*, 127-132.

Lachman, M.E., Steinberg, E.S., & Trotter, S.D. (1987). Effects of control beliefs and attributions on memory self-assessments and performance. *Psychology and Aging, 2*, 266-271.

Levy, B. (1996). Improving memory in old age by implicit self-stereotyping. *Journal of Personality and Social Psychology, 71*, 1092-1107.

Lieppe, M.R. (1980). Effects of integrative and memorial processes on the correspondence of eyewitness accuracy and confidence. *Law and Human Behaviour, 4*, 261-274.

Lieppe, M.R., Manion, A.P., & Romanczyk, A. (1992). Eyewitness persuasion: How and how well do fract finders judge the accuracy of adults' and children's memory reportsd? *Journal of Personality and Social Psychology, 63*, 181-197. In R. Bull, T.Valentine, & T. Williamson (Eds.). *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions,* (pp. 121-136). Wiley-Blackwell: Chichester.

Light, L.L., Patterson, M.M., Chung, C., & Healy, M.R. (2004). Effects of repetition and response deadline on associative recognition in young and older adults. *Memory and Cognition, 32,* 1182-1193.

Likert, R.A. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140, 55.* In Hugh Collican (4th Ed.), *Research Methods and Statistics in Psychology.* Hodder & Stoughton.

Lindsay, R.C.L., Wells, G.L., & O'Connor, F.J. (1989). Mock juror belief of accurate and inaccurate eyewitnesses. *Law and Human Behaviour, 13*, 333-339. In R.

Bull, T.Valentine, & T. Williamson (Eds.), *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions,* (pp. 205-220). Wiley-Blackwell: Chichester.

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.

Lipmann, O. (1935). Methoden der Aussagepsychologie [Methods of eyewitness testimony]. In E. Abderhalden (Ed.), *Handbuch der biologischen Arbeitsmethoden* (Abt. VI, Teil C. II 2, Band 3, pp. 967-1056). Berlin: Urban & Schwarzenberg.

Lipton, J.P. (1977). On the psychology of eyewitness testimony. *Journal of Applied Psychology, 62*, 90-95.

List, J. (1986). Age and schematic differences in the reliability of eyewitness testimony. *Developmental Psychology, 22*, 50-57.

Lockhart, R.S., & Murdock, B.B. (1970). Memory and the theory of signal detection. *Psychological Bulletin, 74*, 100-109.

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

Loftus, E.F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

Loftus, E.F. (1982). Memory and its distortions. In A.G. Kraut (Ed.), *G.Stanley Hall lectures* (pp. 119-154). Washington, DC: American Psychological Association.

Loftus, E. (1989). The psychology of eyewitness testimony. In Raskin, D. (Ed.), *Psychological Methods in Criminal Investigation and Evidence*. New York: Springer.

Loftus, E.F., & Hoffman, H.G. (1989). Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General, 118*, 100-104.

Loftus, E.F., Levidow, B., & Deunsing, (1992), Who remembers best: Individual differences in memory for events that occurred at a science museum. *Applied Cognitive Psychology*, *6*, 93 – 10.

Loftus, E, F., Loftus, G.R., & Messo, J. (1987). Some facts about ‘weapon focus’. *Law and Human Behaviour, 11*, 55-62.

Loftus, E.F., Miller, D.G., & Burns, H.J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning & Memory*, *4*, 19-31.

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

Loftus, E.F., & Pickrell, J.E. (1995). The formation of false memories, *Psychiatric Annals*, *25*, 720-725.

Lovelace, E.A. (1990). Aging and metacognitions concerning memory function. In E.A, Lovelace (Ed.), *Aging and cognition: Mental processes, self awareness, and interventions* (pp. 157-188). Amsterdam: North Holland.

Luo, L. & Craik, F.I.M. (2008). Aging and Memory: A Cognitive Approach. *La Revue canadienne de psychiatrie, 53, 6*.

Lustig, C., Hasher, L., & Tonev, S.T. (2001). Inhibitory control over the present and the past. *European Journal of Cognitive Psychology*, *13*, 107-122.

Luszcz, M.A. & Hinton, M. (1995). Domain- and task-specific beliefs about memory in adulthood: A microgenetic approach. *Australian Journal of Psychology, 47*, 54-59.

Luus, C., & Wells, G.L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology, 79*, 714-723.

Magnussen, S., Andersson, J., Cornoldi, C., De Beni, R., Endestad, T., Goodman, G.S., et al., (2006). What people believe about memory. *Memory, 14*, 595-613.

Malpass, R. (1996). Enhancing eyewitness testimony. In S.L. Sporer, R.S. Malpass and G. Köhnken (Eds.) *Psychological Issues in Eyewitness Identification*. Hillsdale, NJ: Lawrence Erlbaum.

Marquie, J.C., & Huet, N. (2000). Age differences in feeling-of-knowing and confidence judgments as a function of knowledge domain. *Psychology And Aging***,** *15, 3,* 451-461.

McCauley, M.R. & Fisher, R.P. (1995). Facilitating children's eyewitness recall with the revised cognitive interview. *Journal of Applied Psychology, 80*, 510-516.

McCauley, M.R. & Fisher, R.P. (1996). Enhancing children's eyewitness testimony with the cognitive interview. In G. Davies, S. Lloyd-Bostock, M. McMurran, & J.C.

Wilson (Eds.), *Psychology, law and criminal justice* (pp. 127-133). Berlin: de Gruyter.

McCoy, S.L., Tun, P.A., Cox, L.C., Colangelo, M., Stewart, R.A., & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *Quarterly Journal of Experimental Psychology, 58A*, 22-33. In F.I.M. Craik & T.Salthouse (3rd Ed.), *The Handbook of aging and cognition* (251-310). Psychology Press, New York. NY.

McDaniel, M.A., Einstein, G.O., & Jacoby, L.L. (2008). New Considerations in Aging and Memory: The Glass May Be Half Full. In F.I.M. Craik & T.Salthouse (3rd Ed.), *The Handbook of aging and cognition* (251-310). Psychology Press, New York. NY.

McDougall, W. (1908). *An introduction to social psychology*. London: Methuen.

McGroarty, A. & Baxter, J.S. (2007). Interrogative pressure in simulated forensic interviews: The effects of negative feedback: *British Journal of Psychology, 98*, 455-465.

McGroarty, A. & Baxter, J.S. (2009), Interviewer behaviour, interviewee self-esteem and response change in simulated forensic interviews. *Personality and Individual Differences, 47*, 642-646.

McLean, M. (1995). Quality Investigation? Police interviewing of witnesses. *Medicine, Science and the Law, 35*, 116-112.

McMahon, M. (2000). The Effect of the Enhanced Cognitive Interview on Recall and Confidence in Elderly Adults. *Psychiatry, Psychology and Law*, *7 (1)*, 9-32.

Mello, E., & Fisher, R. (1996). Enhancing older adult eyewitness memory with the cognitive interview. *Applied Cognitive Psychology*, *10*, 403-417.

Memon, A., & Bartlett, J. (2002). The effect of verbalisation on face recognition in young and old adults. *Applied Cognitive Psychology, 16,* 635-650.

Memon, A., Bartlett, J., Rose, R., & Gray, C. (2003). The ageing eyewitness: Effects of age on face, delay and source memory. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 6*, 338-345.

Memon, A., & Gabbert, F. (2003a). Improving the identification accuracy of senior witnesses: Do pre-lineup questions and sequential testing help? *Journal of Applied Psychology, 88,* 341-347.

Memon, A., & Gabbert, F. (2003b). Unravelling the effect of sequential lineup presentation in culprit present lineups. *Applied Cognitive Psychology, 17*, 1-12.

Memon, A., Holley, A., Wark, L., Bull, R., & Köhnken, G. (1996). Reducing suggestibility in child witness interviews. *Applied Cognitive Psychology*, *8*, 641-659.

Memon, A., Meissner, C.A., & Fraser, J. (2010). THE COGNITIVE INTERVIEW: A Meta-analytic Review and Study Space Analysis of the Past 25 Years. *Psychology, Public Policy & Law, 16(4)*, 340-372.

Memon, A., & Stevenage, S. (1996). A clarification of the importance of comparison groups and accuracy rates with the cognitive interview. *PSYCOLOQUY 7 (6)* witness-memory.1.memon

Memon, A., Wark, L., Bull, R., & Köhnken, G. (1997). Isolating the effects of the Cognitive Interview techniques. *British Journal of Psychology*, *88(2)*, 179-197.

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.* Chichester, Wiley.

Merckelbach, H., Muris, P., Wessel, I., & Van Koppen, P.J. (1998). The Gudjonsson Suggestibility Scale (GSS): Further data on its reliability, validity, and metacognition correlates. *Social Behaviour and Personality, 26*, 203-210.

Migueles, M., & Garcia-Bajos, E. (1999). Recall, recognition and confidence patterns in eyewitness testimony. *Applied Cognitive Psychology, 13*, 237-256.

Migueles, M., & Garcia-Bajos, E. (2004). Esto es un atracto! Sesgos de la tipicidad en la memoria de testigos. *Estudios de Psicologia, 25*, 331-342.

Milne, R. (1999). The cognitive interview: Current research and applications. *Psychology Crime & Law*, 5 (1-2), 1-2

Milne, R. & Bull, R. (2001). *Investigative interviewing: Psychology and practice.* Chichester: John Wiley & Son.

Milne, R. & Bull, R. (2002). Back to basics: A componential analysis of the original Cognitive Interview mnemonics with three age groups. *Applied Cognitive Psychology*, *16(7),* 743-753.

Milne, R. & Bull, R. (2003). Does the Cognitive Interview help children to resist the effects of suggestive questioning? *Legal and Criminological Psychology*, *8(1)*, 21-38.

Milne, R., & Bull, R. (2006). Interviewing victims of crime, including children and people with intellectual disabilities. In M.R. Kebbell, & G. Davies (Eds.), *Practical psychology for forensic investigations and prosecutions*. Chichester: Wiley.

Milne, R., Clare, I. & Bull, R. (1996). How effective is the cognitive interview as an investigative tool for use with adults with mild intellectual disabilities (mental retardation)? *Poster presented at the American Psychology – Law Society Conference*, Hilton Head, SC, March.

Mitchell, K.J., Johnson, M.K., & Mather, M. (2003), Source Monitoring and Suggestibility to Misinformation: Adult Age-Related Differences. *Applied Cognitive Psychology*, *17*, 107 – 119.

Mitchell, K.J., Johnson, M.K., Raye, C.L., & D'Esposito, M.D. (2000). fMRI evidence of age-related hippocampal dysfunction in feature binding in working memory. *Cognitive Brain Research, 10*, 197-206.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100.

Morrell, H.E.R., Gaitan, S., & Wixted, J.T. (2002). On the nature of the decision axix in signal-detection-based models of recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28*, 1095-1110, in In R. Bull, T.Valentine, & T. Williamson (Eds.). *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions,* (pp. 205-220). Wiley-Blackwell: Chichester.

Mueller-Johnson, K., Toglia, M.P., Sweeney, C.D., & Ceci, S.J. (2007). The Perceived Credibility of Older Adults as Witnesses and its Relation to Ageism. *Behavioural Sciences and the Law, 25*, 355-375.

Murdock, B.B. (1974). *Human memory: Theory and data*. New York: Wiley.

Murdock, B.B. (1982). Recognition memory. In C.R. Puff (Ed.), *Handbook of research methods in human memory and cognition*. New York Academic Press.

