The impact of stereotype threat on older adults' performance on a diagnostic test for dementia

Victoria Page

D.Clin.Psy thesis (Volume 1), 2012

University College London

Overview

This thesis is presented in three parts. The overall focus of the thesis is the impact of stereotype threat on older adults' cognitive performance. Part one presents a systematic review of research that investigates the evidence for a stereotype threat effect on memory in older adults, and further explores possible mediators and moderators which might explain this effect. Part two is an empirical paper that extends this framework to explore issues encountered in clinical practice by investigating the effects of stereotype threat on older adults taking a diagnostic test for dementia. Part three is a critical appraisal of the investigation presented in the empirical paper. Consideration is given to a number of conceptual and methodological issues pertinent to this study in particular, the applicability of the stereotype threat framework to older adults and cognitive performance and the generalisability of these findings given the characteristics of the sample used in this study. The appraisal concludes with some ideas on the experience of the participants in this study and ways this research may have been conducted differently.

Table of contents

Part 1: Literature Review

Is there a stereotype threat effect with older adults and tests of memory; if so what variables moderate and mediate this effect?

Part 2: Empirical Paper

The impact of stereotype threat on older adults' performance on a diagnostic test for dementia

Part 3: Critical Appraisal

Tables and figures

Literature review

Empirical paper

Acknowledgements

First and foremost I would like to thank my research supervisors, Dr Josh Stott & Dr John King for their guidance, support and patience over the last few years. Their input in devising, implementing and writing this research has been invaluable and has made the whole process much less daunting than I had anticipated.

I am eternally grateful to all those who graciously gave up their time to take part in the study. Without them this research would not have been possible; I am indebted to them.

I would like to say a huge thank you to all of my friends and family, in particular special thanks go to my mum who has always believed I can do anything; her unfailing confidence in me has helped enormously on this journey. Thanks also to Granny whose wisdom and calm words of encouragement have always been appreciated in times of stress. And finally enormous thanks to my husband Richard whose patience, understanding and total support I have always been assured of. Richard - thank you for being there for me every step of the way.

Part 1: Literature Review

Is there a stereotype threat effect with older adults and tests of memory; if so what variables moderate and mediate this effect?

Abstract

Aims: Stereotype threat theory predicts that negative stereotypes about ageing and cognition will lead older adults to perform poorly on memory tests. This review examines the literature for evidence of such an effect and explores potential explanatory mechanisms.

Method: To address these questions a search of the literature was conducted on PsychINFO and CINAHL Plus revealing twelve papers which covered seventeen studies relevant to this topic.

Results: Limited evidence was found for a direct stereotype threat effect on recall and recognition. Education was highlighted as an important factor for consideration in this field of research, and possible evidence for a role of a motivational mechanism was found.

Conclusions: Stereotype threat offers some explanation for older adults' performance on memory tests however the picture is more complex than the theory suggests. Based on this recommendations have been made for future research.

Introduction

Ageing & stereotype threat

In Western culture ageing is viewed as a process of inevitable cognitive decline, and as such older adults are routinely subjected to negative stereotypes regarding their mental competence (Kite & Johnson, 1988; Stein, Blanchard-Fields & Hertzog, 2002). Research has suggested that such stereotypes can lead to significant performance decrements on tests of cognitive ability (see Horton, Baker, Pearce & Deakin, 2008 for review).

Stereotype activation, and its impact on performance, has been investigated in relation to older adults and memory using two distinct methods. The first involves implicit primes which are assumed to activate stereotypes about ageing and memory at the subconscious level (e.g. Levy, 1996; Stein et al., 2002; Hess, Hinson & Statham, 2004). Although findings from these studies reveal mixed results there is some evidence that memory performance can, without conscious knowledge, be affected by subtle environmental cues (Hess, Hinson & Hodges, 2009). The second and more widely researched method of stereotype activation involves the use of explicit priming techniques which take effect within the realm of the participants' awareness. The grounding for methods of explicit stereotype activation lies in the work of Steele (1997; Steele & Aronson, 1995) who coined the term 'stereotype threat' to explain the observed effects of negative stereotypes on performance. Stereotype threat can be defined as the situational pressure and emotional distress an individual experiences when they become aware of potentially confirming as self-characteristic a negative stereotype about a group to which they belong, and the fear of being judged and treated according to this stereotype. The consequences of this can include altered behaviour and performance decrements in the stereotyped

domain (Steele & Aronson, 1995; Steele, 1997). Stereotype threat effects were first reported by Steele & Aronson (1995) who observed a decline in the academic performance of Black students when a negative stereotype about their racial group's academic ability was made salient. In other areas of research it was later found that stereotype threat had similar negative effects on women taking maths tests (Spencer, Steele & Quinn, 1999), men's performance on affective tasks (Leyens, Desert, Croizet & Darcis, 2000) and older people's cognitive performance (Horton et al., 2008).

Factors affecting the impact of stereotype threat on performance

In the wider stereotype threat literature there have been repeated attempts to elucidate the mechanism that accounts for performance deficits arising from threat. There is a general consensus that identification with the stereotyped group and domain are necessary prerequisites for stereotype threat to take effect; however the manner in which this proceeds to impair performance remains heavily debated.

Identification with the stereotyped group and domain

According to Steele (1997) for a negative stereotype to be threatening it must be self relevant ie. the individual must firstly identify with the stereotyped group, and secondly have some degree of self-identification with performance in the stereotyped domain.

The definition and understanding of group identity therefore is of primary importance in understanding the effects of stereotype threat. Social identity theory suggests that the existence of a 'group' depends on the psychological activity of a number of individuals who construe and evaluate themselves in terms of common

attributes that differentiate them collectively from others (Hogg, 2006). Individuals have a range of social identities pertaining, for example, to their gender, age, race, ethnicity and religion. The extent to which membership of these social categories or groups forms a central part of their self-identity however can vary. If membership of a given group forms an important part of an individual's identity they should have a stronger motivation to maintain a positive image of that group identity and therefore experience greater threat at the suggestion that this group is in some way inferior to others. For example women who considered their female identity an important aspect of their self-identity were more negatively affected than women with a weaker gender identity by stereotypes about women and maths performance (Schmader, 2002). Older adults who strongly identify with their age group may similarly be more vulnerable to cues that this group is being devalued in terms of its cognitive ability leaving such individuals more susceptible to the effects of age stereotype threat (Kang & Chasteen, 2009).

The second mode of identification in relation to stereotype threat pertains to the stereotyped domain. For an individual to feel threatened by a negative stereotype the performance domain must be important to their self-definition (Schmader, 2002). For example, Steele & Aronson (1995) found that Black students who had a high level of investment in their academic ability were most susceptible to the effects of negative stereotypes relating to racial differences in academic performance. For older adults the theory follows that those who value their cognitive resources the most will be more susceptible to stereotypes pertaining to ageing and memory than those who value this ability less (Wheeler & Petty, 2001).

Mechanisms of performance impairment

Group and domain identification can therefore be viewed as rudimentary requirements for successful induction of stereotype threat; but how then does activating threat actually impair performance? There have been a number of hypotheses offered to explain observed performance deficits following stereotype induction.

Reduction in cognitive resources

Steele & Aronson (1995) proposed that the most likely explanation for performance deficits under conditions of stereotype threat was inefficiency in cognitive processing during task performance. This hypothesis was investigated by Schmader & Johns (2003) who found evidence that, for women taking a maths test, stereotype threat impaired performance via a reduction in working memory capacity. Because working memory involves the ability to focus one's attention on a given task whilst inhibiting irrelevant information they suggest that performance decrements are a result of stereotype activation diverting attention onto task-irrelevant worries. They did not however investigate the specific nature of these worries that were competing for cognitive resources. Cadinu, Maas, Rosabianca & Kiesner (2005) later found evidence that individuals under stereotype threat spontaneously engage in negative task-related thinking which inhibits performance. They conclude that performance deficits are caused not by general worries or feelings of anxiety but rather by domain-specific thought intrusion. One example of such task-related anxiety might be evaluation apprehension which can cause stereotyped individuals to exercise excessive caution when performing a task for fear of being negatively judged in that domain (Cadinu, Maass, Frigerio, Impagliazzo & Latinotti, 2003).

Affective responses

Another line of thinking has examined the role of negative affect and physiological arousal based on the hypothesis that threat increases negative affect which may impair performance through greater levels of arousal (Hess, Hinson et al., 2009). In contrast with the argument above, this would suggest a role for a more general emotional response. A slightly different view of the relationship between threat and affect in older adults is that threat may be associated with a reduction in self-reported positive affect whilst negative affect remains fairly stable (Kang & Chasteen, 2009). Interestingly, cognitive performance has also been noted to share stronger links with self-reported positive affect; Hill, van Boxtel, Ponds, Houx & Jolles (2005) found that positive, rather than negative affect predicted older adults' performance on free recall tasks.

In relation to physiological measures of arousal older adults display greater stress responses than younger adults during cognitive testing which negatively affects test performance (Hess, Hinson et al., 2009). Further evidence for this comes from Neupert, Miller & Lachman (2006) who found increased cortisol levels in older but not younger adults during cognitive testing. The relationship between physiological markers of arousal and cognitive performance is well documented and there is evidence to suggest that memory tasks may be impaired to a greater degree than other cognitive tasks (Wright, Kunz-Ebrecht, Iliffe, Foese & Steptoe, 2005). This could offer a plausible explanation for older adults' heightened arousal during testing and account for reduced memory performance.

Motivational/effort based mechanism

The significance of motivation and effort in explaining the effect of threat on performance has long been cited (e.g. Steele & Aronson, 1995). In the wider stereotype threat literature motivational explanations for task performance under conditions of threat have been investigated using measures of performance expectations and self-efficacy (Steele & Aronson, 1995; Cadinu, Maas, Frigerio et al., 2003; Spencer, Steele & Quinn, 1999). This is based on the hypothesis that stereotype threat lowers expectations about both performance and self-efficacy, (Spencer et al., 1999) and low expectations can negatively affect motivation and effort thus having a detrimental impact on performance (Bandura, 1977).

In relation to effort, two opposing hypotheses have been explored as a means of accounting for the stereotype threat – performance link. One argument is that stereotype threat induces a state akin to learned helplessness which causes individuals to shut down, expending less effort than those who do not experience threat (Smith, 2004). The counterargument is that stereotype threat propels individuals to disprove the stereotype and put in more effort which can cause a 'choking under pressure' effect (Baumeister & Showers, 1986). There have been difficulties and inconsistencies however between studies in operationalising effort with various attempts to do so including measures of time spent on test items, number of items attempted and self-report estimates of effort (Smith, 2004).

Aims of the current review

Several reviews of the stereotype threat literature have been conducted to date (Table 1), the most relevant of these being that of Horton et al. (2008) who carried

out a meta-analysis showing that the impact of stereotype priming on memory performance yielded a small-medium effect size of 0.38. However this did not discriminate between studies which used implicit as opposed to explicit priming techniques. The current review differs from this and other reviews by focusing in detail on the effects of explicitly activated stereotype threat on older adults' memory performance, providing a thorough evaluation of the evidence for and against such an effect. The review will also investigate variables which might explain the mechanism of stereotype threat by examining their role as mediators or moderators of this phenomenon.

Table 1

Previous reviews

Authors	Year	Focus of review	Method of review	Main difference from current review
Horton, Baker, Pearce & Deakin	2008	Impact of stereotype primes on seniors	Meta-analysis	Differential effects of implicit and explicit primes
				Impact on multiple domains (cognitive, physical, physiological, psychological)
				Not concerned with mechanisms
Kitt, Tuokko & Mateer	2008	Implications of stereotype threat for neurological populations	Conceptual review of literature	General review – stereotypes of race, gender and age.
				Extension to neurological populations.
Smith	2004	Review of mediational variables involved in stereotype threat	Conceptual review of literature	General review – stereotypes of race, gender, SES and age.
Wheeler & Petty	2001	Mechanisms involved in stereotype threat effects on behaviour	Conceptual review of literature	General review – stereotypes of race, gender, SES and age.

Method

The current review was limited to quantitative studies which utilised an explicit priming technique with the intention of inducing stereotype threat in older participants relevant to their performance on a memory task.

Inclusion and exclusion criteria

Participant characteristics

Studies were included if they involved non-clinical populations of older adults and excluded if they involved older adults with known memory problems. Studies which investigated stereotype threat in middle aged or young adults were excluded unless older adults were also represented as a comparison group.

Intervention characteristics

This review was concerned solely with studies using explicit methods of stereotype threat activation as there are reported differences in peoples' response to experimental manipulations depending on their level of awareness (Devine, 1989). To assist accurate comparison between studies those using implicit or subliminal priming techniques were therefore excluded.

Outcome measures

Only studies measuring performance on at least one memory task were included in the review, although several studies measured other outcomes as well.

Research design

Only studies which reported experimental data involving more than one group of participants were included. Narrative accounts and reviews of the literature in this area were excluded.

Search Strategy

A search was conducted in PsychINFO using the Ovid interface and in CINAHL Plus and MEDLINE using EBSCO Host. Based on the inclusion and exclusion criteria the search terms in Table 2 were entered individually and then combined.

Table 2

Search categories and terms

Category	Participant characteristics	Intervention characteristics	Outcome measures
Search terms	old* age* eld*	stereotype threat (keyword) stereotyp* threat (keyword and subject heading auto-exploded)	memory (keyword and subject heading auto-exploded)

* indicates terms that were truncated to allow for multiple endings of words.

The combination of these search terms is detailed in Fig. 1. Initially 446 papers were identified on PsychINFO and 325 papers on CINAHL Plus and MEDLINE. Limiting these results to journal paper, English language and human subjects reduced the numbers on PsychINFO to 281 and those on CINAHL Plus and MEDLINE to 305.