Murphy, M.D., Sanders, R.E., Gabriesheski, A.S., & Schmitt, F.A. (1981). Metamemory in the aged. *Journal of Gerontology, 36*, 185-193. In J. Metcalfe & A.P. Shimamura (Eds.), *Metacognition: Knowing about Knowing*, (pp. 227-251)

Naveh – Benjamin, M. (2000). Adult age differences in memory performance: Tests of an Associate Deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26* (5), 1170-1187.

Naveh – Benjamin, M., Hussain, Z., Guez, J., & Bar-on, M. (2003). Adult age differences in episodic memory: Further support for an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 826-837.

Neil v Biggers, 409 U.S. 188 (1972).

Neisser, U. (1988). Time present and time past. In M.M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (*Vol. 2*, pp. 545-560). Chichester, England: Wiley.

Nelson, T.O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95*, 109-113.

Nelson, T.O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.). *The psychology of learning and motivation* (*Vol 26*). New York: Academic Press.

Niemi, H. (2002). Active learning - a cultural change needed in teacher education and schools. *Teaching And Teacher Education, 18, 7*, 763-780.

Nilsson, L.-G., Bäckman, L., Erngrund, K., Nyberg, L., Adolfsson, R., Bucht, G., et al. (1997). The Betula Prospective Cohort Study: Memory, health and aging. *Aging, Neuropsychology, and Cognition, 4,* 1-32.

Norman, D.A., & Wickelgren, W.A. (1969). Strength theory of decision rules and latency in retrieval from short-term memory. *Journal of Mathematical Psychology*, *6*, 192-208.

Nyberg L., Maitland S.B., Ronnlund, M., Bäckman, L., Dixon, R.A., Wahlin, A., & Nilsson, L.G. (2003). Selective adult age differences in an age-invariant multifactor model of declarative memory. *Psychology And Aging, 18, 1,* 149-160.

Oaksford M., Morris F., Grainger B., & Williams, J.M.G. (1996). Mood, reasoning, and central executive processes. *Journal Of Experimental Psychology-Learning Memory And Cognition, 22 (2)*, 476-492.

Odinot, G., Wolters, G., & van Koppen, P. (2009). Eyewitness Memory of a Supermarket Robbery: A Case Study of Accuracy and Confidence After 3 Months. *Law and Human Behaviour*, *33(6)*, 506-514.

Odinot, G., Wolters, G., & Lavender, T. (2009). Repeated Partial Eyewitness Questioning Causes Confidence Inflation but Not Retrieval Induced Forgetting. *Applied Cognitive Psychology*, *23(1)*, 90-97.

Office of National Statistics (2009). Retrieved June 2010 from http://www.statistics.gov.uk/cci/nugget.asp?id=949

Old, S.R. & Naveh-Benjamin, M. (2008). Memory for People and Their Actions: Further Evidence for an Age-Related Associate Deficit. *Psychology and Aging, 23(2)*, 467-472.

Pansky, A., Goldsmith, M., Koriat, A., & Pearlman-Avnion, S. (2009). Memory accuracy in old age: Cognitive, metacognitive, and neurocognitive determinants. *European Journal of Cognitive Psychology, 21 (2/3)*, 303-329.

Park, D.C. & Hedden, T. (2001). Working memory and aging. In M. Naveh – Benjamin, M. Moscovitch, & H.L. Roediger, III (Eds.), *Perspectives on Human Memory and Cognitive Aging: Essays in Honour of Fergus Craik*, (pp. 148-160). New York: Psychology Press.

Park, D. & Minear, M. (2004). Cognitive Aging: New Directions for old theories. In R.A. Dixon, Lars Bäckman & Lars-Göran Nilsson, *New Frontiers in Cognitive Aging* (19-40), Oxford University Press.

Park, D.C., Puglisi, J.T., Smith, A.D., & Dudley, W.N. (1987), Cue utilization and encoding specificity in picture recognition by older adults. *Journal of Gerontology*, *42*, 423-425. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Park, D.C., Smith, A.D., Dudley, W.N., & Lafronza, V.N. (1989). Effects of age and a divided attention task presented during encoding and retrieval on memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15* (6), 1185-1191.

Park, D.C., Smith, A.D., Lautenschlager, G., Earles, J.L., Frieske, D., Zwahr, M., & Gaines, C.L. (1996). Mediators of long-term memory performance across the lifespan, *Psychology and Aging*, *11*, 621-637. In R.A. Dixon, Lars Bäckman & Lars-Göran Nilsson, *New Frontiers in Cognitive Aging* (19-40) Oxford University Press.

Paterniti, S., Verdier-Taillefer, M., Dufouil, C. & Alpérovitch, A. (2002). Depressive symptome and cognitive decline in elderly people. *British Journal of Psychiatry, 181*, 406-410.

Payne, B.K., Lambert, A.J., & Jacoby, L.L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology, 38*, 384-396.

Pearse, J. & Gudjonsson, G.H. (1996). Police interviewing techniques at two south London police stations. *Psychology, Crime, & 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.

Penrod, S. & Cutler, B. (1995). Witness confidence and witness accuracy: assessing their forensic relation. *Psychology, Public Policy and Law*, *1*, 817-845.

Penrod, S., Loftus, E., & Winkler, J. (1982). The reliability of eyewitness testimony.: A psychological perspective. In N.L. Kerr, & R.M. Bray (Eds.), *The psychology of the courtroom* (pp. 119-168). New York: Academic Press.

Perfect, T.J. (2002). When does eyewitness confidence predict performance? In T.J. Perfect, & B.L. Schwartz (Eds.), *Applied Metacognition* (pp. 95-120). Cambridge: Cambridge University Press.

Perlmutter, M. (1978). What is memory aging the aging of? *Developmental Psychology, 14*, 330-345. In J. Metcalfe & A.P. Shimamura (Eds.), *Metacognition: Knowing about Knowing*, (pp. 227-251).

Perlmutter, M., & Nyquist, L. (1990). Relationships between self-reported physical and mental health and performance across adulthood. *Journal of Gerontology: Psychological Sciences, 45,* 145-155.

Pliske, R.M, & Mutter, S.A. (1996). Age differences in the accuracy of confidence judgments. *Experimental Aging Research, 22, 2*, 199-216.

Polczyk, R. (2000). Wplyw spoleczny a różnice indywidulne: Skale Sugestybilności Interrogatywnej Gudjonssona [Social influence and individual differences:

Gudjonsson Suggestibility Scales]. *Kolokwia Psychologiczne ur 8: Wplyw spoleczny: Badania I praktyka.* Warszawa: Instytut Psychologii PAN.

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.

Polczyk, R., Wesolowska, B., Gabarczyk, A., Minakowska, I. Supska, M., & Bomba, E. (2004). Age Differences in Interrogative Suggestibility: A Comparison between Young and Older Adults. *Applied Cognitive Psychology, 18*, 1097-1107.

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.

Ponds, R.W.H.M., & Jolles, J. (1996). Memory complaints in elderly people: The role of memory abilities, metamemory, depression, and personality. *Educational Gerontology, 22*, 341-357.

Ponds, R.W.H.M., Van Boxtel, M.P.J., & Jolles, J. (2000). Age-related changes in subjective cognitive functioning. *Educational Gerontology, 26*, 67-81.

Poole, D.A., & White, L.T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology, 27*, 975-986.

Poole, D.A., & White, L.T. (1993). Two years later: Effects of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology, 29*, 844-853.

Potter, R. & Brewer, N. (1999). Perceptions of witness behaviour-accuracy relationships held by police, lawyers and jurors. *Psychiatry, Psychology and Law, 6*, 97-103. In R. Bull, T.Valentine, & T. Williamson (Eds). *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions,* (pp. 121-134), Wiley-Blackwell: Chichester.

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

Prescott, K., Milne, R., & Clarke, J. (in press). How Effective is the Enhanced Cognitive Interview when Aiding Retrieval of Older Adults including Memory for Conversation. *Journal of Investigative Psychology and Offender Profiling.*

Prideaux, E. (1919). Suggestion and suggestibility. *British Journal of Psychology, 10*, 228-241.

Prull, M.W., Gabrielli, J.D.E., & Bunge, S.A. (2000). Age-related Changes in Memory: A cognitive neuroscience perspective. In F.I.M. Craik & T.A. Salthouse (2nd Ed.), *The handbook of aging and cognition* (pp. 91-153), Mahwah, NJ: Erlbaum.

Puglisi, J.T., Park, D.C. Smith, A.D., & Dudley, W.N. (1988), Age differences in encoding specificity. *Journals of Gerontology: Psychological Sciences*, *43*, 145-150. In F.I.M. Craik & T.Salthouse (2nd Ed), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Rabbitt, P.M.A., Lowe, C., & Shilling, V. (2001). Frontal tests and models for cognitive ageing. *European Journal of Cognitive Psychology, 13*, 5-28.

Rabinowitz, J.C. (1984). Aging and recognition failure. *Journal of Gerontology*, *39*, 65-71. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Rabinowitz, J.C. (1986). Priming in episodic memory, *Journal of Gerontology*, *41*, 204-213. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Rabinowitz, J.C. & Ackerman, B.P. (1982), General encoding of episodic events by elderly adults. In F.I.M. Craik & T.Salthouse (2nd Ed.), *Handbook of aging and cognition* (51-110). Hillsdale, NJ: Erllbaum.

Rahhal, T.A., Hasher, L., & Colcombe, S.J. (2001). Instructional manipulations and age differences in memory: Now you see them, now you don't. *Psychology and Aging, 16*, 697-706.

Rahhal, T.A., May, C.P., & Hasher, L. (2002). Truth and character: Sources that older adults can remember. *Psychological Science, 13*, 101-105.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F.I.M. Craik & T.A. Salthouse (2nd Ed.), *The handbook of aging and cognition* (pp. 91-153), Mahwah, NJ: Erlbaum.

Reder, L.M. (1987). Strategy selection in questioning answering. *Cognitive Psychology, 19*, 90-137.

Reder, L.M., Wible, C., & Martin, J. (1986). Differential memory changes with age: Exact retrieval versus plausible inference. Journal of Experimental Psychology: *Learning, Memory and Cognition, 12*, 72-81.

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

Reece, C.M., & Cherry, K.E. (2006). Effects of Age and Ability on Self-Reported Memory Functioning and Knowledge of Memory Aging. *Journal of Genetic Psychology, 167(2)*, 221-240.

Register, P.A. & Kihlstrom, J.F. (1988). Hypnosis and interrogative suggestibility. *Personality and Individual Differences, 9*, 549-558.

Rhodes, M.G., & Kelley, C.M. (2005). Executive processes, memory accuracy, and memory monitoring: An aging and individual differences analysis. *Journal of Memory and Language, 52*, 578-594.

Rice, F.P. (1986). *Adult Development and Aging.* Boston: Allyn and Bacon. Chapter 12.

Richardson, G., & Smith, P. (1993). The inter-rater reliability of the Gudjonsson Suggestibility Scale. *Personality and Individual Differences, 14,* 251-253.

Richardson, G., Gudjonsson, G.H., & Kelley, T.P. (1995). Interrogative suggestibility in an adolescent forensic population. *Journal of Adolescence, 18*, 211-216.

Riggs, K.M., Lachman, M.E., & Wingfield, A. (1997). Taking charge of remembering: Locus of control and older adults' memory for speech. *Experimental Aging Research, 23*, 237-256.

Robinson, J., & McGuire, J. (2006). Suggestibility and children with mild learning disabilities: the use of the cognitive interview. *Psychology, Crime & Law, 12*, 537-556. In R. N. Kocsis (Ed.). *Applied Criminal Psychology: A Guide to Forensic behavioural Sciences*.

Robinson, M.D. & Johnson, J.T. (1996). Recall memory, recognition memory, and the eyewitness confidence-accuracy relation. *Journal of Applied Psychology, 81*, 587-594.

Rose, R.A., Bull, R., & Vrij, A. (2005). Non-biased lineup instrauctions do matter – A problem for older witnesses. *Psychology, Crime and Law, 11*, 147-159.

Rothman, M.B., Dunlop, B.D., & Entzel, P. (2000). *Elders, crime, and the criminal justice system. Myth, perceptions, and reality in the 21st Century.* New York: Springer. Series on Life Styles and Issues in Aging.

Ryan, E.B. & Cole, R.L. (1990). Evaluative perceptions of interpersonal communication with elders. In H. Giles, N. Coupland, & J. Wiemann (Eds.), *Communication, health and the elderly* (Fulbright Series No. 8, pp. 172-190). Manchester, England: Manchester University Press.

Ryan, E.B., Meredith, S.D., & Shantz, G.B. (1994). Evaluative perceptions of patronizing speed addressed to institutionalized elders in contrasting conversational contexts. *Canadian Journal on Aging, 13*, 236-248.

Salthouse, T.A. (1996). The processing - speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403-428, in R.A. Dixon, Lars Bäckman & Lars-Göran Nilsson, *New Frontiers in Cognitive Aging* (19-40), Oxford University Press.

Salthouse, T.A. (2009). When does age-related cognitive decline begin? *Neurobiology Of Aging, 30(4)*, 507-514.

Salthouse, T.A. & Babcock, R.L. (1991). Decomposing adult age differences in working memory. *Development Psychology*, *27*, 763-776. In R.A. Dixon, Lars Bäckman & Lars-Göran Nilsson, *New Frontiers in Cognitive Aging* (19-40) Oxford University Press.

Salthouse, T.A., Kausler, D.H., & Saults, J.S. (1990). Age, self-assessed health status, and cognition. *Journal of Gerontology: Psychological Sciences, 45*, 156-160.

Sanders, G.S. (1986). The usefulness of eyewitness research from the perspective of police investigators. Unpublished manuscript, state University of New York at Albany.

Sanders, G.S. (1996). On increasing the usefulness of eyewitness research. *Law and Human Behaviour, 10*, 333-336.

Savage, S. & Milne, R. (2006). Miscarriages of Justice – The role of the Investigative process. In T. Newburn, T. Williamson, & A. Wright. (Eds), *Handbook of criminal investigation*. Cullompton: Willan.

Schacter, D. (1990). Introduction to 'Implicit Memory: multiple perspectives', *Bulletin of the Psychonomic Society*, *28 (4)*, 338-340.

Schacter, D.L., Koutstaal, W., Johnson, M.K., Gross, M.S., & Angell, K.E. (1997). False recollection induced by photographs. *Psychology and Aging, 12*, 203-215.

Schacter, D.L., Koutstaal, W., & Norman, K. (1997). False memories and aging. *Trends in Cognitive Sciences, 1*, 229-236.

Schacter, D.L., Norman, K., & Koutstaal, W. (1998). The cognitive neurosciences of constructive memory. *Annual Review of Psychology, 49*, 289-318.

Schneider, B.A., & Pichora-Fuller, M.K. (2000). Implications of perceptual deterioration for cognitive aging research. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 155-219), Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schooler, J.W. & Loftus, E.F. (1986). Individual differences and experimentation: complementary approaches to interrogative suggestibility. *Social Behaviour, 1*, 105-112.

Schwartz, B.L. (1994). Sources of information in metamemory: Judgments of learning and feeling of knowing. *Psychonomic Bulletin and Review, 1*, 357-375.

Scullin, M.H. & Ceci, S.J. (2001). A suggestibility scale for children. *Personality and Individual Differences, 5*, 203-209.

Searcy, J., Bartlett, J.C., & Memon, A. (2000). Influence of post event narratives, line-up conditions and individual differences on false identification by young and older eyewitnesses, *Legal and Criminological Psychology, 5*, 219-235.

Seeman, T., McAvay, G., Merril, S., Albert, M., & Rodin, J. (1996). Self-efficacy beliefs and change in cognitive performance: MacArthur studies of successful aging. *Psychology and Aging, 11*, 538-551.

Sello, E. (1911). *Die Irrtümer der Strafjustiz und ihre Ursachen [Errors of criminal justice and their causes]* (Vol 1). Berlin: R. v. Decker Verlag.

Semmler, C., Brewer, N., & Wells, G.L. (2004). Effects of postidentification feedback on eyewitness identification and nonidentification confidence. *Journal of Applied Psychology, 89*, 334-346.

Sharps, M.J. & Antonelli, J.R.S. (1997). Visual and semantic support for paired-associates recall in young and older adults. *Journal of Genetic Psychology, 158* (3), 347-356.

Shaw, J.S. III (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied*, *2*, 126-146.

Shaw, J.S. III & McClure, K.A. (1996). Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law and Human Behaviour*, 6, 714-723.

Shaw, J.I. & Skolnick, P. (1994). Sex differences, weapon focus, and eyewitness reliability. *Journal of Social Psychology*, *134*, 413-420.

Shepard, R.N. (1967). Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behaviour, 6*, 156-163.

Sheskin, D. (2007). *Handbook of parametric and nonparametric statistical procedures.* **(**4th ed.) Boca Raton : Chapman & Hall.

Shimamura, A.P. (2000). Toward a cognitive neuroscience of metacognition. *Consciousness and Cognition, 9*, 313-323.

Shing, Y.L., Werkle-Bergner, M., Li, S.C., & Lindenberger, U. (2009). Committing memory errors with high confidence: Older adults do but children don't. *Memory 17, 2*, 169-179.

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.

Skagerberg, E.M. & Wright, D.B. (2009). Susceptibility to Postidentification Feedback is Affected by Source Credibility. *Applied Cognitive Psychology, 23 (4),* 257-267.

Smeets, T., Leppink, J., Jelicic, M., & Merckelbach, H. (2009). Shortened versions of the Gudjonsson Suggestibility Scale meets the standards. *Legal and Criminological Psychology, 14,* 149-155.

Smith, (1930). The fallibility of witness testimony. *The American Journal of Police Science, 1*, 487-495.

Smith, V.L., Kassin, S.M., & Ellsworth, P.C. (1989). Eyewitness accuracy and confidence: Within versus between-subjects correlations. *Journal of Applied Psychology, 74*, 356-359.

Spielberger, C.D. (1969). *The State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Sporer, S.L. (2008). Lessons from the Origins of Eyewitness Testimony Research in Europe. *Applied Cognitive Psychology, 22*, 737-757.

Sporer, S.L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence and accuracy: a meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, *118*, 315-327.

Squire, L.R., Wetzel, C., & Slater, P.C. (1979). Memory Compaint after electroconvulsive therapy: Assessment with a new self-rating instrument. *Biological Psychiatry, 147,* 791-801.

Steblay, N.M. (1992). A meta-analytic review of the weapon-focus effect. *Law and Human Behaviour, 16,* 413-424.

Stein, L., & Memon, A. (2006). Testing the efficacy of the Cognitive Interview in a developing country. *Applied Cognitive Psychology, 20*, 597-605.

Stein, R., Blanchard-Fields, & Hertzog, C. (2002). The effects of age-stereotype priming on memory performance in older adults. *Experimental Aging Research, 28,* 169-181.

Stern, L.W. (Ed.). (1903-1906). *Beiträge zur Psychologie der Aussage* (2 Bände) [Contributions to the psychology of testimony (2 Vols.)]. Leipzig: Barth. (1. Folge 1903-04: 2. Folge 1905-06).

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.

Stern, Y. (2009). Cognitive Reserve. Neuropsychologia, 47, 2015-2028. In Chamorro-Premuzic, T., Furnham, A., & von Stumm, S. (Editors), *Handbook of Individual Differences* (in press).

Stevens, J.P. (2009), (5th Ed.). *Applied Multivariate Statistics for the Social Sciences.* Routledge, Hove.

Stock, J., & Cervone, D. (1990). Proximal goal setting and self-regulatory processes on complex decision making perfpormance. *Cognitive Therapy and Research, 14*, 483-494.

Stokes, L.C. and Pankowski, M.L. (1988). Incidental Learning of aging adults via television. *Adult Education Quarterly*, *38*, 88-99.

Stretch, V. & Wixted, J.T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 1379-1396, in In R. Bull, T.Valentine, & T. Williamson (Eds.). *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions,* (pp. 205-220). Wiley-Blackwell: Chichester.

Stroud, J.N., & Wright, D.B. (1997), Identifying your own age: an own-age bias in line-up identifications. Paper presented to the Society of Applied Research in Memory and Cognition, Toronto, July.

Swets, J.A., Tanner, W.P. Jr., & Birdsall, T.G. (1961). Decision processes in perception. *Psychological Review, 68*, 301-340.

Tabachnick, B.G., & Fidell, L.S. (2007), (5th Edition). *Using Multivariate Statistics.* Pearson International.

Tata, P.R. & Gudjonsson, G.H. (1990). The effects of mood and verbal feedback on interrogative suggestibility. *Personality and Individual Differences, 11*, 1079-1085.

Tichner, A., & Poulton, E. (1975). Watching for people and actions. *Ergonomics, 18,* 35-51.

Tousignant, J.P., Hall, D., & Loftus, E.F. (1986). Discrepancy detection and vulnerability to misleading post-event information. *Memory and Cognition, 14*, 329- 338

Trankell, A. (1958). Was Lars sexually assaulted? A study in the reliability of witnesses and of experts. *The Journal of Abnormal and Social Psychology, 56*, 385-395.

Trippi, J.A. (1973). Special class placements and suggestibility of mentally retarded children. *American Journal of Mental Defficiency, 79*, 200-222.

Tucker-Drob, E.M., & Salthouse, T.A. (in press). Individual differences in cognitive aging. In Chamorro-Premuzic, T., Furnham, A., & von Stumm, S. (Eds.), *Handbook of Individual Differences.*

Tulving, E. (1991). Concepts of human memory. In L.R. Squire, N.M. Weinberger, G. Lynch, and J.L. McGaugh (Eds.), *Memory: Organisation and locus of change*, New York: Oxford University Press.

Tulving, E. & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory, *Psychological Review, 80*: 353-370.

Tun, P.A. (1998). Fast noisy speech: Age differences in processing rapid speech with background noise. *Psychology and Aging, 17*, 453-467. In F.I.M. Craik &

T.Salthouse (3rd Ed.), *The Handbook of aging and cognition* (251-310). Psychology Press, New York. NY.

Tun, P.A., McCoy, S., Cox, C., & Wingfield, A. (2006 April). *Divided attention costs for memory for speech depend on both age and hearing ability.* Cognitive Aging Conference, Atlanta, GA. In F.I.M. Craik & T.Salthouse (3rd Ed.), *The Handbook of aging and cognition* (251-310). Psychology Press, New York. NY.