The titles, and if required the abstracts, of these papers were read and those which appeared relevant to the area of stereotype threat, ageing and memory were marked for further consideration. Thirty-three papers were selected from PsychINFO and twenty-three from CINAHL Plus and MEDLINE. It was noted at this stage that all 23 papers from CINAHL Plus and MEDLINE already featured in the PsychINFO search results therefore further analysis focused solely on the broader PsychINFO results. These 33 papers were read in full to establish their suitability for inclusion in the review. Eleven papers were deemed appropriate, involving experimental studies with more than one group of participants of which at least one group was older adults. They utilised explicit rather than implicit methods of inducing stereotype threat, and performance on at least one test of memory was measured. A hand search of the references in these papers was conducted and one further paper of interest was found giving a total of 12 papers for inclusion in this review.

Analysis

Twelve papers detailing a total of seventeen studies were examined to assess whether they found evidence for a stereotype threat effect in older adults on tests of memory and if so, whether they established potential mechanisms to explain this effect.

Figure 1

Search strategy and results

Box	Contents

Old* or
Age* or
Eld*

Memory

Stereotype
threat

Stereotyp*

16 *9*

79 *26*

383 *363*

446 *325*

281 *305*

33 *23*

11

12

Limits applied   Read titles/abstracts   Read paper   Hand search

Searches combined with AND   Searches combined with OR
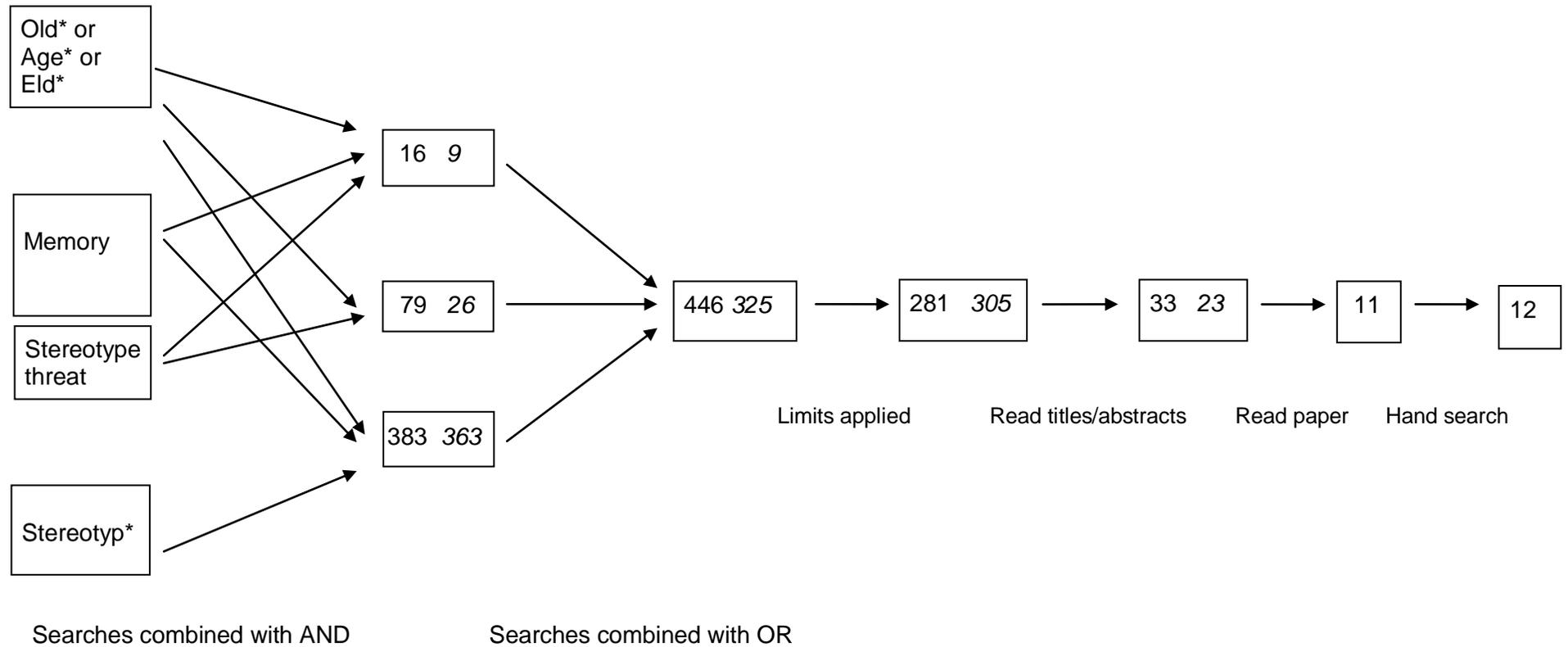
Initial figures indicate number of papers found on PsychINFO; figures in italics refer to papers found on Medline and CINAHL Plus

Results

Seventeen studies from twelve papers were included in this review. The majority of the studies took place in the US (nine) and Canada (five), with one study each from the UK, France and Romania. Sample sizes ranged from 42 – 162 and, with the exception of one female-only study, involved both male and female participants. Ages ranged from 17 – 92 across studies however there was significant variation in terms of defining 'younger' or 'older' adults, with the former ranging from 17 – 56 and the latter ranging from 56 – 92. One study used a 'middle aged' group (40 – 59), whilst another split older adults into young-old (60-70) and old-old (71-82).

Of the 17 studies the majority (14) included some measure of recall performance. Five studies employed tasks of recognition. There was some variation in the way these abilities were measured. Twelve of the seventeen studies compared stereotype threat effects across age groups, including one comparison between young-old and old-old participants. Post-2005 literature indicated a shift away from investigating differences between age groups with the remaining five studies published after this date focusing specifically on older adult participants but using different priming techniques or covariates.

All studies examined the effects of stereotype threat on memory performance and 14 included further analyses of mediators or moderators of this effect.

For the purposes of this review the main findings of these studies have been presented in two parts:

1. The effect of stereotype threat on memory in older adults (Table 3).
2. Mediators and moderators of stereotype threat effects on memory (Table 4).

Table 3

The effect of stereotype threat on memory in older adults

Authors	Year	Participants (age)	N	Experimental manipulation	Dependent measures	Evidence for ST effect	Signific ance (p)	Effect size
Horton, Baker, Pearce & Deakin	2010	Older adults (60-75)	96	3 experimental conditions: *Positive* – participants primed with newspaper article & newscast portraying seniors as high in vitality and wisdom *Negative* – participants primed with newspaper article & newscast highlighting declining memory and physical abilities of seniors *Control* - no priming	Word list recall (30 words)	No	-	n/a
Kang & Chasteen	2009	Female older adults (62-84)	42	2 experimental conditions: *Threat* - tasks framed as diagnostic of memory ability, completed with young confederate and conspicuously timed *Non-threat* - tasks framed as examining thoughts and opinions, completed in same age pairs, inconspicuously timed	Memory for 60 word prose passage:			
					Free recall	Yes	<.05	d= .68
					Cued recall	No	-	
					Recognition	No	-	
Hess, Emery & Queen	2009	Older adults (60-86)	82	2 conditions each with 2 levels: *Stereotype condition* -*Threat* – participants told that the study was interested in age differences in memory ability and that younger adults typically performed much better on this memory task than older adults -*No threat* – participants told that the study was about individual differences in ability and older adults did quite well on this task. *Response condition* -*Deadline* – limited time to make a recognition response -*No deadline* – unlimited response time	50 word remember, know or guess recognition task	Yes, but only in the deadline condition	.02	η_p^2=.07

Author	Year	Age/Group	N	Conditions	Measure	Finding	p	Effect size
Hess, Hinson & Hodges	2009	Young-old (60-70)	53	2 experimental conditions: *Threat* – participants told the study was investigating why younger and older adults perform so differently on memory tests	Word list recall (30 words)	In young-old with higher levels of education	.01	η_p^2=.18
		Old-old (71-82)	50	*No-threat* – participants told that they would be taking a test on which younger and older adults performed similarly				
Hess & Hinson	2006	Adults (24-86)	162	2 experimental conditions: *Positive* – participants read 2 news articles contradicting stereotypical views of ageing and memory *Negative* – participants read 2 news articles supporting stereotypical views of ageing and memory	Word list recall (30 words)	Varied significantly with age	.03	η_p^2=.03
Abrams, Eller & Bryant	2006	Older adults (59-89)	97	2 threat conditions each with 2 levels of intergenerational contact: *Stereotype threat* -*High* – participants told the purpose of the study was to see if old people performed more poorly on intellectual tasks than young people -*Low* – participants told the purpose of the study was to see how people differ in their responses on different tasks *Intergenerational contact* -*Less positive* – participants reporting a smaller amount of recent positive contact with people under 35 -*More positive* – participants reporting a greater amount of recent positive contact with people under 35	Overall Test score based on comprehension, recall, digit span, verbal facility measured using items adapted from WAIS-III and CAMCOG-R[1]	Threat Threat x contact - meaning threat had a greater negative impact on those with less contact	<.001 <.01	Not reported
Chasteen, Bhattacharyya, Horhota, Tam & Hasher	2005	<u>Study 1</u> Young adults (18-25)	40	2 experimental conditions: *Threat* – exercise framed as a memorisation task *Non-threat* – exercise framed as an impression formation task	Recall for sentences using a sentence predicate task (24 items)	Yes – for both age groups	<.01	η^2=.13
		Older adults (61-87)	40					

---

[1] Wechsler Adult Intelligence Scale – third revised edition (Wechsler, 1998) and the Cambridge Cognitive Examination – Revised (Roth, Huppert, Mountjoy & Tym, 1998)

Chasteen, Bhattacharyya, Horhota, Tam & Hasher cont.	2005	<u>Study 2</u> Young adults (17-25) Older adults (65-88)	42 39	Similar study which included additional measures (discussed later in the review) 2 experimental conditions: *Threat* – exercise framed as a memorisation task *Non-threat* – exercise framed as an impression formation task	Recall for sentences using a sentence predicate task (24 items)	Yes –for both age groups	<.001	η^2=.19
		<u>Study 3</u> Young adults (17-22) Older adults (64-84)	43 45	2 experimental conditions: *Threat* – exercise framed as a memorisation task *Non-threat* – exercise framed as an impression formation task	Sentence recognition task (66 items rated as old or new sentence)	Yes - for both age groups	<.05	η^2=.07
Desrichard & Köpetz	2005	<u>Study 1</u> Young adults (19-26) Older adults (56-80)	40 40	2 experimental conditions: *Threat* – task presented as a memory task *Non-threat* – task presented as an orientation task	Running an errand task involving list recall and route planning	Age x threat interaction – older people did significantly worse in memory emphasis condition	.01	Not reported
		<u>Study 2</u> Young adults (17-56) Older adults (61-92)	60 60	2 experimental conditions: *Threat* – task presented as an assessment of memory ability *Non-threat* – task presented as an assessment of cognitive skills not including memory	Story recall	Only older adults	<.01	
					Shape recall	Only older adults	<.01	
					Visuo-visual span	Only older adults	<.01	

Andreoletti & Lachman	2004	Young (21-39)	46	3 experimental conditions: *Stereotype* – participants told that there were age differences on the test with younger adults remembering more words than older adults	Word list recall (30 words)	Significant condition x education effect – meaning threat had a greater negative effect on those with less education	<.01	η^2=.07
		Middle aged (40-59)	53	*Counterstereotype* – participants told that there were no age differences on the test with younger and older adults remembering the same number of words				
		Older adults (60-80)	50	*Control* – participants given no information about age differences on the test				
Hess, Hinson & Statham	2004	Study 1 Young adults (17-27)	62	2 conditions each with 2 levels: *Stereotype prime* -*Positive* – scrambled sentence task containing words reflecting positive views about ageing	Free recall	No	-	n/a
		Older adults (57-81)	59	-*Negative* - scrambled sentence task containing words reflecting negative views about ageing *Awareness* -*Explicit priming* – age-related words highlighted -*Implicit priming* – words were not highlighted				
		Study 2 Young adults (17-35)	36	2 conditions each with 2 levels: *Stereotype prime* -*Positive* – lexical decision task involving positive age-related words -*Negative* - lexical decision task involving negative age-related words	Free recall	No	-	n/a
		Older adults (59-82)	36	*Awareness* -*Aware* – words presented at a speed at which participants could consciously perceive them -*Unaware* – words presented at slowest speed at which participants could not identify *[Only results of explicit priming conditions are included in the review]*				

Hess, Auman, Colcombe & Rahhal	2003	Young adults (18-30)	48	3 experimental conditions: *Negative* – participants were presented with 2 research reports discussing findings that older adults' memory skills were worse than those of younger adults *Positive* – participants were presented with 2 research reports discussing more positive findings regarding the relationship between ageing and memory *Control* – no information presented	Free recall	Significant in older group	.02	$\eta^2 = .17$
		Older adults (62-84)	48					
Rahhal, Hasher & Colcombe	2001	Study 1 Young adults (17-24)	48	2 experimental conditions: *Threat* – task instructions emphasized memory component *Non-threat* – task instructions did not emphasize memory component	Memory accuracy (recall & recognition)	Significant age x threat interaction	<.05	$d = .24$
		Older adults (61-75)	48					
		Study 2 Young adults (18-24)	56	2 experimental conditions: *Threat* – task instructions emphasized memory component *Non-threat* – task instructions did not emphasize memory component	Memory accuracy (recall & recognition)	Significant age x threat interaction	<.02	$d = .65$
		Older adults (60-74)	56					

The effect of stereotype threat on memory in older adults

The first part of this review examines the evidence for a stereotype threat effect on older adults' memory performance. The studies reviewed are presented in Table 3.