Turtle, J.W. & Yuille, J.C. (1994). Lost but not forgotten details: Repeated eyewitness recall leads to reminiscence but not hypermnesia. *Journal of Applied Psychology*, *79*, 260 – 271.

Undeutsch, U. (1982). Statement reality analysis. In A. Trankell (Ed.), *Reconstructing the past: The role of psycologists in criminal trials* (pp. 27-56). Denver: Kluwer.

Vakil, E., Weise, M., & Enbar, S. (1997). Direct and indirect memory measures of temporal order: Younger versus older adults. *International Journal of Aging and Human Development*, *45* (3), 195-206.

Van Bergen, S., Brands, I. Jelicic, M., & Merckelbach, H. (2010). Assessing trait memory distrust: Psychometric properties of the Squire Subjective Memory Questionnaire. Legal And Criminological Psychology, 15 (2), 373-384.

Van Bergen, S., Jelicic, M., & Merckelbach, H. (2009). Are subjective memory problems related to suggestibility, compliance, false memories, and objective memory performance? *American Journal of Psychology, 122(2),* 249-257.

Van Hooren, S.A.H., Valentijn, A.M., Bosma, H., Ponds, R.W.H.M., Van Boxtel, M.P.J., & Jolles, J. (2007). Cognitive Functioning in Healthy Older Adults Aged

64-81: A Cohort Study into the Effects of Age, Sex, and Education. *Aging, Neuropsychology, and Cognition, 14*, 40-54.

Visher, C.A. (1987). “Incapacitation and crime control: Does a “Lock em up” strategy reduce crime?” *Justice Quarterly 4*, 513-543.

Warren, A., Hulse-Trotter, K., & Tubbs, E.C. (1991). Inducing resistance to suggestibility in children. *Law and Human Behaviour, 15*, 273-285.

Watson, D., Clark, L.A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 47*, 1063–1070.

Weber, N., & Brewer, N. (2003). The effect of judgment type and confidence scale on confidence-accuracy calibration in face recognition. *Journal of Applied Psychology, 88*, 490-499.

Weber, N., & Brewer, N. (2004). Confidence-accuracy calibration in absolute and relative face recognition judgments. Journal of Experimental Psychology: *Applied 10*, 156-172.

Weber, N., & Brewer, N. (2006). Positive versus negative face recognition decisions: Confidence, accuracy and response latency. *Applied Cognitive Psychology, 20*, 17-31.

Wechsler, D. (1945). *The Wechsler Memory Scale*. San Antonio: The Psychological Corporation.

Wechsler, D. (2001). *Wechsler Test of Adult Reading* (Harcourt Assessment, San Antonio).

Weingardt, K.R., Leonesio, R.J., & Loftus, E.F. (1994). Viewing eyewitness research from a metacognitive perspective. In J. Metcalfe and A.P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 155-184). Cambridge, MA: MIT Press.

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, G.L. (1978). Applied Eyewitness- testimony Research: System Variables and Estimator Variables, *Journal of Personality and Social Psychology*, *36 (12)*, 1546-1557.

Wells, G.L. & Bradfield, A.L. (1998). "Good you identified the suspect": Feedback to eyewitness distorts their reports to the witnessing experience. *Journal of Applied Psychology, 83*, 360-376.

Wells, G.L., & Bradfield, A.L. (1999). Distortions in eyewitnesses' recollections: Can the postidentification-feedback effect be moderated? *Psychological Science, 10*, 138-144.

Wells, G.L., & Lieppe, M.R. (1981). How do triers of fact infer the accuracy of eyewitness identification? Using memory for peripheral detail can be misleading. *Journal of Applied Psychology, 66*, 682-687.

Wells, G.L., & Lindsay, R.C.L. (1985). Methodological notes on the accuracy-confidence relation in eyewitness identifications. *Journal of Applied Psychology, 70*, 413-419.

Wells, G.L., & Loftus, E.F. (Eds.), (1984). *Eyewitness testimony: Psychological Perspectives.* Cambridge, England: Cambridge University Press.

Wells, G.L., & Murray, D.M. (1984). Eyewitness Confidence. In G.L. Wells & E.F. Loftus (Eds.), *Eyewitness Testimony: Psychological Perspectives* (155-170) Cambridge: Cambridge University Press.

Wells, G.L., Small, M., Penrod, S., Malpass, R.S., Fulero, S.M., & Brimacombe, C.A.E. (1998). Eyewitness identification procedures: recommendations for line-ups and photo spreads. *Law and Human Behaviour*, *22*, 603 -647.

West, R.L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120,* 272-292.

West, R.L., Bagwell, D.K., & King, A. (2006). The impact of group size on memory and memory beliefs: A methodological report. *Experimental Aging Research 32(3)*, 297–316.

West, R.L., & Berry, J.M. (1994). Age declines in memory self – efficacy: General or limited to particular tasks and measures? In J.D. Simon (Ed), *Interdisisciplinary handbook of adult lifespan learning* (pp 426–446). Westport, CT: Greenwood Press.

West, R.L., Dennehy-Basile, D., & Norris, M.P. (1996). Memory self-evaluation: The effects of age and experience. *Aging, Neuropsychology, and Cognition, 3*, 67-83.

West, R.L., Thorn, R.M., & Bagwell, D.K. (2003). Memory performance and beliefs as a function of goal setting and aging. *Psychology and Aging*, *18*, 111–125.

West, R.L., Welch, D.C., & Thorn, R.M. (2001). Effects of goal setting and feedback on memory performance and beliefs among older and younger adults. *Psychology and Aging*, *16*, 240–250.

White, N., & Cunningham, W.R. (1982). What is the evidence for retrieval problems in the elderly? *Experimental Aging Research, 8*, 169-171.

Whiting, W.L. & Smith, A.D. (1997). Differential age-related processing limitations in recall and recognition tasks. *Psychology and Aging, 12*, 216-224.

Wilcock, R.A. (2010). The Aging Eyewitness. In J. Adler & J. Gray (2nd Ed.), *Forensic Psychology: Concepts, Debates and Practice* (pages 162-180), Willan.

Wilcock, R.A., Bull, R., & Vrij, A. (2005). Aiding the performance of older eyewitnesses: Enhanced non-biased lineup instructions and presentation. *Psychiatry, Psychology and Law, 12*, 129-140.

Wilcock, R.A., Bull, R., & Vrij, A. (2007). Are older witnesses always poorer witnesses? Identification accuracy, context reinstatement, own age bias. *Psychology, Crime, and Law, 13,* 305-316.

Wilcock, R., Bull, R., & Milne, R. (2008). *Witness identification in criminal cases: Psychology and Practice.* Oxford University Press.

Williamson, T. (2007). Psychology and criminal investigation. In T. Newburn, T. Williamson, & A. Wright (Eds.), *Handbook of criminal investigation* (pp. 68-91). Devon: Willan Publishing.

Wilson, R.S., Beckett, L.A., Barnes, L.L., Schneider, J.A., Bach, J., Evans, D.A., & Bennettt, D.A., (2002). Individual differences in rates of change in cognitive abilities of older persons. *Psychology and Aging, 17*, 179-193.

Wingfield, A., Tun, P.A., & McCoy, S.L. (2005). Hearing loss in older adulthood: What it is and how it interacts with cognitive performance. *Current Directions in Psychological Science, 14*, 144-148. In F.I.M. Craik & T. Salthouse (3rd Ed.), *The Handbook of aging and cognition* (251-310). Psychology Press, New York. NY.

Wixted, J.T. & Stretch, V. (2005). In defence of the signal detection interpretation of remember / know judgments. *Psychonomic Bulletin & Review, 11*, 616-641.

Woodruff-Pak, D.S. (1997). *The neuropsychology of aging.* Malden, MA: Blackwell.

Wright, D.B. & Davies, G.M. (1999). Eyewitness Testimony. In F.T. Durso, R.S. Nickerson, R.W. Schvaneveldt, S.T. Dumnis, D.S. Lindsay & T.H. Chi (Eds.), *Handbook of Applied Cognition*, 789-818, John Wiley: New York.

Wright, D.B. & Stroud, J.N. (2002). Age differences in lineup identification accuracy: People are better with their own age. *Law And Human Behavior 26 (6),* 641-654.

Wright, A.M., & Holliday, R. E. (2007). Enhancing the Recall of Young, Young – Old and Old – old Adults with the Cognitive Interview and a Modified Version of the Cognitive Interview. *Applied Cognitive Psychology, 21*, 19-43.

Wright, D.B., Memon, A., Skagerberg, E.M., & Gabbert, F. (2009). When Eyewitnesses Talk. *Current Directions in Psychological Science*, *18 (3)*, 174-178.

Wright, D.B. & Stroud, J.N. (1998). Memory quality and misinformation for peripheral and central objects. *Legal and Criminological Psychology*, *3*, 273-286.

Yarmey, A.D. (1982). Eyewitness identification and stereotypes of criminals. In A. Trankell (Ed.) *Reconstructing the past. The role of psychologists in criminal trials.* (pp. 205-225) Deventer. The Netherlands: Kluwer.

Yarmey, A.D. (1993), Adult age and gender differences in eyewitness recall in field settings. *Journal of Applied Social Psychology*, *23*, 1921-1932.

Yarmey, A. D., Jones, H.T., & Rashid, S. (1984). Eyewitness memory of elderly and young adults. In D.J. Muller, D.E. Blackman, & A.J. Chapman (Eds.), *Psychology and Law*, (pp. 215-228), Chichester: Wiley.

Yarmey, A. D., & Kent, J. (1980). Eyewitness identification by elderly and young adults. *Law and Human Behaviour, 4*, 359-371.

Yarmey, A. D., & Yarmey, M.J. (1997). Eyewitness recall and duration estimates in field settings. *Journal of Applied Social Psychology*, *27*, 330-371.

Young, K., Powell, M.B., & Dudgeon, P. (2003). Individual differences in children's suggestibility: a comparison between intellectually disabled and mainstream samples. *Personality and Individual Differences, 35*, 31-49.

Yuille, J.C., & Cutshall, J.L. (1986), A case study of eyewitness memory of a crime. *Journal of Applied Psychology*, *71*, 291-301.