Recall

Tests involving some measure of recall featured in 14 of the 17 studies reviewed, however the evidence for stereotype threat effects on older adults' performance on such tasks is mixed.

Findings demonstrating an effect on recall in older adults

In studies solely involving older participants Kang & Chasteen (2009) and Abrams, Eller & Bryant (2006) provide evidence for significant stereotype threat effects on recall. Arguing for the specificity of this effect to older adults Desrichard & Köpetz (2005), Hess, Auman, Colcombe & Rahhal (2003) and Rahhal, Hasher & Colcombe (2001) demonstrate, by comparing the effects of stereotype threat across age groups, that older participants' recall performance was impaired by threat induction whilst their younger counterparts remained unaffected. Effect sizes reported indicate a moderate-large effect in the findings of Rahhal et al. (2001, study 2), Hess et al. (2003) and Kang & Chasteen (2009), whereas the results reported by Rahhal et al. (2001, study 1) showed a small effect. Desrichard & Köpetz (2005) and Abrams et al. (2006) failed to provide an indication of effect size pertaining to their findings.

Issues with outcome measures/bias in some of these studies

The memory accuracy task used by Rahhal et al. (2001) involved both recall and recognition, thus the differential impact of threat on each cannot be ascertained. Similarly the nine-part cognitive test used by Abrams et al. (2006) included a

measure of recall ability however their results were presented only as total score rendering it impossible to quantify the impact of threat specifically on recall. The dependent variable used by Desrichard & Köpetz (2005; study 1) was not a straightforward recall task as it involved simultaneous list recall and route planning, the spatial component of which may have activated the stereotype of women having poorer abilities in this domain than men (e.g. McGlone & Aronson, 2006). In addition the younger adult group in this study were all currently attending university, whereas their older adult group were retired. It is possible therefore that groups differed in education as well as age with the younger group clearly well educated, although we have no information on the educational histories of the older participants. Their age categorisation of participants differs in study 1 and 2 and is therefore questionable. Other sampling issues are noted in the research by Kang & Chasteen (2009) which involved an all female sample whom the authors allege are more susceptible to stereotype threat effects. This however makes it hard to compare with other studies in this area. The study by Hess et al. (2003) represents a high standard of research; they used a control group for comparison with their positive and negative stereotype conditions and report that their randomisation was successful in controlling for factors that could confound the effects of the stereotype threat manipulation. Their intervention is clearly defined and their outcomes well reported.

Findings demonstrating no effect on recall

Horton, Baker, Pearce & Deakin (2010) assert that their stereotype threat induction had no significant effect on recall in their sample of older adults. Hess, Hinson & Statham (2004) found no effects for stereotype threat in either young or older adults. The research reported by Horton et al. (2010) was of a high standard; being a single blind randomised controlled trial with selection of a population, intervention and

outcome measures explicitly based on the findings of previous research. Manipulation checks revealed successful randomisation across co-varying factors such as age, medication and education. The two studies by Hess et al. (2004) were also grounded in prior research findings and used a randomised design; however the researcher was not blind to the participants' condition. There was also a difference reported between participants, with older adults having significantly more education than younger adults, which may have biased their results.

The influence of age

The findings of Hess et al. (2004) reported above raise a question about the relevance of age in predicting stereotype threat effects. Chasteen, Bhattacharyya, Horhota, Tam & Hasher (2005; studies 1 and 2), and Andreoletti and Lachman (2004) found stereotype threat effects on recall in all age groups in their studies. Hess, Hinson et al. (2009) and Hess & Hinson (2006) found that age did influence vulnerability to stereotype threat but that this was strongest for those in their late 60s, with older participants appearing less affected by threat. The influence of age on stereotype threat will be discussed in more detail below.

Recognition

Comparatively fewer studies have examined the effect of stereotype threat on recognition in older adults. Rahhal et al. (2001) found that older participants' performance was significantly worse in the threat condition compared to the non-threat condition whilst younger participants' performance remained relatively stable across conditions. However as discussed previously these studies employed a composite test of both recall and recognition and it is unclear to what extent the results pertain to each of these memory domains. Partial support for a stereotype

threat effect on recognition is offered by Hess, Emery & Queen (2009) who found a significant and moderately sized effect with older adults but only when response time was limited. Chasteen et al. (2005; study 3) found support for such an effect but, as with their previous studies involving recall tasks, the negative impact of stereotype threat on performance was evident across all age groups. The size of this effect was moderate. In contrast to these results Kang & Chasteen (2009) found no evidence of their threat induction impacting on older adults' performance on a recognition task.

Summary

Fourteen of the seventeen studies reviewed here provided some evidence to suggest that stereotype threat affected older adults' memory performance. Findings relating to the effects of stereotype threat on recall performance were mixed, suggesting a more complicated picture than the theory, at face value, might indicate with two of the best designed studies suggesting that threat affected recall regardless of age and many of the those suggesting an age difference in response to threat being poorly designed studies. One possible suggestion is that education may have a role in determining the impact of stereotype threat on recall. In relation to recognition, the evidence is weak, the only studies which hint at such an effect being those of Rahhal et al. (2001) whose outcome measure is poorly designed and Hess, Emery et al. (2009) who's findings suggest that recognition is only affected in the presence of a deadline.

31

Table 4

Mediators and moderators of stereotype threat effects on memory

Study	Age	Age group identification	Intergenerational contact	Domain identification/ importance of memory	Education	Stigma consciousness/ perceived stereotype threat	Task related anxiety	Evaluation apprehension	Task demands ie. time limit	Disrupted working memory	Strategy use	General anxiety	Negative affect	Physiological arousal	Memory self-efficacy	Performance expectations
Kang & Chasteen (2009)		Mod				Mod							X			
Hess, Hinson et al. (2009)	Mod				Mod	Mod				X	X	X	X	X		Med[2]
Hess, Emery et al. (2009)									Mod				X			
Hess & Hinson (2006)				X			Mod	X			X	X				
Abrams et al. (2006)			Mod				Med									

---

[2] This mediated the moderating effect of education

Study	Age	Age group identification	Intergenerational contact	Domain identification/ importance of memory	Education	Stigma consciousness/ perceived stereotype threat	Task related anxiety	Evaluation apprehension	Task demands ie. time limit	Reduction in working memory	Strategy use	General anxiety	Negative affect	Physiological arousal	Memory self-efficacy	Performance expectations
Chasteen et al. (2005)																
Study 2								X				X			X	
Study 3												X			X	
Desrichard & Köpetz (2005) Study 1															Mod	
Study 2																Med
Andreoletti & Lachman (2004)				Mod											X	
Hess et al. (2004) Study 1				X		X					X	X				X
Study 2						X					X	X				X
Hess et al. (2003)				Mod		X					Med					

Note: X=tested but not significant; Mod = evidence of a moderating effect; Med = evidence of a mediating effect

Mediators and moderators of stereotype threat effects on memory

The relationship between threat induction and performance deficits as well as the association with age is, as discussed above, unclear. Thirteen of the seventeen studies included in this review investigated the role of other variables in mediating or moderating the link between threat and performance (Table 4). For convenience these are divided into participant characteristics and threat based effects.

Participant characteristics

The six variables conceptualised as participant characteristics can be considered broadly representative of Steele's original factors deemed necessary for stereotype threat to occur ie. identification with the stereotyped group and domain.

Identification with the stereotyped group

Age, age group identification and intergenerational contact were investigated to ascertain the extent to which identification with the stereotyped group affects performance.

Age

There was mixed data relating to this: It was reported by Horton et al. (2010) that recall in their older adult sample deteriorated significantly as age increased with no additional evidence for threat effects on performance. In addition Chasteen et al. (2005; studies 1 and 2), and Andreoletti and Lachman (2004) found that their threat inductions had significant detrimental effects on recall performance across all age groups in their studies suggesting that stereotype threat effects may in fact be non-age specific. In Andreoletti and Lachman's (2004) study participants were stratified by age, education and gender before being randomly assigned to a group. Their effect size was moderate. The sizes of the effects reported by Chasteen et al.

(2005) were moderate for study 1 and moderate-large for study 2. However they do not report whether assignment to experimental condition was randomised and state that their testing procedure involved two participants taking the test in the same session which may have affected performance.

Hess, Hinson et al. (2009) found a trend towards significance of threat affecting recall performance in young-old but not old-old participants. Hess & Hinson (2006) found that threat had an effect up to age 68 and that this effect then declined with no evidence of threat based effects on recall for their oldest participants. The effect size for both these reported observations, however, is small. Furthermore Hess, Hinson et al. (2009) report that the young-old participants in the threat condition had significantly lower levels of education which may have influenced their results.

Age group identification

Bearing in mind the findings above it is conceivable that chronological age may be experienced differently within the older adult population with some feeling a strong connection to an age-related identity and others rejecting the association. Kang and Chasteen (2009) used age-group identification to assess the relevance of ageing stereotypes to their older adult participants finding that, irrespective of threat, as this increased recall performance decreased. They argue that age group identification could moderate the effect of threat on recall performance but do not demonstrate any evidence that this is the case. Furthermore, given that this is based on a regression model using a very small sample (N=42) these findings are not statistically robust hence it is difficult to place much weight on their assertion.

Intergenerational contact

Abrams et al. (2006) argue that older individuals with lots of positive contact with younger people should have weaker identification with their own age group and thus used this as a proxy variable. They found that the effect of threat on older adults' cognitive performance varied as a function of the amount of recent positive contact with someone under the age of 35. In the high threat condition those who reported more positive intergenerational contact during the previous week performed significantly better than those with less positive contact. The authors stated that intergenerational contact moderated the effect of stereotype threat on test performance, though this moderating effect was mediated by test-related anxiety (see below for further discussion). No other studies have addressed the effects of intergenerational contact on memory performance.

Identification with the stereotyped domain

The extent to which participants valued their memory ability was examined by direct quantification of domain identification and also approximated through measures of education and stigma consciousness.

Domain identification/importance of memory

Hess et al. (2003), using the Memory Achievement subscale of the Metamemory in Adulthood Questionnaire (MIA-Ach; Dixon & Hultsch, 1984), found that the degree to which threat impaired older adults' memory performance increased in tandem with the value they placed on memory ability concluding that domain identification moderated the impact of stereotype threat on older adults' recall performance. The studies reported by Hess et al. (2004) and Hess & Hinson (2006) using the same

questionnaire however failed to replicate this moderating effect. The fact that these studies used the same measure of domain identification facilitates accurate comparison between of their findings. One potential confounding factor however is the variation in administering this measure. Hess et al. (2004) and Hess & Hinson (2006) asked all their participants to complete the MIA-Ach questionnaire prior to attending the test session. Hess et al. (2003) did likewise for their older participants, but younger participants completed this measure immediately prior to their test, which may plausibly have had some influence on the differential significance of age on the threat x MIA-Ach interaction. Furthermore, the effect size of this interaction was small suggesting that this difference was weak in any case.

Education

Education is a useful proxy for importance placed on cognitive abilities and those with more education might be assumed more susceptible to threat (Hess, Hinson et al., 2009). Andreoletti & Lachman (2004) found a significant threat x education effect with 'low-education' individuals performing worse in both the negative and positive information conditions than in the control condition, indicating that any information pertaining to ageing and memory, whether positive or negative, had a detrimental effect on recall. Conversely those with higher levels of education showed relative immunity to stereotype threat in the positive condition, performing significantly better than those in the negative or control conditions. The authors conclude that across age groups education moderated the effects of the stereotype manipulation on recall such that higher education enabled participants to benefit from positive information about memory and ageing.

Support for this finding comes from Hess, Hinson et al. (2009) who used similar experimental conditions as above though without a younger adult control

group. Using years of education as a variable they found a significant threat x education interaction in their young-old group (60-70) with recall increasing in the positive information condition as education increased. They also found evidence for education as a moderator of stereotype threat effect on memory such that the relationship between threat and recall varied significantly as a function of education. Further analyses revealed that the moderating role of education was itself mediated by performance expectation (see below for further discussion).

Stigma consciousness/perceived stereotype threat

Those who are aware they may be stigmatised in a particular domain are likely to exhibit specific threat-related performance decrements in this domain (Hess, Hinson et al., 2009). These authors argued for a moderating effect of stigma consciousness on recall performance however this was based on a trend towards a significant association between increasing stigma consciousness and decreasing recall in their older groups in the stereotype threat condition but not in the non-threat condition.

Chasteen et al. (2005; studies 2 and 3) used the conceptually similar construct of perceived stereotype threat (PST) finding that although PST could explain age differences on recall and recognition tasks it did not account for the effect of stereotype threat on performance.

Deconstructing PST into trait or state based factors Kang & Chasteen (2009) measured both general awareness and situation specific feelings of stereotype threat in relation to ageing and memory. Examining performance on free recall, cued recall and recognition tasks they concluded that PST moderates the effect of threat on memory in older adults. This assertion however is based solely on decrements in cued recall with no significant effects on either free recall or recognition as a function of threat interacting with either trait- or state-PST.

Threat based effects

Three mechanisms for stereotype threat have been examined in the reviewed studies: reduction in cognitive resources, affective responses, and a motivational mechanism.

Reduction in cognitive resources

This mechanism was investigated using measures of task-related anxiety, task demands, reduction in working memory, strategy use and evaluation apprehension.

Task related anxiety

Hess et al. (2003) using the Memory Anxiety subscale of the Metamemory In Adulthood questionnaire (MIA-Anx; Dixon & Hultsch, 1984) found that memory related anxiety was unrelated to threat manipulation. Hess et al. (2004; study 2) similarly found that threat had no effect on either the Memory Controllability Inventory (MCI) or Ageing Concern Scales (ACS; Lachman, Bandura, Weaver & Elliot, 1995). Abrams et al. (2006) conversely did find threat effects on participants' retrospective ratings of anxiety during testing. They also found that anxiety was significantly related to performance, concluding that anxiety partially mediates the main effect of threat on performance. A confounding issue with the three studies just described is that they all measured anxiety following completion of memory tests. It is thus impossible to ascertain whether anxiety reports were due to threat or a sense of doing badly on the test.

Hess & Hinson (2006) used the ACS and MCI but asked participants to complete these measures twice to calculate change. They report a significant threat x time (pre- vs. post-test) effect on beliefs and concerns about memory. They further

found that those whose memory beliefs became more positive performed better than those whose beliefs became more negative following completion of the recall task and consequently suggest that this change in beliefs moderated memory performance. Although the idea of measuring change in anxiety is in theory a useful one, as discussed previously, administering the second measure of anxiety after test completion is unreliable.

Evaluation apprehension

Performance interfering anxiety that arises in the presence of an evaluative audience has been cited as a potential mediator of threat effects on memory performance (Spencer et al., 1999). Hess & Hinson (2006) and Chasteen et al. (2005; study 2) however reported no significant effects of threat on evaluation apprehension.

Task demands

Hess, Emery et al. (2009) investigated the effect of time constraints on recognition using a remember, know or guess paradigm which proposes that to remember, as opposed to vaguely know, that an item has been seen before participants need to attend well at the learning stage. A reduction in 'remember' responses at the recognition stage suggests divided attention at encoding. They found that independent of threat older adults' performance was significantly poorer when time was limited, and report a trend towards a significant threat x time interaction. Planned contrasts revealed that stereotype threat reduced performance only in the time limited condition suggesting that task demands moderated the effect of threat on recognition. However, given that the planned contrasts were based on a non-significant interaction these results should be treated with caution.

Disrupted working memory

It has been suggested that stereotype threat leads to a disruption in working memory which in turn causes reduced performance on explicit tests of memory. However, the only study to directly investigate this found no evidence of threat effects on working memory (Hess, Hinson et al., 2009).

Strategy use

Free recall tasks have a strong effortful component to them which necessitates the development and implementation of some kind of strategy to maximise performance (Hess, Emery et al., 2009). Using the Adjusted Ratio of Clustering (ARC; Roenker, Thompson & Brown, 1971) to assess the strategy of semantic clustering ie. semantically linked words being recalled together, Hess et al. (2003) found that clustering was significantly lower in the threat condition across age groups. Further analysis within age groups revealed that independent of threat, older adults' recall performance was positively associated with clustering and this strategy accounted for over half of the variance of stereotype threat related effects on recall. They conclude that strategy use partially mediated the impact of threat on recall. Subsequent studies however have failed to replicate this finding. Hess et al. (2004) found no significant effects for clustering. Hess & Hinson (2006) report significant effects of age on strategy and an interaction with threat however found no evidence for strategy use as a mediator. Hess, Hinson et al. (2009) found a significant main effect of threat on strategy use but found no mediating role of this variable on recall.

Affective response mechanism

Widely known to inhibit memory performance, affect and arousal were explored using measures of general anxiety, affect and physiological arousal.

General anxiety

General anxiety has long been thought to mediate the effects of stereotype threat on performance (e.g. Steele, 1995). Several studies measured state anxiety prior to memory testing using a short form of the State-Trait Anxiety Inventory (STAI-S; Spielberger, Gorsuch & Lushene, 1970) but failed to find any significant results pertaining to stereotype threat (Hess, Hinson et al., 2009; Hess & Hinson, 2006; Chasteen et al., 2005, study 2; Hess et al., 2004). Chasteen et al. (2005; study 3) did find heightened levels of state anxiety in their threat condition however this occurred across age groups.

Negative affect

Three studies measured affect using the Positive and Negative Affect Schedule (PANAS; Watson, Clark & Tellegen, 1988). Hess, Hinson et al. (2009) report a marginal influence of threat on negative affect; however this was not associated with subsequent recall performance. Hess, Emery et al. (2009) found no significant impact of threat on affect measured following a recognition task. Kang & Chasteen (2009) measured pre- to post-test change in affect finding a trend towards a decrease in positive affect under conditions of threat. They suggest that this moderated the effects of threat on recall though do not provide evidence to support this assertion. Furthermore, as discussed previously it is difficult to ascribe post-test reports of affect solely to the effects of threat.

Physiological arousal

Hess, Hinson et al. (2009) recorded Skin Conductance Response (SCR) as a measure of psychophysiological arousal reporting that although mean SCR was significantly greater in the threat than the non-threat condition, there was no evidence linking this with performance.

Motivational mechanism

The studies in this review investigated motivation via performance expectations and memory self-efficacy; however, effort which is linked to this mechanism in the wider stereotype threat literature was not considered in these papers.

Memory self-efficacy

Andreoletti & Lachman (2004) found no main effects of age or threat but a significant effect of education on memory self-efficacy ratings. They concluded that although education impacted positively on memory self-efficacy, such beliefs were unrelated to threat and therefore did not account for stereotype threat effects on performance.

Expanding on this 1-item measure, Desrichard & Köpetz (2005; study 1) used a 10-item Memory Self-Efficacy (MSE) Questionnaire devised by Berry, West & Dennehey (1989). They found that threat had a greater effect on performance on their errand running task when MSE was low, concluding that the relationship between threat and performance was moderated by memory self-efficacy. However, it is difficult to place much weight on this study as they counterbalanced the order in which participants completed the self-efficacy questionnaire and the experimental

task. Ratings may therefore, in half their sample, have been more affected by actual performance than threat induction.

Conversely, Chasteen et al. (2005; studies 2 & 3), using a 5-item memory self-efficacy questionnaire, did find an effect of threat on self-efficacy ratings but this generalised across age groups. Furthermore self-efficacy did not influence memory performance and thus could not mediate the relationship between threat and performance on either recall or recognition tasks.

Performance expectations

Hess et al. (2004), operationalising performance expectations as recall predictions, found that threat did not affect predicted recall in either of their studies. Desrichard & Köpetz (2005; study 2) on the other hand found evidence of mediation with threat producing lower performance expectations which led to reduced scores on memory tests. Similarly Hess, Hinson et al., (2009) found that in their young-old participants the moderating effect of education on recall was itself mediated by performance expectations as, under conditions of threat, participants with more education had lower performance expectations.

Discussion

This review comprised two parts; the first examined 17 studies to ascertain the effects of stereotype threat on older adults' memory. The studies were grouped according to their investigation of recall or recognition performance. The second part further analysed 13 of these studies to consider factors which might explain such an

effect. Studies were grouped according to the particular variables they explored, and these were grouped conceptually into participant characteristics and threat based effects. Several issues of interest arose from the review.

Does stereotype threat affect older adults' memory performance?

Abrams et al., (2006) and Rahhal et al., (2001) found evidence for stereotype threat effects on composite measures of cognitive performance however it was not clear which aspect of cognition explained these findings. In terms of recall performance, with the exception of the study by Hess et al. (2003), it was difficult to find robust evidence to support the claim that stereotype threat had any direct impact, and two reliable papers provided convincing arguments against such an effect (Hess et al., 2004; Horton et al., 2010). No evidence was found for the effect of threat on recognition. One possible reason for this subtle difference in findings is that recall tasks are widely known to involve greater cognitive resources than recognition tasks and as people grow older these resources are depleted. As a consequence older people naturally perform more poorly on recall than on recognition tasks (Hess, Emery et al., 2009; Craik & McDowd, 1987). Performance on recognition tasks is therefore less likely to be disturbed by stereotype threat.

Against this backdrop of relatively weak evidence in support of the predicted stereotype threat effect on older adults' memory performance, Hess, Hinson et al. (2009) and Hess & Hinson (2006) reported that those in the early years of older adulthood were most susceptible to threat effects, whilst Andreoletti & Lachman (2004) and Chasteen et al. (2005) found that stereotype threat affected all age groups on tests of recall and recognition.

Age group identification

The findings above suggest that susceptibility to threat does not necessarily relate directly to chronological age. This poses conceptual difficulties in defining a group of individuals who might therefore be affected by ageing stereotypes and arouses curiosity regarding age-group identification. The theory predicts that older adults who strongly identify with their age group should be more susceptible to ageing stereotypes. However, Kang & Chasteen (2009) directly measured age group identification and although they reported an inverse relationship between identification and performance their results were not significant and their methodology weak.

The majority of research in this area takes chronological age to imply susceptibility to ageing stereotypes. However, compared to groups with more delineated boundaries e.g. sex or ethnicity, those with more ambiguous limits such as age groups may respond differently to manipulations that invoke comparisons with other groups (O'Brien & Hummert, 2006). Thus it may be that, with respect to group identification, old age does not fit well into the stereotype threat model. In other areas of research identification or membership is more distinct and there is generally no scope for movement between the stereotyped and non-stereotyped group. Old age however is an identity which most of us will someday assume. Thus it is possible that identification with old age represents not only current belonging but also anticipated future belonging to that group. This may explain why the effects of an age-related stereotype threat were found by Andreoletti & Lachman (2004) and Chasteen et al. (2005) to generalise across age groups. In relation to the findings of Hess, Hinson et al. (2009) and Hess & Hinson (2006) who report that their oldest participants were least affected by threat it may be that younger members of the

older adult group exhibit concerns about imminent decline, whereas older members may experience a sense of comfort in knowing that they have thus far escaped such decline. Alternatively it may be that for the young-old their newly acquired age-related identity is more prominent in their minds, for example they may be recently retired, have started to receive a pension, or be eligible for new age-related services.

Chronological age evidently serves only as a crude approximation for age group identification, and whilst direct measures of this construct as employed by Kang & Chasteen (2009) have yielded encourageing results there is clearly scope for further research in this area. Finally, whilst Abrams et al. (2006) suggest measures such as intergroup contact may reduce effects of threat, possibly through weakening age group identification, no other studies have measured this.

Identification with the domain of memory

According to Steele (1995) identifying with the stereotyped domain is linked to heightened experience of threat and impaired performance in that domain. However, direct measures of domain identification/importance of memory yielded weak evidence in support of this.

Using education as a proxy for importance placed on cognitive abilities revealed that higher education enabled individuals to benefit from positive information about ageing and memory whereas those with less education found any information pertaining to ageing and memory, whether positive or negative, detrimental to performance (Andreoletti & Lachman, 2004). The findings of Hess et al. (2009) offer additional support for this effect. Given these results it appears clear that education has a role to play in determining susceptibility to threat however it

may be that education offers more than a representation of domain identification, perhaps tapping into constructs such as resilience to negative stereotypes. Future research may wish to investigate the role of education in more detail.

Individuals who believe themselves to be stigmatised in a particular domain are likely to be more susceptible to stereotype threat when performance in that domain is tested (Hess, Hinson et al., 2009). Awareness of stigma regarding memory and ageing is therefore likely to influence the extent to which threat affects older adults' performance on memory tests. There is good evidence for this assertion with Kang & Chasteen (2009) and Hess, Hinson et al. (2009) both finding evidence for stigma consciousness as a moderator. Although conceptualised here within the realm of domain identification it could also be argued that stigma consciousness relies equally on the presence of group identification. The content validity of questionnaires used in this area should be further explored to elucidate the precise constructs they measure.

Mechanisms of stereotype threat

Threat based effects were examined to investigate a number of mechanisms which might account for observed performance decrements under conditions of threat.

Reduction in cognitive resources

Task demands were reported to moderate the effects of threat on performance however these findings were deemed flawed due to statistical methods used (Hess, Emery et al., 2009). Strategy use was reported by Hess et al. (2003) to mediate the effect of threat on performance however subsequent studies failed to replicate this (Hess, Hinson et al., 2009; Hess & Hinson, 2006). Examining test-related anxiety as

a possible mediator yielded mixed results, due in part to the range of methods and tools used to measure this construct. Any evidence which was found for anxiety as a mediator is likely to be unreliable by virtue of the use of post-test measures (Hess & Hinson, 2006) which may be more affected by perceptions of performance than threat induction; and the use of retrospective reports (Abrams et al., 2006) which are not necessarily reliable indicators of emotional states. A more suitable suggestion could be to use pre- and post-induction measures to ascertain how much test-related anxiety the threat induction itself induces. No evidence was found for evaluation apprehension or disruption in working memory explaining the effects of threat on memory. In terms of evaluating this mechanism the overall supporting evidence is scarce, and that which exists is, for the most part, methodologically flawed. It is thus unlikely that a reduction in cognitive resources is the mechanism by which stereotype threat takes effect. Previous research which supports the existence of such a mechanism has been conducted in relation to gender and racial stereotypes thus it may be something about older adults and their natural age-related variation in cognitive resources which makes this mechanism particularly difficult to evaluate in this population.

Affective response mechanism

Investigating mediators or moderators linked to affective states generated no evidence for general anxiety as measured by the STAI-S, negative affect as measured by the PANAS, or physiological arousal as measured by skin conductance response. It therefore doesn't seem likely that affective responses explain the mechanism of stereotype threat. This is in keeping with the findings of Smith's (2004) wider review of mediational variables in stereotype threat across different populations.