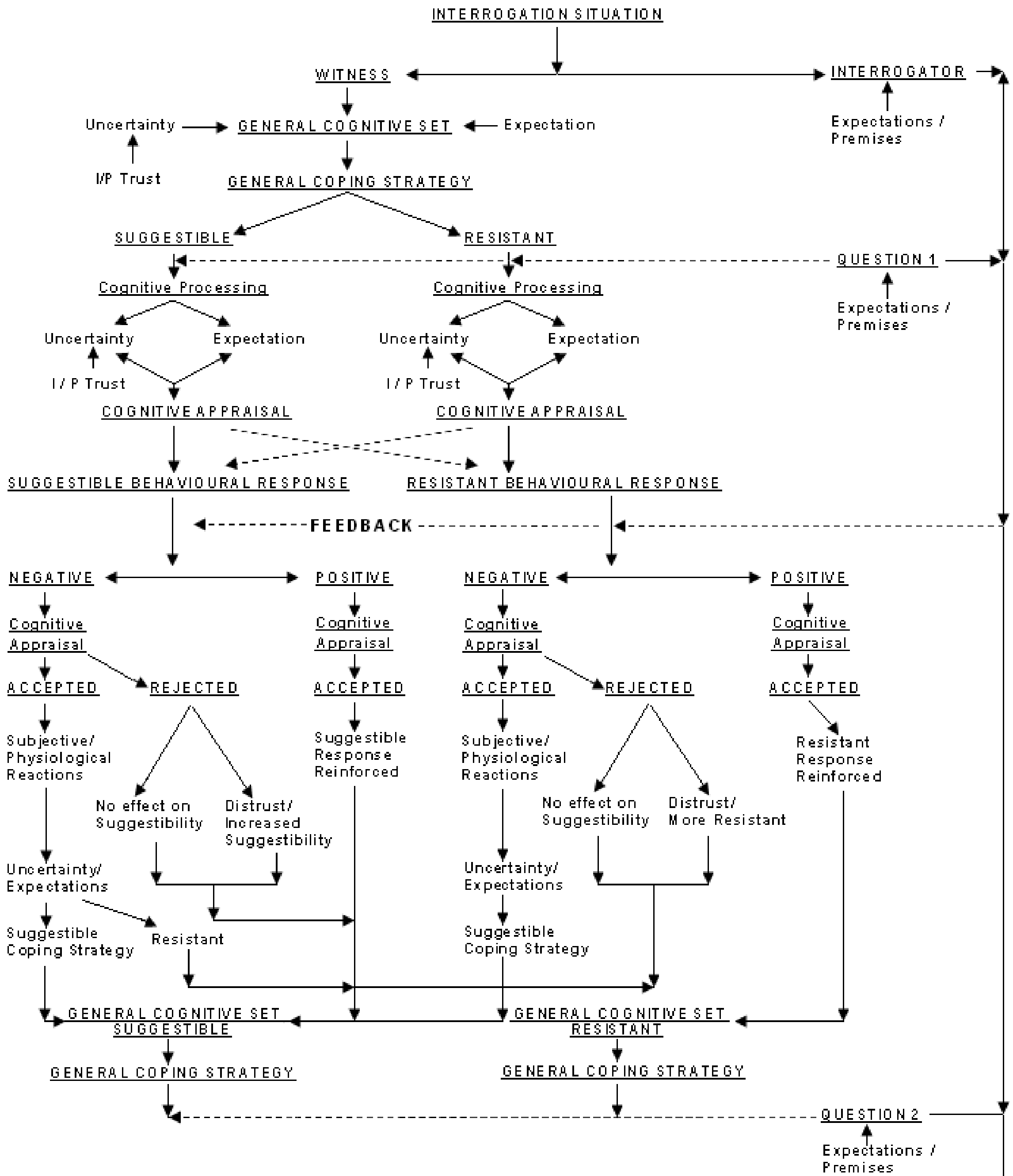
Zacks, R.T. & Hasher, L (1988). Capacity theory and the processing of inferences. In L.L. Light & D.M. Burke (Eds.), *Language, memory and aging* (154-170). New York: Cambridge University Press.

Zacks, R.T., Hasher, L., & Li, K.Z.H. (2000). Human Memory. In F.I.M. Craik & T.A. Salthouse (2nd ed.), *The handbook of aging and cognition* (pp. 293-357). Mahwah, NJ: Erlbaum.

Zander, M. & Henderson, P. (1993). Crown court study. *Royal Commission on Criminal Justice*. *Research Study*, *19*. HMSO: London.

Zelinski, E.M., Gilewski, M.J., & Anthony-Bergstone, C.R. (1990). Memory Functioning Questionnaire: Concurrent validity with memory performance and self-reported memory failures. *Psychology and Aging, 5*, 388-399.

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 maximum score 40 (range 0-40) Total confabulation 1
30 minute delay	-
Interviewees asked again for free recall	Delayed recall score maximum score 40 (range 0-40) Total confabulation 2
Interviewees asked 20 questions about the narrative (15 leading questions)	**Yield 1** – number of leading items accepted by interviewee 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 maximum score 20 (range 0-20) **Yield 2** – number of leading items accepted by interviewee maximum score 15 (range 0-15) Total suggestibility = Yield 1 + Shift 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 maximum score 40 (range 0-40) Total confabulation 1
50 minute delay	-
Interviewees asked again for free recall	Delayed recall score maximum score 40 (range 0-40) 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 maximum score 20 (range 0-20)

Appendix D - Participant Details Questionnaire

NAME:……………………………………

ADDRESS………………………………………………………………
……………………………………………………………………………
….………………………..………………………………………………

TELEPHONE No………………………………………………………

EMAIL ADDRESS (if you have one)………………………………………

Sex……………………….

Age……………………….

Date of Birth…………………………………

Do you wear glasses or contact lenses? Yes / No

Do you experience colour blindness? Yes / No

If yes, are there certain colours which you find it difficult to distinguish?
(please specify)……………………………………..

Is English your first language? Yes / No

What is your occupation / pre – retirement
occupation……………………………………………………………

How many years of education do you have?…………………………….
e.g. left school at 16 = 12 years of education; educated to degree level = 17

Please circle your response to the following questions, with 1 being excellent and 5 being very poor

How would you describe your eyesight?	1 2 3 4 5
How would you describe your hearing?	1 2 3 4 5
How would you describe your physical health?	1 2 3 4 5
How would you describe your mental health?	1 2 3 4 5
How would you describe your everyday memory?	1 2 3 4 5

Please provide any other information you feel the experimenter should be aware of before you start the testing (e.g. health status not covered above, the best times to contact you)

………………………………………………………………………………
………………………………………………………………………………
………………………………………………………………………………
………………………………………………………………………………

Appendix E - Directions film summary

	SCORING (Units of Info.)	**SPEECH**
PERIPHERAL **Panorama of Todd's Bar area:-** Barony Hall Parked cars Taxi pulls away (A) Irn Bru logo Car pulls away (A) Pedestrian beside parked cars Woman Wearing – black coat (2) black trousers (2) black bag over shoulder (2) Bushes Trees Wooden Benches X 2 Outdoor Lights X 6 Buildings – Flats Dustbin	19 - SURROUNDINGS	

<u>PERIPHERAL</u> **Todd's Bar area:-** Signs in window X 2 Pentagon shape Orange Main windows X 2 Doorways X 2 Doors X 4 Grey Single glass panel Table Metal Chairs X 4 Seat & Back – wicker Arms & Legs – metal Sponge on Table Green	16 - SURROUNDINGS	
<u>PERIPHERAL</u> Waiter Aged 35 - 40 Ethnicity = White Scottish Accent Dark hair – Short Wearing - Black trousers (2) - White shirt (2) Slim Build Clean shaven	12 - PERSON	

PERIPHERAL Waiter approaches table from screen right (A) Waiter has notebook - puts in back right trouser pocket (A) Waiter straightens 2 of 4 chairs (A) Waiter cleans table with sponge (A) - using right hand (A) Waiter enters bar (A) through left door	 7 - PERSON	
CENTRAL Girls X 2 enter screen left (A) Girl 1 Aged 25 - 30 Ethnicity = White Dark hair Collar length Wearing - Black coat - (2) - Dark Jeans – (2) - with turnups - Black shoes – (2) Slim Build Girl 1 Carries hand bag & shopping bag	 20 - PERSON	

– over right shoulder (A) - Bag is black Shopping bag is large - yellow Speaks with a North American Accent		
<u>CENTRAL</u> Girl 2 Aged 20 – 25 Ethnicity = white Dark Hair Collar Length Silver Streak Wearing – Brown coat - purple scarf - brown trousers Slim build Girl 2 - Carries shopping bag over right shoulder (A) Bag – cream & green with brown trim Girl 2 – Carries Lakeland bag in left hand (A) Scottish Accent	 19 - PERSON	

CENTRAL Girl 1 speaks to Girl 2	1 - PERSON 2 - DIALOGUE	Girl 1 "Shall we sit down and have a coffee" Girl 2 "Yeah why not"
CENTRAL Girl 1 pulls out seat with left hand - sits down (A) Girl 1 takes both bags from her shoulder (A) Girl 1 puts bags on ground to her right (A)	3 - PERSON	
CENTRAL Girl 2 takes green & cream bag from her shoulder (A) Places in her left hand (A) Girl 2 sits in seat on right of table (A) - puts Lakeland bag on the ground to her left (A) - puts green & cream bag on her knee (A) - puts green & cream bag on ground to her left (A) Girl 2 crosses left over right leg (A)	7 - PERSON	

CENTRAL Girl 1 speaks to Girl 2	 1 - PERSON 2 - DIALOGUE	 Girl 1 “My feet are killing me” Girl 2 “Too much walking – I know. Too many nice things to choose from”
PERIPHERAL Waiter approaches table from screen right (A) Waiter has notepad in left hand - pen in right hand Waiter walks to left hand side of the table (A)	 4 - PERSON	
CENTRAL Waiter speaks to Girl 1 & 2	 1 - PERSON 3 - DIALOGUE	 Waiter “Hiya – what would you like?” Girl 1 “ I’ll have a white coffee please” Waiter (repeats) “White coffee” Girl 2 “And a latte thanks” Waiter “White coffee and a latte – is that all?” Girl 1 & 2 “Yeah”
PERIPHERAL Waiter leaves table to screen right (A)	5 - PERSON	

- note pad & pen in right hand Opens left hand door (A) - with left hand Waiter enters bar (A)		
CENTRAL Girl 2 taps table with both hands (A) Girl 2 speaks to Girl 1	2 - PERSON 3 - DIALOGUE	Girl 2 "So tell me – shoes" Girl 1 "Finally – I got the perfect shoes to go with that dress. I'm so happy – such a relief" Girl 2 "The problem is – there's too many things to choose from"
CENTRAL **Passerby** Male Aged 25 - 30 Ethnicity = White Brown Hair - short Beard & moustache (2) Wearing - Brown jacket (2) - Blue jeans (2) - Back pack - Over left shoulder	16 - PERSON	

Medium Build Irish Accent		
<u>CENTRAL</u> Passerby enters screen right (A) - Looks off screen to left (A) - Turns round (A) - Approaches table from screen left (A) Passerby carries street map in both hands (A) - looks at street map (A)	6 - PERSON	
<u>CENTRAL</u> Passerby speaks to Girl 1 & Girl 2 Passerby points to street map(A) Passerby places street map on the table(A) Girl 1 looks at street map(A) Girl 1 points to street map(A)	5 - PERSON 3 - DIALOGUE	Passerby “Sorry, sorry to interrupt. You couldn’t tell me where the ticket office is? I think its somewhere in the Merchant City” Girl 1 & Girl 2 “Yeah” Girl 1 “Do you have a map there?” Girl 1 “OK so we are here”
<u>CENTRAL</u> Passerby rubs beard (A) - with left hand	9 - PERSON	Passerby “OK”

Passerby leans right hand on table(A) Girl 1 points to screen left(A) Passerby looks to screen left (A) Girl 2 looks at street map(A) Passerby exits screen left(A) Girl 1 waves (A) - with right hand	8 - DIALOGUE	Girl 1 “And the Merchant City is just down here” Passerby “OK” Girl 1 “So what you’re gonna wanna do is go down to the main road” Girl 1 “And then you’ll see it right at the corner of Ingram and Candleriggs” Passerby (repeats) “Ingram and Candleriggs” Girl 2 “Are you sure its Candleriggs – is that what it says? – It’s definitely on a corner – maybe its Albion Street – is that right?” Girl 1 “It could be” Girl 2 “I know the map’s not very clear” Passerby “OK There both quite close together – I’ll try both then. Excellent thanks very much” Girl 1 & 2 “Bye”
CENTRAL Girl 1 speaks to Girl 2(A) Waiter enters screen right(A) - walks to right side of table	8 - PERSON 3 - DIALOGUE	Girl 1 “That necklace you bought is absolutely beautiful” Girl 2 “Expensive though”