Motivational mechanism

Stereotype threat lowers expectations about performance and self-efficacy (Spencer, Steele and Quinn, 1998) and low expectations can negatively affect motivation and effort thus having a detrimental impact on performance (Bandura, 1977). The idea that threat may impact on feelings of self-efficacy in relation to a specific task and, in an almost self-fulfilling way, disrupt performance is therefore theoretically appealing. However, with the exception of Desrichard & Köpetz (2005) investigations into this proposed mechanism have yielded mostly null findings. One explanation may be that beliefs of self-efficacy are relatively unimportant in determining stereotype threat effects as it is not necessary to believe the stereotype or be concerned it applies to oneself; it is the mere awareness that others may believe it to be true and treat you differently as a result (Steele, 1995). In relation to performance expectations the findings of this review present more encourageing findings with two of the four studies investigating this variable reporting that their threat manipulation led to poorer performance via lower expectations (Hess, Hinson et al., 2009; Desrichard & Köpetz, 2005). In one of these studies this effect was moderated by education (Hess, Hinson et al., 2009). On balance the evidence presented here suggests that a motivational mechanism as indicated by performance expectations may account for some of the performance decrements under conditions of threat. Findings from the wider stereotype threat research however have failed to find statistical evidence for effort as a mediator.

Conclusion

Stereotype threat theory implies that negative stereotypes about ageing and memory should lead to significant performance decrements on tests of cognitive ability. A review of the literature found limited evidence for such an effect on tests of recall and no evidence for an effect on tests of recognition. Difficulties with research in this area were highlighted; namely the apparent problem with identifying those vulnerable to age-related stereotypes given the weak link between age and age-group identification, and the confounding issue of controlling for natural age-related cognitive decline. Education was highlighted as an important factor for consideration in this field of research, and possible evidence for a role of a motivational mechanism was found.

References

Abrams, D., Eller, A., & Bryant, J. (2006). An age apart: The effects of intergenerational contact and stereotype threat on performance and intergroup bias. *Psychology and Aging*, *21*(4), 691-702.

Andreoletti, C., & Lachman, M. (2004). Susceptibility and resilience to memory aging stereotypes: Education matters more than age. *Experimental Aging Research*, *30*, 129–148.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191-215.

Baumeister, R. F. & Showers, C. J. (1986) A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology, 16*, 361-383*.*

Berry, J. M., West, R. L. & Dennehey, D. M. (1989). Reliability and validity of the memory self-efficacy questionnaire. *Developmental Psychology, 5,* 701-713.

Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L. & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology, 33,* 267-285.

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why Do Women Underperform Under Stereotype Threat?: Evidence for the Role of Negative Thinking. *Psychological Science*, *16*(7), 572–578.

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.

Craik, F. I. M. & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology. Learning Memory and Cognition, 13(3),* 474-479.

Dixon, R. A. & Hultsch, D. F. (1984). The Metamemory in Adulthood (MIA) instrument. *Psychological Documents, 14,* 3.

Desrichard, O., & Köpetz, C. (2005). A threat in the elder: The impact of task-instructions, self-efficacy and performance expectations on memory performance in the elderly. *European Journal of Social Psychology*, *35*, 537–552.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology*, *56*(1), 5-18.

Hess, T. M., Auman, C., Colcombe, S. J., & Rahhal, T. A. (2003). The impact of stereotype threat on age differences in memory performance. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *58B*(1)*,* 3-11.

Hess, T. M., Emery, L., & Queen, T. L. (2009). Task demands moderate stereotype threat effects on memory performance. *The Journals of Gerontology: Psychological Sciences*, *64B*(4), 482-486.

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., & Hodges, E. A. (2009). Moderators of and mechanisms underlying stereotype threat effects on older adults' memory performance. *Experimental Aging Research*, *35*, 153-177.

Hess, T. M., Hinson, J. T., & Statham, J. A. (2004). Explicit and implicit stereotype activation effects on memory: Do age and awareness moderate the impact of priming? *Psychology and Aging*, *19*(3), 495-505.

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.

Hogg, M. A. (2006). Social identity theory. In P. J. Burke (Ed.), *Contemporary social psychological theories* (pp. 111-136). Palo Alto, CA: Stanford University Press.

Horton, S., Baker, J., Pearce, G. W., & Deakin, J. M. (2008). On the malleability of performance: Implications for seniors. Journal of *Applied Gerontology, 27,* 446-465.

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.

Kang, S. K., & Chasteen, A. L. (2009). The moderating role of age-group identification and perceived threat on stereotype threat among older adults. *The International Journal of Aging and Human Development, 69,* 201–220.

Kit, K. A., Tuokko, H. A. & Mateer, C. A. (2008). A review of the stereotype threat literature and its application in a neurological population. *Neuropsychology Review, 18,* 132-148.

Kite, M. E. & Johnson, B.T. (1988). Attitudes toward older and younger adults: A meta-analysis. *Psychology and Aging, 3*(3), 233-244.

Lachman, M. E., Bandura, M., Weaver, S. L. & Elliott, E. (1995). Assessing memory control beliefs: The Memory Controllability Inventory. *Aging and Cognition, 2,* 67-84.

Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology, 71*(6), 1092-1107.

Leyens, J-P., Desert, M., Croizet, J-C & Darcis, C. (2000). Stereotype Threat: Are lower status and history of stigmatization preconditions of stereotype threat? *Personality and Social Psychology Bulletin, 26,* 1189-1199.

McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology*, *27*(5), 486–493.

Neupert, S. D., Miller, L. M. S., & Lachman, M. E. (2006). Physiological reactivity to cognitive stressors: Variations by age and socioeconomic status. *The International Journal of Aging and Human Development*, *62*(3), 221–235.

O'Brien, L. T., & Hummert, M. L. (2006). Memory performance of late middle–Aged Adults: Contrasting Self–Stereotyping and Stereotype Threat Accounts of Assimilation to Age Stereotypes. *Social Cognition*, *24*(3), 338–358.

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*(4), 697-706.

Roenker, D. L., Thompson, C. P. & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin, 76,* 45-48.

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38,* 194-201.

Schmader, T. & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*(3), 440-452.

Smith, J. (2004). Understanding the process of stereotype threat: A review of the mediational variables and new performance goal directions. *Educational Psychology Review, 16*(3), 177-206.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4–28.

Spielberger, C. D., Gorusch, R. L. & Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory.* Palo Alto, California: Consulting Psychologists Press.

Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *American Psychologist, 52* (6), 613-629.

Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69* (5), 797-811.

Stein, R., Blanchard-Fields, F. & Hertzog, C. (2002). The effects of age-stereotype priming on the memory performance of older adults. *Experimental Aging Research, 28,* 169-181.

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, 54*, 1063-1070.

Wheeler, S. C. & Petty, R. E. (2001). The effects of stereotype activation on behaviour: A review of possible mechanisms. *Psychological Bulletin, 127*(6), 797-826.

Wright, C. E., Kunz-Ebrecht, S. R., Iliffe, S., Foese, O. & Steptoe, A. (2005). Psychological correlates of cognitive functioning in an elderly population. *Psychoneuroendocrinology, 30*(9), 826-838.

Part 2: Empirical Paper

The impact of stereotype threat on older adults' performance

on a diagnostic test for dementia

Abstract

Aims: In keeping with the stereotype threat framework this study aimed to show that activating a negative stereotype about ageing and memory would impair older adults' performance on memory tests. Extending this idea further it was hypothesised that labelling the tests as diagnostic for dementia would lead to even greater performance impairments.

Method: Participants were stratified by gender and level of education and assigned randomly to one of three groups: no threat, age stereotype threat, and dementia diagnosis threat, before completing recall, recognition and dementia screening tests.

Results: Negative age stereotypes had a detrimental impact on cognitive performance and state anxiety but only when individuals believed the threat induction. The threat of a dementia diagnosis was associated with higher levels of anxiety. However performance on recall, recognition and dementia screening tests was not impaired. Younger age and higher level of education predicted performance independent of threat induction.

Conclusions: This study provides some support for the impact of negative age stereotypes on cognitive performance however rejects the hypothesis that labelling tests as diagnostic for dementia should impair performance to a greater extent.

Introduction

Stereotypes and behaviour

Negative group stereotypes are plentiful in our society. They are often unconscious; perpetuated by the media, families, education systems and countless other social or environmental influences (Kit, Tuokko & Mateer, 2008; Scholl & Sabat, 2008; Steele, 1997). Research has shown that negative stereotypes have the potential to cause psychological distress to group members leading to altered behaviour and performance decrements in the stereotyped ability (Steele & Aronson, 1995; Steele, 1997). These authors explained this phenomenon as something called 'stereotype threat', defined as the situational pressure and emotional distress an individual experiences when they become aware of potentially confirming as self-characteristic a negative stereotype about a group to which they belong, and the fear of being judged and treated according to this stereotype. Stereotype threat effects were first reported by Steele & Aronson (1995) who observed a decline in the academic performance of Black students when a negative stereotype about their racial group's academic ability was made salient. In other areas of research it is reported that stereotype threat has similar negative effects on women taking maths tests (Spencer, Steele & Quinn, 1999), men's performance on affective tasks (Leyens, Desert, Croizet & Darcis, 2000) and older people's cognitive performance (Horton, Baker, Pearce & Deakin, 2008).

The effects of stereotype threat are thought to be greatest when an individual strongly identifies with the stereotyped group e.g. their race is a central part of their self-identity; they are highly invested in the performance domain e.g. academic ability is important to them; and when test diagnosticity is high e.g. an IQ test (Steele, 1997; Hess, Auman, Colcombe & Rahhal, 2003; Hess & Hinson, 2006).

According to stereotype threat theory, emotional reaction to threat is one mechanism which can potentially explain the effects of stereotype on performance (Steele, 1997; Steele & Aronson, 1995). In the stereotype threat literature physiological indicators and self-reported anxiety have been implicated as mediating the effect of threat on performance (Scholl & Sabat, 2008).

Ageing & stereotype threat

In Western culture ageing is viewed as a process of inevitable cognitive decline, and as such older adults are routinely subjected to negative stereotypes regarding their mental competence (Kite & Johnson, 1988; Stein, Blanchard-Fields & Hertzog, 2002). Consequently older adults fall prey to the stereotype that they have a deteriorating memory, are forgetful and less competent than their younger counterparts (Kit et al., 2008; Chasteen, Bhattacharyya, Horhota, Tam & Hasher, 2005). Older adults themselves appear to anticipate such decline and report less control over memory function (Chasteen et al., 2005). In relation to older adults' performance on memory tests, stereotype threat theory would predict that those who value their memory ability the most and who identify strongly with their age group are likely to be most vulnerable to threat effects (Wheeler & Petty, 2001; Kang & Chasteen, 2009). Such effects are thought to be greatest when test diagnosticity is high (Hess & Hinson, 2006; Steele & Aronson, 1995) and may be explained by an emotional response to threat which interferes with performance (Steele & Aronson, 1995; Steele, 1997). The role of age, education, test diagnosticity and emotional arousal in relation to ageing and stereotype threat are discussed below.

The role of age

The research pertaining to stereotype threat effects in older adults has reported significant performance decrements on tests of memory for example, story and word list recall (Kang & Chasteen, 2009; Desrichard & Köpetz, 2005; Hess et al., 2003) though lesser impact is reported on recognition tasks (Hess, Emery & Queen, 2009). Although these studies provide some evidence for direct effects of stereotype threat on memory performance, other findings indicate that threat impairs performance via an interaction with other variables such as age (Hess & Hinson, 2006; Hess, Hinson & Hodges, 2009). These authors suggest that based on their findings stereotype threat is likely to be most salient for those in their late 60s. It is also possible that age affects performance independently of threat manipulations, for example Horton, Baker, Pearce & Deakin (2010) found no evidence of stereotype threat in their study however report that recall performance deteriorated significantly as age increased.

The role of education

There is also evidence that threat impacts on performance via an interaction with education. Older adults with higher levels of education might be assumed to have greater investment in cognitive abilities which, according to stereotype threat theory, would imply that more educated individuals should be most susceptible to stereotype threat (Steele & Aronson, 1995; Horton et al., 2010). However research by Andreoletti & Lachman (2004) and Hess, Hinson et al. (2009) revealed significant effects in the opposite direction than the theory would predict, such that lower education was associated with greater vulnerability to stereotype threat. Their findings suggest that education may provide a buffer against negative age stereotypes. Horton et al. (2010) speculate that their failure to demonstrate an effect

of threat on performance may be due to the highly educated sample used in their study, suggesting that education offers some immunity to such effects.

Test diagnosticity

The stereotype threat framework proposes that the detrimental impact of stereotypes on performance is greatest when test diagnosticity is high (Hess & Hinson, 2006; Steele & Aronson, 1995). This hypothesis is well supported in the older adult literature with healthy individuals exhibiting greater performance deficits when told that tests are diagnostic of memory ability (Kang & Chasteen, 2009; Desrichard & Köpetz, 2005; Chasteen et al., 2005; Rahhal, Hasher & Colcombe, 2001).

Emotional arousal

It has been argued that threat increases anxiety and negative affect which may impair performance through greater levels of emotional arousal (Steele & Aronson, 1995; Hess, Hinson et al., 2009). Others have claimed that negative affect is reported less often as individuals age (Charles, Reynolds & Gatz, 2001) and thus an emotional response to threat in older adults may be more likely to be represented by a reduction in positive affect (Kang & Chasteen, 2009). Despite the assumed centrality of emotions, there is little evidence within the older adult literature for the role of anxiety or negative affect mediating the effects of stereotype threat on performance. This may be due in part to measurement issues with some studies recording mood and state anxiety prior to testing (Hess, Hinson et al., 2009; Hess & Hinson, 2006; Chasteen et al., 2005; Hess, Hinson & Statham, 2004); others measuring mood post testing (Hess, Emery et al., 2009); and one study which measured mood pre- and post-testing to calculate change in affect (Kang &

Chasteen, 2009). A more reliable approach might be to measure emotional states during testing. Another issue pertaining to the investigation of the relationship between stereotype threat and state anxiety in previous research is the failure to take into account individual differences in trait anxiety. This construct is characterised as the general disposition to experience transient states of anxiety, thus in the face of threat an individual with high trait anxiety might be more vulnerable to state anxiety (Tovilović, Novović, Mihić & Jovanović, 2009).