Waiter carries a white tray (A) - in both hands - 2 white mugs Waiter places each mug down on the table(A) Waiter speaks to Girl 1 & Girl 2		Girl 1 & 2 “Hiya” Waiter “There you go – white coffee and a latte – ok”
PERIPHERAL Waiter exits screen right (A) Waiter holds tray in right hand (A)	2 - PERSON	

TOTALS	
PERSON	**128**
DIALOGUE	**24**
SURROUNDINGS	**35**
GRAND TOTAL	**187**

Appendix F (1) - Action Questions (Cued-Recall)

Q1 What type of vehicle was seen first pulling away?
A Black taxi

Q8 Where did the waiter put his notebook?
A Back right trouser pocket

Q9 How many chairs did the waiter re-position at the table?
A Two

Q10 What did the waiter use to clean the table?
A A green sponge

Q13 How did girl 1 carry her bags?
A Over her right shoulder

Q15 After girl 2 had sat down where did she place her bags?
A On ground to her left

Q20 On leaving the table - Where did the waiter put his notepad & pen?
A In his right hand

Q21 What did girl 2 tap on the table?
A Both her hands

Q22 From what direction did the passer-by enter the scene?
A Screen right

Q23 What was the passerby holding?
A A street map

Q24 Why did the passerby stop at the table?
A To ask for directions

Q30 How did the waiter carry the drinks tray?
A In both hands

Action Questions (4 – Alternative Forced Choice)

Q1 What type of vehicle was seen first pulling away?

a) private car
b) motor bicycle
c) black taxi
d) bus

Q8 Where did the waiter put his notebook?

a) shirt pocket
b) back right trouser pocket
c) back left trouser pocket
d) on a chair

Q9 How many chairs did the waiter re-position at the table?

a) one
b) two
c) three
d) four

Q10 What did the waiter use to clean the table?

a) shirt sleeve
b) cloth
c) scrubbing brush
d) green sponge

Q13 How did girl 1 carry her bags?

a) in her left hand
b) over left shoulder
c) over right shoulder
d) in right hand

Q15 After girl 2 had sat down where did she place her bags?

a) on ground to her left
b) on her knee
c) on another seat
d) on ground to her right

Q20 On leaving the table - Where did the waiter put his notepad & pen?

- **a) in his right hand**
- b) in his left hand
- c) in his front right trouser pocket
- d) in his shirt pocket

Q21 What did girl 2 tap on the table?

- a) Menu
- b) Pencil
- **c) Both her hands**
- d) Newspaper

Q22 From what direction did the passer-by enter the scene?

- a) Screen left
- **b) Screen right**
- c) In front of the table
- d) Other

Q23 What was the passerby holding?

- **a) A street map**
- b) A cigarette
- c) Copy of the Big Issue
- d) A computer case

Q24 Why did the passerby stop at the table?

- a) To ask for a date
- b) To ask for a light
- c) To sell the Big Issue
- **d) To ask for directions**

Q30 How did the waiter carry the drinks tray?

- **a) in both hands**
- b) in his right hand
- c) in his right hand – over his shoulder
- d) in left hand – over his shoulder

Appendix F (2) - Description Questions (Cued-Recall)

Q2 What was the passing taxi advertising?
A Irn Bru

Q3 How many pedestrians stood beside the parked cars?
A One

Q4 Where did the scene take place?
A Todd’s Bar

Q5 What was the background colour of the sign?
A orange

Q6 What material was the table made of?
A Metal

Q7 What colour was the waiter’s hair?
A Dark

Q11 How many people first approached the bar?
A 2 Girls

Q14 What colour was the shopping bag of Girl 1?
A Yellow

Q16 What accent did girl 1 have?
A North American

Q17 Did girl 2 have any distinguishing features?
A Silver streak to front of hair

Q28 Did the passerby have any distinguishing features?
A Beard & moustache

Q29 What colour was the passerby’s jacket?
A Brown

Description Questions (4 – Alternative Forced Choice)

Q2 What was the passing taxi advertising?
- a) Tennents Lager
- b) Guinness
- c) Volvic Mineral Water
- **d) Irn Bru**

Q3 How many pedestrians stood beside the parked cars?
- a) zero
- **b) one**
- c) two
- d) three

Q4 Where did the scene take place?
- a) Tam's Bar
- b) Toad's Bar
- **c) Todd's Bar**
- d) Top Bar

Q5 What was the background colour of the sign?
- **a) orange**
- b) yellow
- c) green
- d) red

Q6 What material was the table made of?
- a) wooden
- **b) metal**
- c) glass
- d) other

Q7 What colour was the waiter's hair?
- a) blonde
- b) brown
- **c) dark**
- d) dark with blonde streaks

Q11 How many people first approached the bar?

a) 2 girls
b) 1 girl
c) 2 men
d) 1 girl & 1 man

Q14 What colour was the shopping bag of Girl 1?

a) Black
b) Red
c) Yellow
d) Green

Q16 What accent did girl 1 have?

a) English
b) Irish
c) Australian
d) North American

Q17 Did girl 2 have any distinguishing features?

a) silver streak to front of hair
b) wore glasses
c) tattoo on back of hand
d) nose pierced

Q28 What colour was the passerby's jacket?

a) black
b) blue
c) brown
d) grey

Q29 What was girl 1's shopping purchase?

a) dress
b) shoes
c) necklace
d) bag

Appendix F (3) - Dialogue Questions (Cued-Recall)

Q12 Why did they decide to stop at the bar?
A To have coffee

Q18 How did the waiter greet the girls?
A Hiya – what would you like?

Q19 What did girl 2 order?
A Latte

Q25 Where did the passerby want to go?
A The ticket Office

Q26 Where did girl 1 say it was located?
A Ingram & Candleriggs

Q27 What was girl 1's shopping purchase?
A Necklace

Dialogue Questions (4 – Alternative Forced Choice)

Q12 Why did they decide to stop at the bar?
- a) to have lunch
- b) to have an alcoholic drink
- **c) to have coffee**
- d) other

Q18 How did the waiter greet the girls?
- a) Hello – what can I get you?
- b) Would you like to order?
- c) Good afternoon – what can I get you?
- **d) Hiya – what would you like?**

Q19 What did girl 2 order?

a) Gin & tonic
b) latte
c) cappuccino
d) a chicken salad

Q25 Where did the passerby want to go?

a) Theatre Royal
b) The Ticket Office
c) Cineworld
d) The Pavillion Theatre

Q26 Where did girl 1 say it was located?

a) Ingram & Candleriggs
b) Albion Street & Candleriggs
c) Argyle Street
d) Sauchiehall Street

Q27 Did passerby have any distinguishing features?

a) curly hair
b) beard & moustache
c) wore glasses
d) tattoo on back of hand

Appendix G (1) - Memory Questionnaire - Condition 1

In this questionnaire we would like you to tell us what you can remember about the film that you saw earlier. You will be asked a number of questions about the people, places, dialogue and events depicted within the film.

Please give your answer in the space provided.

We are particularly interested in the accuracy of your memory account therefore
IF YOU ARE NOT SURE OF THE ANSWER TO A PARTICULAR QUESTION, YOU MAY CHOOSE NOT TO GIVE A RESPONSE.

Q1 What type of vehicle was seen first pulling away?
A ……………………..

Q2 What was the passing taxi advertising?
A ………………………….

Q3 How many pedestrians stood beside the parked cars?
A ……………………….

Q4 Where did the scene take place?
A ……………………………..

Q5 What was the background colour of the sign?
A ……………………………

Q6 What material was the table made of?
A ……………………………..

Q7 What colour was the waiter's hair?
A ……………………………..

Q8 Where did the waiter put his notebook?
A ………………………………..

Q9 How many chairs did the waiter re-position at the table?
A …………………………..

Q10 What did the waiter use to clean the table?
A ……………………………………………

Q11 How many people first approached the bar?
A ……………………………………..

Q12 Why did they decide to stop at the bar?
A ………………………………….

Q13 How did girl 1 carry her bags?
A …………………………….

Q14 What colour was the shopping bag of Girl 1?
A ………………………………….

Q15 After girl 2 had sat down where did she place her bags?
A ………………………………….

Q16 What accent did girl 1 have?
A …………………………………

Q17 Did girl 2 have any distinguishing features?
A ………………………………………

Q18 How did the waiter greet the girls?
A …………………………..

Q19 What did girl 2 order?
A ………………………..

Q20 On leaving the table - Where did the waiter put his notepad & pen?
A ……………………………….

Q21 What did girl 2 tap on the table?
A ……………………………….

Q22 From what direction did the passer-by enter the scene?
A …………………………………

Q23 What was the passerby holding?
A ……………………………..

Q24 Why did the passerby stop at the table?
A ………………………………

Q25 Where did the passerby want to go?
A ……………………………….

Q26 Where did girl 1 say it was located?
A ………………………………

Q27 Did the passerby have any distinguishing features?
A ………………………………….

Q28 What colour was the passerby's jacket?
A ……………………………………..

Q29 What was girl 1's shopping purchase?
A ………………………………..

Q30 How did the waiter carry the drinks tray?
A ………………………………….

Appendix G (2) - Memory Questionnaire - Condition 2

In this questionnaire we would like you to tell us what you can remember about the film that you saw earlier. You will be asked a number of questions about the people, places, dialogue and events depicted within the film.

Please give your answer in the space provided.

We are particularly interested in the accuracy of your memory account.

One point will be awarded for each correct answer however one point will be deducted for each incorrect answer.

YOU MAY AVOID THE POINTS SYSTEM IF YOU CHOOSE NOT TO ANSWER A PARTICULAR QUESTION. YOU WILL NOT BE PENALIZED OR REWARDED FOR OMITTED RESPONSES.

Q1 What type of vehicle was seen first pulling away?
A ……………………..

Q2 What was the passing taxi advertising?
A ………………………….

Q3 How many pedestrians stood beside the parked cars?
A ……………………….

Q4 Where did the scene take place?
A ……………………………..

Q5 What was the background colour of the sign?
A ……………………………

Q6 What material was the table made of?
A ……………………………..

Q7 What colour was the waiter's hair?
A ……………………………..

Q8 Where did the waiter put his notebook?
A ………………………………..

Q9 How many chairs did the waiter re-position at the table?
A …………………………..

Q10 What did the waiter use to clean the table?
A ………………………………………

Q11 How many people first approached the bar?
A ……………………………………..

Q12 Why did they decide to stop at the bar?
A ………………………………….

Q13 How did girl 1 carry her bags?
A …………………………….

Q14 What colour was the shopping bag of Girl 1?
A ………………………………….

Q15 After girl 2 had sat down where did she place her bags?
A ………………………………….

Q16 What accent did girl 1 have?
A …………………………………

Q17 Did girl 2 have any distinguishing features?
A ……………………………………….

Q18 How did the waiter greet the girls?
A …………………………..

Q19 What did girl 2 order?
A ………………………..

Q20 On leaving the table - Where did the waiter put his notepad & pen?
A ……………………………….

Q21 What did girl 2 tap on the table?
A ……………………………….

Q22 From what direction did the passer-by enter the scene?
A ……………………………………

Q23 What was the passerby holding?
A ……………………………..