Ageing and dementia

The term dementia refers to a collection of symptoms including memory loss and cognitive decline, for example a reduction in reasoning and communication skills that impair everyday function (Dementia UK, 2012). It is a common feature of old age with one in three over 65s expected to develop the disease. The current UK prevalence of dementia is estimated to be about 820,000, though given our ageing population this is expected to rise to 1,700,000 by 2051 (Alzheimer's Society, 2012). Although there is significant evidence that cognitive dysfunction in dementia can be attributed to neuropathology in the brain it has been argued that psychosocial factors also play an important role in explaining older adults' performance on neuropsychological tests (Scholl & Sabat 2008). These authors question whether observed performance decrements on such tests reflect the effects of psychosocial factors such as stereotype threat over and above the neuropathology of dementia; however this hypothesis has not yet been tested. A similar line of thinking has queried the effects of stereotype threat on neurological populations (Kit et al., 2008) and some preliminary research has been carried out in this area (Suhr & Gunstad, 2002, 2005). These authors examined the effect of threat on cognitive performance

in individuals with a history of mild head injury and found that exposure to information that highlighted their 'diagnosis' and emphasised the potential for poor cognitive performance in their diagnostic group led to poorer test performance. They use the term 'diagnosis threat' to describe this phenomenon, suggesting the potential for such an effect to negatively impact on the neuropsychological test performance of clinical populations.

Clinical relevance

Given the prevalence of dementia, the prospect of a positive diagnosis can cause great anxiety in many older adults who may focus on innocuous slips of memory as signs of the disease (Scholl & Sabat, 2008). With test diagnosticity purported to increase the effect of stereotype threat it is likely that this phenomenon has some impact on performance on neuropsychological tests for dementia thereby increasing the likelihood of a diagnosis. This study has potential implications for the way in which such tests are introduced and administered in clinical settings.

Aims of the study

There is well documented research on the effects of stereotype threat on older adults' performance on memory tests. Also appearing in the literature is a concern that this phenomenon extends to neurological populations (Suhr & Gunstad, 2002, 2005; Kit et al., 2008; Scholl & Sabat, 2008). Given that diagnosticity increases stereotype threat (Hess et al., 2003; Hess & Hinson, 2006) it was hypothesised that a diagnostic test for dementia in an ageing population would have a compounded effect of activating an age-related stereotype threat and increasing fear of positive

diagnosis. In light of previous research findings the impact of age and education on performance were examined and the role of anxiety during testing was explored.

This study manipulated stereotype threat and test diagnosticity to create three experimental groups: no threat, age stereotype threat, dementia diagnosis threat, and examines performance on free recall, recognition, and a dementia screening test.

The following pattern of results was hypothesised:

1. Recall, recognition and dementia screening test scores will decrease in a linear fashion as a function of threat: no threat condition > age stereotype threat condition > dementia diagnosis threat condition.

2. Controlling for individual differences in trait anxiety, state anxiety will increase in a linear fashion as a function of threat: dementia diagnosis threat condition > age stereotype threat condition > no threat condition. Negative affect will follow a similar pattern, with positive affect changing in the opposite direction.

3. Higher state anxiety and negative affect will correlate with poorer recall and recognition scores and lower scores on a dementia screening test.

4. The effects of age and education on performance will be examined as an exploratory measure.

Method

Participants

The sample for this study was 60 older adults from the Greater London area, recruited via community or educational groups for seniors. There was no financial

incentive for taking part in the study. It has been suggested from previous research (Hess & Hinson, 2006; Hess, Hinson et al., 2009) that stereotype threat is most likely to exert an effect on individuals in the earlier years of older adulthood, therefore individuals aged 65 – 70 were the target for recruitment. A cut-off age of 70 was proposed to reduce the incidence of participants naturally meeting diagnostic criteria for dementia as this would potentially contaminate data. Due to some difficulty finding sufficient participants in this age range three participants above this age were included. The age range of participants was therefore 65-72 (M = 67.38, SD = 1.88). The 60 participants consisted of 39 women and 21 men. This was a cognitively able sample with 78% (47) having proceeded to further education. The remaining 22% (13) had not progressed beyond secondary school. Participants were considered eligible for the study if they had no history of psychiatric illness or neurological conditions. They were required to be fluent English speakers, to have no difficulty with reading or writing, and no visual or hearing impairments that would affect their ability to complete the research tasks. It was also required that they were not currently engaged in a primary caring role for someone with dementia.

Design

A randomised experimental design was used in this study. There was one factor with three levels – no threat, age stereotype threat, and dementia diagnosis threat. Participants were stratified by gender and level of education and assigned randomly to one of the three experimental groups. Statistical analyses were conducted using SPSS.

In terms of power analysis, a study by Suhr and Gunstad (2002) was found to be the most relevant in calculating the sample size needed for the present

research. They administered the Auditory Verbal Learning Test (AVLT) to participants with a history of head injury and found an effect size of $\eta^2 = 0.2$. Taking this as an indicator of the potential effects of threat on memory performance, power calculation was carried out using G*Power 3 (Faul, Erdfelder, Lang and Buchner, 2007). Assuming equal group sizes and specifying alpha = 5% and desired power = 80% the required sample size was estimated at 42. It was deemed sensible in this circumstance to take a more conservative estimate while maintaining practical feasibility, so a total sample size of 60 was decided allowing 20 participants per group.

Materials

Measures of affective responses

The State-Trait Anxiety Inventory (STAI) was developed in 1970 by Spielberger, Gorsuch and Lushene. It is a self report measure which takes approximately 10 minutes to complete and comprises a 20 item state anxiety scale that evaluates on a four point scale (not at all, somewhat, moderately so, very much so) the intensity of how respondents feel "right now, at this moment" (e.g. 'I feel frightened', 'I am jittery'); and a 20 item trait anxiety scale which assesses on a four point scale (almost never, sometimes, often, almost always) the frequency of how respondents feel "generally" (e.g. 'I lack self-confidence', 'I feel nervous and restless'). Scores range from 20 to 80 on each scale. The alpha coefficients, based on a sample of working adults, for both state and trait anxiety scales are high at 0.93 and 0.91 respectively. Furthermore, the STAI is reported to have excellent psychometric

properties for the assessment of anxiety in elderly populations (Spielberger, Gorusch, Lushene, Vagg & Jacobs, 1983).

The Positive and Negative Affect Schedule - Extended (PANAS-X) is an extension of the PANAS developed by Watson, Clark, and Tellegen (1988). It is a 60 item measure which contains two 10-item higher order scales for Positive Affect (e.g. excited, enthusiastic, attentive) and Negative Affect (e.g. afraid, irritable, hostile) as well as 11 lower order scales for specific affects: Fear, Sadness, Guilt, Hostility, Shyness, Fatigue, Surprise, Joviality, Self-Assurance, Attentiveness, and Serenity. Participants are asked to rate the extent to which they feel each of the 60 emotions listed on a five point scale where 1 = very slightly/not at all; 2 = a little; 3 = moderately; 4 = quite a bit; 5 = extremely. The PANAS-X can be completed in 10 minutes and can be administered in relation to different time frames. The protocol used in this study instructed participants to rate their feelings "at this moment". For this temporal instruction the internal consistency reliabilities (Cronbach's coefficient alpha) for the two higher order scales is reported to be between 0.83 and 0.91, with the internal reliability of specific affects ranging from 0.72 - 0.93.

Measures of free recall and recognition

The Auditory-Verbal Learning Test (AVLT), developed by Rey (1964) is a six trial list learning test comprising of two separate lists of 15 words. It can be used with individuals aged 7-89 and is routinely administered in 10-15 minutes although a further period of 20 minutes is required for the delayed recall (Ivnik, Malec, Tangalos, Petersen, Kokmen & Kurland, 1990). Instructions, in the absence of a test manual, were taken from Lezak (1995). The psychometric properties of the AVLT are good with reported reliability coefficients above 0.60 (Mitrushina, Boone & D'Elia, 1999). It generates scores for a number of measures including word list recall

over five learning trials (scored out of 75) and recognition (scored out of 30) and is thus useful in evaluating verbal memory.

Measure of cognitive functioning

The Addenbrooke's Cognitive Examination – Revised (ACE-R) is a modified version of the ACE developed by Mathuranath, Nestor, Berrios, Rakowicz and Hodges (2000). It is a brief cognitive test battery for dementia that assesses five cognitive domains including memory. It is reported to be easy to use and is acceptable to patients, with an administration time of 12-20 minutes. It has excellent diagnostic accuracy and is widely used in clinical practice (Larner, 2007). It is scored out of 100 and there are two commonly used clinical cut-offs. A score of <88 gives 94% sensitivity and 89% specificity for dementia, whereas a score of <82 gives 84% sensitivity and 100% specificity for dementia (Mioshi, Dawson, Mitchell, Arnold & Hodges, 2006). For the purposes of this study the full scale score and memory subscale score were utilised.

Procedures

Requests for participants were circulated by email to members of educational or community groups in the Greater London area. Interested individuals were invited to contact the researcher by telephone or email. A brief screening questionnaire, based on the inclusion/exclusion criteria, was conducted by telephone and some basic demographic information was obtained. Participants were then invited to attend a 1 hour appointment at University College London.

Upon arrival at the university participants were given an information sheet explaining what the tests would involve and outlining their rights as a participant

(See Appendix A). The sheet contained information about the aims of the study thus inducing one of the three experimental conditions. This information was also transmitted verbally to ensure that the desired induction occurred.

Participants in the neutral/no threat condition were told:

> *'The aim of this study is to examine individual differences in performance on a range of tests for attention, memory, and language skills. Research has shown that people vary a great deal in these abilities and that performance on these tests is unrelated to age. For example the strategies we use for learning lists of words tend to stay stable throughout the lifespan. The tests I will carry out today will help me to find out a bit more about this'*

Participants in the age stereotype threat condition were told:

> *'The aim of this study is to examine performance on a range of tests for attention, memory, and language skills. Research has shown that the normal ageing process can have a negative impact on these abilities. For example as we get older, from around age 65, we find it more difficult to remember details of recent events or to learn new information. The tests I will carry out today will help me to find out a bit more about this'*

Participants in the dementia diagnosis threat condition were told:

> *'The aim of this study is to examine performance on a range of tests for attention, memory, and language skills. Research has shown that as people age they perform worse on such tests. One possible reason for this could be age-related illnesses such as dementia. The term dementia refers to a collection of symptoms including memory loss and cognitive decline e.g. a reduction in reasoning and communication skills. According to the Alzheimer's Society there are about 750,000 people in the UK with dementia. It most commonly occurs in people aged 65 or over. The tests I will carry out today will help me to find out a bit more about this'*

Having read the information sheet fully, participants were then asked to read and sign the consent form (See Appendix B). Participants were asked to complete the

first of the two mood measures, the PANAS, being instructed to rate the extent to which each of the 60 items reflected how they were feeling at the present moment. Introducing the AVLT the researcher again reiterated the experimental induction prior to testing participants on free recall. State and trait anxiety was measured next using the STAI. The ACE-R was then administered followed by AVLT recognition. Participants were debriefed and fully informed about the rationale for the study and an explanation of the different experimental conditions was given (See Appendix C; see below for further discussion). Finally participants were given their test results from the ACE-R and AVLT.

Ethics

Ethical approval for this study was obtained from the University College London Research Ethics Committee (See Appendix D). Ethical considerations were noted prior to undertaking this research, for example administering dementia tests to an elderly population gives rise to the possibility that some participants may score within the clinical range. In the case of two participants there were some clinical concerns which the researcher discussed fully with them and advised them to see their GP for further investigation. It was also thought possible that some participants may be distressed by the nature of the tests. Appropriate measures were thus taken to ensure participants' safety and wellbeing e.g. making sure they were aware of what their participation involved, that they did not feel pressurised into participating or suffer any harm or unnecessary demands as a result of their taking part. As participants were given different rationales for the study a comprehensive information sheet was provided to all participants upon completion of the testing. This explained the concept of stereotype threat, further explained the focus of the

study, and clarified the reasons for giving people different information prior to testing (See Appendix C). Provisions were made for people to contact the researcher should any issues or concerns arise following their participation in the study.

Results

Preliminary analysis

In total 61 participants were tested. One participant was removed from the analysis as her ACE-R score was in the clinical range. Twenty participants were in each of the no threat, age stereotype threat and diagnosis threat conditions. No significant differences were found between the groups in terms of participants' age, level of education or gender. There were no systematic age differences between males (M = 67.90, SD = 2.119) and females (M = 67.10, SD = 1.698).

Table 1

Demographic variables for participants in each condition (N = 60)

	Condition		
	No threat (N=20)	Age stereotype threat (N=20)	Dementia diagnosis threat (N=20)
Age (mean, SD)	67.05 (2.01)	67.70 (1.98)	67.38 (1.88)
Education			
Low	5	3	5
High	15	17	15
Gender			
Male	7	7	7
Female	13	13	13

The role of age and education

Independent of experimental condition, the effect of education on performance was examined using one way ANOVAs. A main effect of education was found on AVLT recall F (1, 59) = 17.793, p = .000, η^2 = .24 and recognition F (1, 59) = 7.742, p = .007, η^2 = .12 such that those with higher levels of education did significantly better than those with lower education. No effects were found on the ACE-R memory subscale or total score.

A correlation analysis was carried out to examine the relationship between age and performance on the AVLT and ACE-R. Age correlated negatively with total ACE-R score r(60) = -.267, p = .039, ACE-R memory r(60) = -.278, p = .032, **AVLT recall r(60) = -.383, p = .003**, and **AVLT recognition r(60) = -.489, p = .000**. The results in bold are those which survived Bonferroni correction to a more conservative alpha level (p = .0125) which controls for Type I error in the case of multiple comparisons.

In light of the significant associations between AVLT recall and recognition scores with participants' age and education a multiple regression analysis was used to test if age and education significantly predicted recall and recognition scores. The results of the regression indicated that for recall age and education explained 33.9% of the variance (R^2 = .339, F(2, 59) = 14.61, p = .000). It was found that age independently predicted recall (β = -.326, t(58) = -2.997, p = .004), as did education (β = .442, t(58) = 4.069, p = .000). For recognition age and education explained 31.8% of the variance (R^2 = .318, F(2, 59) = 13.314, p = .000). It was found that age independently predicted recognition (β = .284, t(58) = -4.096, p = .000), as did education (β = -.452, t(58) = 2.576, p = .013).

Emotional arousal and performance

It was hypothesised that higher state anxiety as measured by the STAI, and negative affect as measured by the PANAS would correlate negatively with scores on the ACE-R and the AVLT. The results however show that neither state anxiety nor negative affect showed significant correlations with any of the performance measures (Table 2).

Table 2

Correlations between emotional arousal and performance (N = 60)

	Negative affect	State anxiety	AVLT Recall	AVLT Recog.	ACE-R Full	ACE-R Memory
Negative affect	1					
State anxiety	.563**	1				
AVLT Recall	-.107	-.009	1			
AVLT Recog.	-.058	-.059	.677**	1		
ACE-R Full	-.015	.038	.513**	.601**	1	
ACE-R Memory	-.110	-.020	.396**	.522**	.865**	1

**p < .01 level

Stereotype threat effects on state anxiety and mood

A second hypothesis was that self-reported positive and negative affect measured immediately following threat induction would vary as a function of threat with negative affect being greater, and positive affect being lower, in the age stereotype and dementia diagnosis threat conditions than in the no threat condition. ANOVA revealed no differences by group in either positive affect $F(2, 59) = .078$, $p = .925$, or negative affect $F(2, 59) = .266$, $p = .767$.

It was also hypothesised that controlling for variation in trait anxiety, state anxiety would be greater in the age stereotype and dementia diagnosis threat conditions than in the no threat condition. The overall trend in the data supported the hypothesis that higher levels of state anxiety would occur in conditions of greater threat: no threat (M = 32.2; SD = 7.911); age stereotype threat (M = 35.75; SD = 8.996); dementia diagnosis threat (M = 37.4; SD = 8.768). As the data were normally distributed an ANCOVA [between-subjects factor: condition (no threat, age stereotype threat, dementia diagnosis threat); covariate: trait anxiety] revealed a significant main effect of trait anxiety on state anxiety, $F(1, 60) = 26.424$, $p = .000$, $\eta_p^2 = .329$, but no main effect of condition, $F(2, 60) = .507$, $p = .605$, $\eta_p^2 = .018$, and no interaction between condition and trait anxiety, $F(2, 60) = 1.262$, $p = .291$, $\eta_p^2 = .045$.

Stereotype threat effects on performance

The primary hypothesis was that recall, recognition and dementia screening test scores would vary as a function of threat with the highest scores expected in the no threat condition, followed by the age stereotype threat condition with the lowest scores expected in dementia diagnosis threat condition. As the data were normally

distributed this was investigated using a one way ANOVA. Variance in total ACE-R scores according to condition approached statistical significance with a moderate effect size, $F_{(2, 57)} = 3.069$, $p = .054$, $\eta^2 = .097$. Post-hoc analyses using Tukey's HSD indicated that this was due to significantly lower scores in the age stereotype threat condition compared to the dementia diagnosis threat condition ($p = .045$). No significant effect was found for condition on the ACE-R memory subscale or on AVLT recall and recognition. An overall trend in the data indicated that participants in the age stereotype threat condition performed worse than those in the no threat and dementia diagnosis threat conditions (Table 1) but when assessed using a MANOVA procedure this covariation was not found significant.

Table 3

Mean score (standard deviation) on each performance measure (N = 60)

| | Condition | | |
	No threat (N=20)	Age stereotype threat (N=20)	Dementia diagnosis threat (N=20)
AVLT Recall	58.8 (10.14)	56.8 (7.865)	58.25 (8.149)
AVLT Recognition	22.55 (4.989)	20.95 (3.456)	23.25 (2.425)
ACE-R Total	96.55 (2.856)	95.25 (3.041)	97.35 (2.134)
ACE-R Memory	24.25 (2.074)	23.3 (2.494)	24.65 (1.137)

It became apparent during testing that not all participants believed the experimental induction so a 7-point Likert rating scale was introduced to measure belief with 1 = strongly disbelieve and 7 = strongly believe. Forty-nine participants completed this

scale at the end of testing; filtering out those who disbelieved the induction further analysis was conducted on the remaining dataset (N = 40). Group sizes varied with eight in the no threat condition and sixteen in each of the other conditions. Significant differences between conditions were found for total ACE-R score, $F_{(2, 37)} = 4.167$, $p = .023$, $\eta^2 = .18$; ACE-R memory, $F_{(2, 37)} = 5.603$, $p = .007$, $\eta^2 = .23$; and AVLT recognition $F_{(2, 37)} = 4.035$, $p = .026$, $\eta^2 = .18$. Post-hoc analyses using Bonferroni reveal that this was due to participants in the age stereotype threat condition performing significantly worse than those in the no threat condition (total ACE-R score, $p = .025$; ACE-R memory, $p = .006$; and AVLT recognition $p = .026$). Recall scores did not differ significantly between conditions.

Table 4

Mean score (standard deviation) on each performance measure for individuals who believed the threat induction (N = 40)

	Condition		
	No threat (N=8)	Age stereotype threat (N=16)	Dementia diagnosis threat (N=16)
AVLT Recall	62.25 (7.61)	58.25 (7.55)	58.25 (8.903)
AVLT Recognition	25.13 (3.09)	21.56 (3.14)	23.31 (2.7)
ACE-R Total	98.5 (1.51)	95.5 (3.01)	97.13 (2.28)
ACE-R Memory	25.75 (0.7)	23.63 (1.93)	24.56 (1.21)

Discussion

This study demonstrated that negative stereotypes about ageing had a detrimental effect on cognitive performance for individuals in the age stereotype threat condition who believed the threat induction. Significant differences with large effects sizes were found on recognition, ACE-R total, and ACE-R memory scores such that those in the age stereotype threat condition were significantly lower than those in the non-threat condition. However, the hypothesis that labelling these tests as diagnostic for dementia should increase the effects of threat was rejected, with those in the dementia diagnosis condition performing consistently better than those in the age stereotype threat condition on these three dependent measures. In contrast to the recognition performance no effects were found for recall which is surprising given that previous research has reported that recall is more likely than recognition to be impaired as a function of threat (Hess, Emery et al., 2009).

Given the failure to demonstrate any performance decrements in the dementia diagnosis threat condition it is possible that the sample used in this study was somehow immune to the suggestion that their performance could be indicative of dementia. Recruiting from educational or community groups, these individuals represented a cognitively able cohort of older adults who were confident enough to volunteer for a memory study. It may be that this afforded them some protection against the relevance of a dementia diagnosis identity, thereby reducing the impact of threat. The literature on diagnosis threat utilises a population with known neurological histories and explicitly mentions membership of this diagnostic group to induce threat (Suhr & Gunstad, 2002, 2005). Although the threat induction clearly stated the prevalence of dementia in the over 65s, the fact that this study relied solely on the power of suggestion within a healthy sample of older adults may

account for the failure to demonstrate effects for the dementia diagnosis threat condition.

Another factor which may have influenced the results of this study is the incidence of individuals scoring at ceiling level on the ACE-R. The lack of sensitivity of this measure in this population makes it difficult to ascertain the presence of group differences.

Although the greatest performance decrements were not seen, as predicted in the dementia diagnosis threat condition, there was a trend towards these individuals reporting higher levels of state anxiety during testing. This finding possibly suggests that the threat of dementia diagnosis affected participants on an emotional level but this did not translate to performance decrements as expected, as state anxiety did not correlate with any of the performance measures.

Interestingly there were no systematic differences between conditions on self-reported positive or negative affect immediately following threat induction. It may be that it takes some time for the induction to take effect, or that subsequent test performance relevant to the stereotyped domain, regardless of actual performance, is seen as confirming the stereotype. This may then influence self-reported anxiety as measured part way through testing.

Research has shown that stereotype threat effects tend to be greater in people who value the domain being assessed (Aronson et al., 1999; Hess et al., 2003), thus a highly educated sample might be assumed more susceptible to threat relating to cognitive ability. This study found some evidence for an effect in the opposite direction with education predicting higher scores on recall and recognition This is in keeping with the findings of Andreoletti & Lachman (2004) and Hess, Hinson et al. (2009) and offers empirical support for the conjecture of Horton et al.

(2010) that a highly educated sample might offer some kind of immunity to stereotype threat effects.

Age also emerged as a significant predictor of recall and recognition performance, in keeping with the findings of Horton et al. (2010) who found similar age differences in recall. It is also fitting with the general consensus about the existence of age related decrements in cognitive ability (Zacks, Hasher & Li, 2000).

Clinical implications

This study aimed to demonstrate that a diagnostic test for dementia in an ageing population would impair memory performance over and above that which might be attributable to an age stereotype threat effect which would have potential implications for the way in which such tests are introduced and administered in clinical settings. The results of this research however indicated that, although individuals in the dementia diagnosis threat condition reported greater anxiety during testing, their performance was not impaired. There are a number of possible explanations for this. One is that there are particular profiles of individuals who are vulnerable or resilient to the effects of threat. The participants in this research were educated and routinely engaged in social, educational, or occupational activities. It may be that such individuals are somehow immune to the effects of threat as suggested by Horton et al. (2010). Another possibility is that the open and upfront disclosure about the diagnostic features of the tests enabled individuals to utilise their resources to the best of their ability in order to reject the diagnosis. Alternatively the participants may have employed an optimistic outlook in relation to test diagnosticity; in the age stereotype threat condition the induction states that ageing and poor cognitive performance are undeniably linked, however in the

dementia diagnosis threat condition the suggestion is perhaps more ambiguous in its assertion that some people over 65 have dementia which this leaves open the possibility that some over 65s do not have dementia. This may enable individuals to disidentify from the possibility that the threat relates to them. Applying these ideas to clinical practice suggests that it may be beneficial to adopt an open and upfront stance about the purpose of tests and their diagnostic potential as this may allow patients to expend their full effort into their performance. Furthermore, given the potential to make use of an optimistic outlook it may also be helpful to introduce the idea that there are people who attend memory clinics who do not receive a diagnosis, thereby enabling patients to make use of the potential to identify with the non-diagnosed group of patients.

Conclusion

This study provides support for the effect of negative age stereotypes on cognitive performance when individuals believe the threat induction. The threat of a dementia diagnosis was associated with higher levels of state anxiety however performance on recall, recognition and cognitive screening was not impaired. Age and education significantly predicted performance independent of threat induction which is in keeping with previous research findings.

Directions for future research arising from this study include the suggestion to increase the relevance of a dementia diagnosis threat by recruiting individuals who are worried about their memory or have noticed memory decline, and by querying or drawing attention to recent slips of memory or highlighting a family history of dementia. In addition the induction script could more explicitly state the diagnostic ability of the tests. In a healthy population it is recommended that a more

challenging measure of cognitive ability is utilised as ceiling effects were found on the ACE-R.

References

Alzheimer's Society (2012). *Dementia 2012: A national challenge.* Retrieved from http://alzheimers.org.uk/infographic.

Andreoletti, C., & Lachman, M. (2004). Susceptibility and resilience to memory aging stereotypes: Education matters more than age. *Experimental Aging Research*, *30*, 129–148.

Charles, S. T., Reynolds, C. A., & Gatz, M. (2001). Age-related differences and change in positive and negative affect over 23 years. *Journal of personality and social psychology*, *80*(1), 136-151.

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.

Dementia UK (2012). *About dementia.* Retrieved from http://www.dementiauk.org/information-support/about-dementia/#2

Desrichard, O., & Köpetz, C. (2005). A threat in the elder: The impact of task-instructions, self-efficacy and performance expectations on memory performance in the elderly. *European Journal of Social Psychology*, *35*, 537–552.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis for the social, behavioural, and biomedical sciences. *Behaviour Research Methods, 39*, 175-191.

Hess, T. M., Auman, C., Colcombe, S. J., & Rahhal, T. A. (2003). The impact of stereotype threat on age differences in memory performance. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *58B*(1), 3-11.

Hess, T. M., Emery, L., & Queen, T. L. (2009). Task demands moderate stereotype threat effects on memory performance. *The Journals of Gerontology: Psychological Sciences*, *64B*(4), 482-486.

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., & Hodges, E. A. (2009). Moderators of and mechanisms underlying stereotype threat effects on older adults' memory performance. *Experimental Aging Research*, *35*, 153-177.

Hess, T. M., Hinson, J. T., & Statham, J. A. (2004). Explicit and implicit stereotype activation effects on memory: Do age and awareness moderate the impact of priming? *Psychology and Aging*, *19*(3), 495-505.

Horton, S., Baker, J., Pearce, G. W., & Deakin, J. M. (2008). On the malleability of performance: Implications for seniors. Journal of *Applied Gerontology, 27,* 446-465.

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.