Q24 Why did the passerby stop at the table?
A ………………………………

Q25 Where did the passerby want to go?
A ……………………………….

Q26 Where did girl 1 say it was located?
A ………………………………

Q27 Did the passerby have any distinguishing features?
A ………………………………….

Q28 What colour was the passerby's jacket?
A ……………………………………..

Q29 What was girl 1's shopping purchase?
A ………………………………..

Q30 How did the waiter carry the drinks tray?
A ………………………………….

Appendix G (3) - Memory Questionnaire - Condition 3

In this questionnaire we would like you to tell us what you can remember about the film that you saw earlier. You will be asked a number of questions about the people, places, dialogue and events depicted within the film.

Please give your answer in the space provided.

We are particularly interested in the accuracy of your memory account.

Five points will be awarded for each correct answer however one point will be deducted for each incorrect answer. An additional monetary prize payable to charity, of £25.00 will be awarded for the most accurate performance.

Nominated Charity (please specify)………………………………………………………

YOU MAY AVOID THE POINTS SYSTEM IF YOU CHOOSE NOT TO ANSWER A PARTICULAR QUESTION. YOU WILL NOT BE PENALIZED OR REWARDED FOR OMITTED RESPONSES.

Q1 What type of vehicle was seen first pulling away?
A …………………..

Q2 What was the passing taxi advertising?
A ………………………….

Q3 How many pedestrians stood beside the parked cars?
A ……………………….

Q4 Where did the scene take place?
A ……………………………..

Q5 What was the background colour of the sign?
A ……………………………

Q6 What material was the table made of?
A ……………………………..

Q7 What colour was the waiter's hair?
A ……………………………..

Q8 Where did the waiter put his notebook?
A ………………………………..

Q9 How many chairs did the waiter re-position at the table?
A …………………………..

Q10 What did the waiter use to clean the table?
A ………………………………………

Q11 How many people first approached the bar?
A ……………………………………..

Q12 Why did they decide to stop at the bar?
A ………………………………….

Q13 How did girl 1 carry her bags?
A …………………………….

Q14 What colour was the shopping bag of Girl 1?
A ………………………………….

Q15 After girl 2 had sat down where did she place her bags?
A ………………………………….

Q16 What accent did girl 1 have?
A …………………………………

Q17 Did girl 2 have any distinguishing features?
A ……………………………………….

Q18 How did the waiter greet the girls?
A …………………………..

Q19 What did girl 2 order?
A ………………………..

Q20 On leaving the table - Where did the waiter put his notepad & pen?
A ……………………………….

Q21 What did girl 2 tap on the table?
A ……………………………….

Q22 From what direction did the passer-by enter the scene?
A ……………………………………

Q23 What was the passerby holding?
A ……………………………..

Q24 Why did the passerby stop at the table?
A ………………………………

Q25 Where did the passerby want to go?
A ……………………………….

Q26 Where did girl 1 say it was located?
A ………………………………

Q27 Did the passerby have any distinguishing features?
A ………………………………….

Q28 What colour was the passerby's jacket?
A ……………………………………..

Q29 What was girl 1's shopping purchase?
A ………………………………..

Q30 How did the waiter carry the drinks tray?
A ………………………………….

Appendix G (4) - Memory Questionnaire - Condition 4

In this questionnaire we would like you to tell us what you can remember about the film that you saw earlier. You will be asked a number of questions about the people, places, dialogue and events depicted within the film. Please take your time and answer each of the questions to the best of your ability.

Each question is followed by four choices. Draw a circle around the letter corresponding to your choice. Mark **only one** letter for each question.

We are particularly interested in the accuracy of your memory account therefore IF YOU ARE NOT SURE OF THE ANSWER TO A PARTICULAR QUESTION YOU MAY CHOOSE NOT TO GIVE A RESPONSE.

Q1 What type of vehicle was seen first pulling away?

e) private car
f) motor bicycle
g) black taxi
h) bus

Q2 What was the passing taxi advertising?

e) Tennents Lager
f) Guinness
g) Volvic Mineral Water
h) Irn Bru

Q3 How many pedestrians were seen standing beside the parked cars?

e) zero
f) one
g) two
h) three

Q4 Where did the scene take place?

e) Tam's Bar
f) Toad's Bar
g) Todd's Bar
h) Top Bar

Q5 What was the background colour of the sign?

e) orange
f) yellow
g) green
h) red

Q6 What material was the table made of?

e) wood
f) metal
g) glass
h) other

Q7 What colour was the waiter's hair?

e) blonde
f) brown
g) dark
h) dark with blonde streaks

Q8 Where did the waiter put his notebook?

e) shirt pocket
f) back right trouser pocket
g) back left trouser pocket
h) on a chair

Q9 How many chairs did the waiter re-position at the table?

e) one
f) two
g) three
h) four

Q10 What did the waiter use to clean the table?

e) shirt sleeve
f) cloth
g) scrubbing brush
h) green sponge

Q11 How many people first approached the bar?

e) 2 girls
f) 1 girl
g) 2 men
h) 1 girl & 1 man

Q12 Why did they decide to stop at the bar?

e) to have lunch
f) to have an alcoholic drink
g) to have coffee
h) other

Q13 How did girl 1 carry her bags?

e) in her left hand
f) over left shoulder
g) over right shoulder
h) in right hand

Q14 What colour was the shopping bag of Girl 1?

e) Black
f) Red
g) Yellow
h) Green

Q15 After girl 2 had sat down where did she place her bags?

e) on ground to her left
f) on her knee
g) on another seat
h) on ground to her right

Q16 What accent did girl 1 have?

e) English
f) Irish
g) Australian
h) North American

Q17 Did girl 2 have any distinguishing features?

e) silver streak to front of hair
f) wore glasses
g) tattoo on back of hand
h) nose pierced

Q18 How did the waiter greet the girls?

e) Hello – what can I get you?
f) Would you like to order?
g) Good afternoon – what can I get you?
h) Hiya – what would you like?

Q19 What did girl 2 order?

e) Gin & tonic
f) latte
g) cappuccino
h) a chicken salad

Q20 On leaving the table, where did the waiter put his notepad & pen?

e) in his right hand
f) in his left hand
g) in his front right trouser pocket
h) in his shirt pocket

Q21 What did girl 2 tap on the table?

e) Menu
f) Pencil
g) Both her hands
h) Newspaper

Q22 From what direction did the passer-by enter the scene?

e) Screen left
f) Screen right
g) In front of the table
h) Other

Q23 What was the passerby holding?

e) A street map
f) A cigarette
g) Copy of the Big Issue
h) A computer case

Q24 Did passerby have any distinguishing features?

e) curly hair
f) beard & moustache
g) wore glasses
h) tattoo on back of hand

Q25 What colour was the passerby's jacket?

e) black
f) blue
g) brown
h) grey

Q26 Why did the passerby stop at the table?

e) To ask for a date
f) To ask for a light
g) To sell the Big Issue
h) To ask for directions

Q27 Where did the passerby want to go?

e) Theatre Royal
f) The Ticket Office
g) Cineworld
h) The Pavillion Theatre

Q28 Where did girl 1 say it was located?

e) Ingram & Candleriggs
f) Albion Street & Candleriggs
g) Argyle Street
h) Sauchiehall Street

Q29 What was girl 1's shopping purchase?

e) dress
f) shoes
g) necklace
h) bag

Q30 How did the waiter carry the drinks tray?

e) in both hands
f) in his right hand
g) in his right hand – over his shoulder
h) in left hand – over his shoulder

Appendix G (5) - Memory Questionnaire - Condition 5

In this questionnaire we would like you to tell us what you can remember about the film that you saw earlier. You will be asked a number of questions about the people, places, dialogue and events depicted within the film. Please take your time and answer each of the questions to the best of your ability.

Each question is followed by four choices. Draw a circle around the letter corresponding to your choice. Mark **only one** letter for each question.

We are particularly interested in the accuracy of your memory account.

One point will be awarded for each correct answer however one point will be deducted for each incorrect answer.

YOU MAY AVOID THE POINTS SYSTEM IF YOU CHOOSE NOT TO ANSWER A PARTICULAR QUESTION. YOU WILL NOT BE PENALIZED OR REWARDED FOR OMITTED RESPONSES.

Q1 What type of vehicle was seen first pulling away?

i) private car
j) motor bicycle
k) black taxi
l) bus

Q2 What was the passing taxi advertising?

i) Tennents Lager
j) Guinness
k) Volvic Mineral Water
l) Irn Bru

Q3 How many pedestrians were seen standing beside the parked cars?

i) zero
j) one
k) two
l) three

Q4 Where did the scene take place?
i) Tam’s Bar
j) Toad’s Bar
k) Todd’s Bar
l) Top Bar

Q5 What was the background colour of the sign?
i) orange
j) yellow
k) green
l) red

Q6 What material was the table made of?
i) wood
j) metal
k) glass
l) other

Q7 What colour was the waiter’s hair?
i) blonde
j) brown
k) dark
l) dark with blonde streaks

Q8 Where did the waiter put his notebook?
i) shirt pocket
j) back right trouser pocket
k) back left trouser pocket
l) on a chair

Q9 How many chairs did the waiter re-position at the table?
i) one
j) two
k) three
l) four

Q10 What did the waiter use to clean the table?

i) shirt sleeve
j) cloth
k) scrubbing brush
l) green sponge

Q11 How many people first approached the bar?

i) 2 girls
j) 1 girl
k) 2 men
l) 1 girl & 1 man

Q12 Why did they decide to stop at the bar?

i) to have lunch
j) to have an alcoholic drink
k) to have coffee
l) other

Q13 How did girl 1 carry her bags?

i) in her left hand
j) over left shoulder
k) over right shoulder
l) in right hand

Q14 What colour was the shopping bag of Girl 1?

i) Black
j) Red
k) Yellow
l) Green

Q15 After girl 2 had sat down where did she place her bags?

i) on ground to her left
j) on her knee
k) on another seat
l) on ground to her right

Q16 What accent did girl 1 have?

i) English
j) Irish
k) Australian
l) North American

Q17 Did girl 2 have any distinguishing features?

i) silver streak to front of hair
j) wore glasses
k) tattoo on back of hand
l) nose pierced

Q18 How did the waiter greet the girls?

i) Hello – what can I get you?
j) Would you like to order?
k) Good afternoon – what can I get you?
l) Hiya – what would you like?

Q19 What did girl 2 order?

i) Gin & tonic
j) latte
k) cappuccino
l) a chicken salad

Q20 On leaving the table - Where did the waiter put his notepad & pen?

i) in his right hand
j) in his left hand
k) in his front right trouser pocket
l) in his shirt pocket

Q21 What did girl 2 tap on the table?

i) Menu
j) Pencil
k) Both her hands
l) Newspaper

Q22 From what direction did the passer-by enter the scene?

i) Screen left
j) Screen right
k) In front of the table
l) Other

Q23 What was the passerby holding?

i) A street map
j) A cigarette
k) Copy of the Big Issue
l) A computer case

Q24 Did passerby have any distinguishing features?

i) curly hair
j) beard & moustache
k) wore glasses
l) tattoo on back of hand

Q25 What colour was the passerby's jacket?

i) black
j) blue
k) brown
l) grey

Q26 Why did the passerby stop at the table?

i) To ask for a date
j) To ask for a light
k) To sell the Big Issue
l) To ask for directions

Q27 Where did the passerby want to go?

i) Theatre Royal
j) The Ticket Office
k) Cineworld
l) The Pavillion Theatre

Q28 Where did girl 1 say it was located?

i) Ingram & Candleriggs
j) Albion Street & Candleriggs
k) Argyle Street
l) Sauchiehall Street

Q29 What was girl 1's shopping purchase?

i) dress
j) shoes
k) necklace
l) bag

Q30 How did the waiter carry the drinks tray?

i) in both hands
j) in his right hand
k) in his right hand – over his shoulder
l) in left hand – over his shoulder

Appendix G (6) - Memory Questionnaire - Condition 6

In this questionnaire we would like you to tell us what you can remember about the film that you saw earlier. You will be asked a number of questions about the people, places, dialogue and events depicted within the film. Please take your time and answer each of the questions to the best of your ability.

Each question is followed by four choices. Draw a circle around the letter corresponding to your choice. Mark **only one** letter for each question.

We are particularly interested in the accuracy of your memory account.

Five points will be awarded for each correct answer however one point will be deducted for each incorrect answer. An additional monetary prize payable to charity, of £25.00 will be awarded for the most accurate performance.

Nominated Charity (please specify)…………………………………………………..

YOU MAY AVOID THE POINTS SYSTEM IF YOU CHOOSE NOT TO ANSWER A PARTICULAR QUESTION. YOU WILL NOT BE PENALIZED OR REWARDED FOR OMITTED RESPONSES.

Q1 What type of vehicle was seen first pulling away?
- m) private car
- n) motor bicycle
- o) black taxi
- p) bus

Q2 What was the passing taxi advertising?
- m) Tennents Lager
- n) Guinness
- o) Volvic Mineral Water
- p) Irn Bru

Q3 How many pedestrians were seen standing beside the parked cars?
- m) zero
- n) one
- o) two
- p) three

Q4 Where did the scene take place?

m) Tam’s Bar
n) Toad’s Bar
o) Todd’s Bar
p) Top Bar

Q5 What was the background colour of the sign?

m) orange
n) yellow
o) green
p) red

Q6 What material was the table made of?

m) wood
n) metal
o) glass
p) other

Q7 What colour was the waiter’s hair?

m) blonde
n) brown
o) dark
p) dark with blonde streaks

Q8 Where did the waiter put his notebook?

m) shirt pocket
n) back right trouser pocket
o) back left trouser pocket
p) on a chair

Q9 How many chairs did the waiter re-position at the table?

m) one
n) two
o) three
p) four

Q10 What did the waiter use to clean the table?

m) shirt sleeve
n) cloth
o) scrubbing brush
p) green sponge

Q11 How many people first approached the bar?

m) 2 girls
n) 1 girl
o) 2 men
p) 1 girl & 1 man

Q12 Why did they decide to stop at the bar?

m) to have lunch
n) to have an alcoholic drink
o) to have coffee
p) other

Q13 How did girl 1 carry her bags?

m) in her left hand
n) over left shoulder
o) over right shoulder
p) in right hand

Q14 What colour was the shopping bag of Girl 1?

m) Black
n) Red
o) Yellow
p) Green

Q15 After girl 2 had sat down where did she place her bags?

m) on ground to her left
n) on her knee
o) on another seat
p) on ground to her right

Q16 What accent did girl 1 have?

m) English
n) Irish
o) Australian
p) North American

Q17 Did girl 2 have any distinguishing features?

m) silver streak to front of hair
n) wore glasses
o) tattoo on back of hand
p) nose pierced

Q18 How did the waiter greet the girls?

m) Hello – what can I get you?
n) Would you like to order?
o) Good afternoon – what can I get you?
p) Hiya – what would you like?

Q19 What did girl 2 order?

m) Gin & tonic
n) latte
o) cappuccino
p) a chicken salad

Q20 On leaving the table - Where did the waiter put his notepad & pen?

m) in his right hand
n) in his left hand
o) in his front right trouser pocket
p) in his shirt pocket

Q21 What did girl 2 tap on the table?

m) Menu
n) Pencil
o) Both her hands
p) Newspaper

Q22 From what direction did the passer-by enter the scene?

m) Screen left
n) Screen right
o) In front of the table
p) Other

Q23 What was the passerby holding?

m) A street map
n) A cigarette
o) Copy of the Big Issue
p) A computer case

Q24 Did passerby have any distinguishing features?

m) curly hair
n) beard & moustache
o) wore glasses
p) tattoo on back of hand

Q25 What colour was the passerby's jacket?

m) black
n) blue
o) brown
p) grey

Q26 Why did the passerby stop at the table?

m) To ask for a date
n) To ask for a light
o) To sell the Big Issue
p) To ask for directions

Q27 Where did the passerby want to go?

m) Theatre Royal
n) The Ticket Office
o) Cineworld
p) The Pavillion Theatre

Q28 Where did girl 1 say it was located?

m) Ingram & Candleriggs
n) Albion Street & Candleriggs
o) Argyle Street
p) Sauchiehall Street

Q29 What was girl 1's shopping purchase?

m) dress
n) shoes
o) necklace
p) bag

Q30 How did the waiter carry the drinks tray?

m) in both hands
n) in his right hand
o) in his right hand – over his shoulder
p) in left hand – over his shoulder

Appendix H - Robbery Summary

	SCORING (Units of Info.)	SPEECH
ROADSIDE Road has a single yellow line Building in the background Building walls are pebble-dashed (office-type) Walls also have large vents (air conditioning-type)	4 - SURROUNDINGS	
CAR Parked at the side of a road Parked next to opening / entrance Car is red / burgundy in colour Car model is a Rover 200 / Rover 25 Silver wheel trims Registration number is N51 OGS	6 - OBJECT	
GIRL # 1 Aged 25-35 Ethnicity = White Dark hair –	12 - PERSON	

Long - tied back in a pony tail Wearing - Blue coat (denim) - Grey skirt - White blouse - Black jumper - Black boots - Black tights Speaks with an English accent		
<u>Girl # 1</u> walks from driver's side to front of car	1 - PERSON	
<u>Girl # 1</u> opens bonnet of car	1 - PERSON	
<u>MAN # 1</u> Aged 35-45 Ethnicity = White Short grey hair Wearing - Black coat - Black trousers - Black shoes - Burgundy shirt <u>Man # 1</u> Carries bag Bag is in his right hand Bag is black Bag is laptop / computer case-type bag	16 - PERSON	

Bag is being held by handle Bag has a shoulder strap Speaks with an English Accent Clean Shaven		
Man # 1 enters scene from right of screen	1 - PERSON	
Man # 1 speaks to Girl # 1	1 - PERSON 3 - DIALOGUE	Man # 1 : Having problems with your car? Girl # 1 : Yeah, it was making a strange rattling noise earlier and now it's just stopped Man # 1 : Yeah, it looks like it's overheating
Man # 1 walks towards car & Girl # 1	1 - PERSON	
Man # 1 puts bag down on the roadside	1 - PERSON	
Man # 1 and Girl # 1 look at car engine	1 - PERSON	
MAN # 2 Aged 30-40 Ethnicity = White	9 - PERSON	

Wearing - Blue jeans - Dark jacket Short dark hair Both hands in jacket pockets Speaks with a Scottish accent Clean Shaven		
<u>Man # 2</u> enters scene from right of screen	1 - PERSON	
<u>Man # 2</u> speaks to <u>Girl # 1</u> & <u>Man # 1</u>	1 - PERSON 2 - DIALOGUE	Man # 2 : Alright, mate. Can I give you a hand? My brother's a mechanic. Man #1 : Yeah, it's the lady's car. Overheating, I'd say.
<u>Man # 2</u> speaks to <u>Girl # 1</u> & <u>Man # 1</u>	2 - DIALOGUE	Man # 2 : Could be the head gasket's gone. Happens all the time. This your car? Aye, well if the head gasket's gone, you're stuck. You'll need a tow. Man #1 : I mean, it might just be a water leak. We should try putting some water in first.
<u>**MAN # 3**</u> Aged 20-30 Ethnicity = White Wearing - Grey / blue Jacket	11 - PERSON	

- Brown scarf / collar - Jeans Rucksack on shoulder Short dark hair Both hands in jacket pockets Man # 3 walks past (from left to right) Clean Shaven		
MAN # 4 Aged 45-55 Ethnicity = White Wearing - Brown jacket - Dark trousers Short dark receding hair Heavy Build Man # 4 walks past (from left to right) Clean Shaven	9 - PERSON	
Man # 4 picks up bag with right hand	1 - PERSON	
Man # 2 exits the scene to the left	1 - PERSON 1 - DIALOGUE	Man #2 : aye well, suit yourself

Girl # 1 takes telephone out of her right–hand jacket pocket	1 - PERSON 1 - DIALOGUE	Girl # 1 : I'll have to call my husband
Man # 1 realises bag is missing	1 - PERSON 7 - DIALOGUE	Man # 1 : wait a minute! Where's my case? Girl # 1 : what is it? Man # 1 : I don't believe this! My case. Someone's lifted my case. I put it down there. Girl # 1 : Did it have anything valuable in it? Man # 1 : Had my laptop in it and a whole load of confidential stuff, and my keys. All my keys.
Man # 1 looks to the right of the screen	1 - PERSON 2 - DIALOGUE	Girl # 1 : I didn't see anyone take it but I think a couple of people have walked by. Look, I'm really sorry about this.
Man # 1 looks to the left of the screen	1 - PERSON 5 - DIALOGUE	Man # 1 : Did you see what they looked like. Girl # 1 : No, not really. I was watching what you were doing. Man # 1 : Suppose I'll have to call the police. Much good that'll do. Girl # 1 : Here, use my phone. Man # 1 : Wait a minute. I think it was…

Man # 1 leaves screen Runs off to the left side of the screen (Wrong direction / in direction of Man # 2) Woman is left standing on her own	1 - PERSON 1 - PERSON	

TOTALS	
PERSON	**74**
DIALOGUE	**22**
OBJECT	**6**
SURROUNDINGS	**4**
GRAND TOTAL	**106**

Appendix I - 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 the scene to the left or 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 passer-by 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 the scene to the left or the right?	1		Right	Left	Y	A
18	The second passer-by 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 the scene to the left or 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 taken?	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 J - 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 much so. Please answer honestly and do not omit any items.

1) How easy did you find it to recall details of the crime scene during questioning?

Not at all				Very much so
1	2	3	4	5

2) To what extent did you feel distracted during the questioning?

Not at all				Very much so
1	2	3	4	5

3) To what extent did you feel influenced during the questioning?

Not at all				Very much so
1	2	3	4	5

4) To what extent did you feel comfortable during the questioning?

Not at all				Very much so
1	2	3	4	5

5) How easy was it for you to concentrate during the questioning?

Not at all				Very much so
1	2	3	4	5

6) Did you feel that the questions were reasonable?

Not at all				Very much so
1	2	3	4	5

7) To what extent did you feel under pressure during the questioning?

Not at all				Very much so
1	2	3	4	5

8) Overall how difficult did you find the interview?

Not at all				Very much so
1	2	3	4	5

Appendix K - 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 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 maximum score 29 (range 0-29) **% response change** – number of responses changed as percentage of no. of questions **Response change towards inaccuracy** – number of responses which change towards inaccuracy following negative feedback Maximum score 29 (0-29) RC toward inaccuracy as % of no. of questions RC toward inaccuracy as % of total changes