Ivnik, R. J., Malec, J. F., Tangalos, E. G., Petersen, R. C., Kokmen, E. & Kurland, L. T. (1990). The Auditory-Verbal Learning Test (AVLT): Norms for ages 55 years and older. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*(3), 304-312.

Kang, S. K., & Chasteen, A. L. (2009). The moderating role of age-group identification and perceived threat on stereotype threat among older adults. *The International Journal of Aging and Human Development*, *69*, 201–220.

Kit, K. A., Tuokko, H. A. & Mateer, C. A. (2008). A review of the stereotype threat literature and its application in a neurological population. *Neuropsychology Review, 18,* 132-148.

Kite, M. E. & Johnson, B.T. (1988). Attitudes toward older and younger adults: A meta-analysis. *Psychology and Aging, 3*(3), 233-244.

Larner, A. J. (2007). Addenbrooke's Cognitive Examination – Revised (ACE-R) in day-to-day clinical practice. *Age and Aging, 36*(6) 685-686. Retrieved from http://ageing.oxfordjournals.org/content/36/6/685.full.pdf.

Leyens, J-P., Desert, M., Croizet, J-C & Darcis, C. (2000). Stereotype Threat: Are lower status and history of stigmatization preconditions of stereotype threat? *Personality and Social Psychology Bulletin, 26,* 1189-1199.

Lezak, M. D. (1995). *Neuropsychological assessment (3rd Ed).* New York: Oxford University Press.

Mathuranath, P. S., Nestor, P. J., Berrios, G. E., Rakowicz, and Hodges, J. R. (2000). A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology, 55* (11), 1613-1620.

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R. & Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): A brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry, 21,* 1078-1085.

Mitrushina, M., Boone, K. B. & D'Elia, L. F. (1999). *Normative data for neuropsychological assessment.* New York: Oxford University Press.

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*(4), 697-706.

Scholl J. M. & Sabat, S. R. (2008). Stereotypes, stereotype threat and ageing: implications for the understanding and treatment of people with Alzheimer's disease. *Ageing and Society, 28,* 103-130.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35,* 4–28.

Spielberger, C. D., Gorusch, R. L., Lushene, R., Vagg, P. R. & Jacobs, G. A. (1983). State-trait anxiety inventory for adults: Sampler set, manual, test, scoring key. Palo Alto, California: Mind Garden.

Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613-629.

Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811.

Stein, R., Blanchard-Fields, F. & Hertzog, C. (2002). The effects of age stereotype priming on the memory performance of older adults. *Experimental Aging Research, 28,* 169-181.

Suhr, J. A. & Gunstad, J. (2002). "Diagnosis threat": The effect of negative expectations on cognitive performance in head injury. *Journal of Clinical and Experimental Neuropsychology, 24*(4), 448-457.

Suhr, J. A., & Gunstad, J. (2005). Further exploration of the effect of 'diagnosis threat' on cognitive performance in individuals with mild head injury. *Journal of the International Neuropsychological Society, 11*(1), 23-29.

Tovilović, S., Novović, Z., Mihić, L. & Jovanović, V. (2009). The role of trait anxiety in induction of state anxiety. *Psihologija, 42*(4), 491-504.

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, 54*, 1063-1070.

Wheeler, S. C. & Petty, R. E. (2001). The effects of stereotype activation on behaviour: A review of possible mechanisms. *Psychological Bulletin, 127*(6), 797-826.

Zacks, R. T., Hasher, L. & Li, K. Z. H. (2000). Human memory. In T. A. Salthouse and F. I. M. Craik (Eds.) *Handbook of Aging and Cognition*, (2nd ed., pp. 293-357). Mahwah, NJ: Lawrence Erlbaum.

Part 3: Critical Appraisal

Introduction

This appraisal will critically reflect upon some of the key conceptual issues that arose in attempting to explore the impact of stereotype threat on older adults' performance on a diagnostic test for dementia. It will also address some of the methodological issues encountered in recruiting an older adult sample for this research, paying particular attention to the problem of selection bias. Some ideas regarding the experiences of participants in this research will be considered and the appraisal will conclude with reflections on how various aspects of this research might have been carried out differently.

Conceptual issues

The main conceptual issues identified in this research were related to the limitations of applying the stereotype threat framework to age related stereotypes about memory.

Limitations of the stereotype threat framework

In the wider stereotype threat literature proficiency in the stereotyped performance domain appears to have no tangible relationship to the characteristics of the stereotyped group, for example Black students are not academically inferior and women are not innately poor at maths. However as people grow older their cognitive resources do tend to be depleted, leading to poorer performance on tests of such abilities (Craik & McDowd, 1987). Thus any assertion pertaining to stereotype threat

effects on memory in this population needs to be cautiously interpreted within this context.

A further issue which delineates investigations of stereotype threat in older adults from other stereotype threat research is that in other areas, group identification or membership is clearly defined, with no scope for movement between the stereotyped and non-stereotyped group. Old age however is an identity which most of us will someday assume. Thus it is possible that identification with this stereotyped group represents not only current belonging but also anticipated future belonging to that group. This may explain why some authors have found the effects of an age-related stereotype threat generalised across age groups (Andreoletti & Lachman, 2004; Chasteen, Bhattacharyya, Horhota, Tam & Hasher, 2005).

The literature on diagnosis threat, described as a special case of stereotype threat, utilises a population with known neurological histories and explicitly mentions membership of this diagnostic group to induce threat (Suhr & Gunstad, 2002, 2005). The fact that this study relied solely on the power of suggestion within a healthy cohort of older adults posed some difficulties in judging the applicability of a dementia diagnosis threat in this sample.

The apparent lack of applicability of the stereotype threat framework to this area of research perhaps offers some explanation for the research results.

Methodological issues

My approach to recruitment had a clear impact on the types of people who volunteered for this study. I will discuss this in terms of factors influencing participation which may have led to selection bias. Another methodological issue

which may help to explain the results of this study was the questionable efficacy of the threat induction.

Recruitment and selection bias

The majority of participants for this research were recruited via educational groups for seniors. Other sources of recruitment which proved fruitful were the UCL student body and word of mouth. Although recruitment was not necessarily problematic in this research, in terms of the potential audience that my requests reached, the numbers recruited from each area were surprisingly low. One possible reason for this may have been the restricted age range of 65-70 that I had imposed; informally I did come to learn that many members of the educational groups I contacted were over 70 and thus precluded from participating. In relation to word of mouth advertising, many participants seemed confident that their friends or family would be willing to participate; however in most cases this did not transpire. In attempting to widen the potential participant pool for this research I pursued other avenues of recruitment such as advertising in libraries; however posters and leaflets placed in five libraries across two London boroughs yielded only one participant. Endeavours which proved similarly futile were advertising in seniors exercise classes, drama groups and online pensioners forums.

In attempting to gain some insight into who was responding to my research request I examined the characteristics of the participants. They were a mainly White British (93%) sample of females (65%) who had been educated beyond secondary school (78%). The majority of participants were familiar with the internet and used email (97%) and had a mobile phone (98%). Whilst such individuals were the highest responders to the research advertisements, I queried whether this was a true reflection of the demographics of those who attended the educational groups I

was recruiting through or if there were systematic differences between those who responded and those who did not. I explored this by examining factors that might influence participation.

Travelling to central London

Participation in this research necessitated travel into central London by tube or bus, followed by a five to ten minute walk to the university building. The cost, time and level of ability required to do this were considered as potential barriers to participation. Several potential participants who lived outside the Greater London area reported being deterred by travel costs, although a similar number of people willingly travelled into London at their own expense in order to partake in this research. For those living in London however the issue of travel costs was bypassed by the eligibility of this population for free public transport.

For people with busy lives or commitments such as caring, a one hour appointment coupled with the time needed to travel to the university might have seemed unappealing. Those who did participate spoke of the time factor in a number of ways - some reported an abundance of free time and welcomed the opportunity to get out and do something a bit different, others scheduled their appointment to fit with existing commitments nearby, and some chose to use their time in central London as an opportunity to visit some of the local attractions, meet up with friends or family who lived or worked nearby.

Given the location of the university, participation in this research was difficult for those who were not comfortable or physically able to travel into central London for the appointment. The tube station nearest to the building did not have step-free access; other stations nearby were more accessible but necessitated a slightly

longer walk to the university. All those who attended were physically able and comfortable taking public transport into central London.

Seeing a trainee clinical psychologist for a memory test

The advertised description of this research project is likely to have aroused strong reactions in older people. The association between ageing and memory is rarely a positive one and this may well have been off-putting for some potential participants. On the other hand some may have seen it as an opportunity to test their memory skills or as one lady told me 'it was a challenge'. Informal feedback from some participants suggested that non-participants may have been fearful of undertaking a memory test in case they performed badly. This implies that those who did take part were perhaps representative of a small section of this age group who were confident about their cognitive abilities. It is possible therefore that those who did not volunteer might have been more suggestible to the threat induction.

Given that this research essentially involved people being tested, previous experiences of being tested were likely to be important in determining whether to volunteer for the study. It is reasonable to assume that for people with a stronger educational background, testing was something they were likely to be familiar with, and therefore fairly confident about. It is possible that those with more education felt more able to undertake such a task, whilst those with relatively less experience in this area might have been put off by the prospect of taking a test.

My requests for participants mentioned my job as a trainee clinical psychologist. It is possible that this set up some sort of expectations regarding my role, and perhaps people who were keen to gauge the opinion of a psychologist were attracted by the prospect. One participant explained that she had volunteered because she feared she had dementia and thought I could diagnose it; another

confided that he had once been told that he had 'psychological issues' which he hoped I could offer some insight into. Several participants told me of personal troubles prompting me to wonder whether in some way my research was attracting people who wanted a relatively informal way to meet with a psychologist.

It was revealed by some participants that they had been encouraged by friends or family members to partake in the research. I experienced this in different ways, for example, one lady attended at the suggestion of her daughter who apparently thought she was 'losing her marbles'. This situation needed to be managed sensitively though without jeopardising the experimental manipulation of threat. Others attended with their spouse or in sibling or friendship pairs which appeared to incite a sense of competition between participants. This was something I acknowledged in discussion with participants but remained ethically removed from, declining to comment on the performance of others when participants asked how they had fared in comparison with their companion. I wondered about the distinction between public and private participation, whether knowing that a loved one was going to ask about your experience and possibly about your test results might bring about different feelings than if you were attending of your own volition, with no expectation to report back on the experience.

Attitudes to research

Although participation in this research was voluntary, attendance rates were high (97%). I wondered whether this was a feature of how I had set up the study, sending a somewhat formal email to confirm appointment times, or if this apparent commitment to the research was more reflective of the character and values of the participants.

It is reported that one of the major reasons for older adults' non-participation in research is the feeling that 'it won't help me' (Kaye, Lawton & Kaye, 1990) thus it is reasonable to assume that the lack of direct personal benefits on offer in this research contributed to decisions to decline participation. It is possible however that people did perceive some indirect personal benefits to their participation such as a chance to prove their ability in the memory tests or a chance to ask questions about ageing and memory.

Another major incentive for older adults to volunteer for research is the altruistic potential to help others through participation Kaye et al. (1990). Several participants in this study commented on the importance of contributing to research, one gentleman described it as a civic duty which everyone should uphold. Others seemed to appreciate the attention given to issues regarding ageing and wanted to demonstrate this though their participation in the study, perhaps in the hope that their enthusiasm might encourage research to continue in this area. Thus it is possible that the prospect of a collective benefit to the older adult community was sufficient enticement to participate.

The efficacy of threat induction in this study

Given the difficulties highlighted in defining a group who might be susceptible to age stereotype threat and the further difficulties in ensuring the relevance of a diagnosis threat, it is unlikely that the manipulation of threat in this study worked as intended. In the first instance a significant number of people did not believe the induction for the no threat condition. This led me to doubt whether there was a genuine difference between the no threat and the age stereotype threat groups and whether participants' preconceived ideas or stereotypes were in fact more powerful than my

threat induction. Related to this point I noted the frequent assumption across all experimental conditions that this study was related to dementia research. This poses the question whether the mere hint of a relationship between ageing and memory conjured up ideas of dementia without me necessarily mentioning it. These issues led me to believe that the three experimental groups in this study were not as distinct as intended.

The experience of participation

The motivations for participating in research have been discussed above, yet given the centrality of the participants' role in research I felt it might be useful to further reflect on their experience. Consideration is given to the impact of being given feedback on test performance and being informed of the findings of the completed study.

Feeding back the results of a dementia screening test

I had not initially anticipated feeding back the results of the tests however I found that I became adept at scoring as I went along which opened up the possibility of offering participants immediate feedback on their performance. Bearing in mind the lack of direct personal benefit to participants in this study I felt this was a respectful and fitting way to give them something back by way of a thank-you for taking part in the research. Participants were offered the choice whether or not to receive feedback; no one declined the offer.

The experience of giving feedback was an interesting one with many participants commenting on the benefits of receiving their test results. This sentiment was summed up well in an email received from one of the participants who had been quite anxious about her memory:

> 'Thank you, Victoria, for this morning's session. It was lovely to meet you and I really enjoyed taking part in your project. I hadn't anticipated being given feedback on my results and was delighted to have that. I found it very re-assuring.'

Given that almost all participants performed well I was able to give a lot of positive feedback to which seemed to assuage much anxiety regarding memory and ageing. I encountered evident relief from participants and comments such as 'you've made my day' and 'thank goodness I'm ok' which really affirmed for me the value in offering quick, easily administered screening tests for dementia to this population. Several participants commented that such tests should be offered to all older adults. This experience led me to agree with this sentiment in the belief that much of the worry that exists for people in this age group could be vastly reduced, if not eliminated, by such simple screening.

For those who performed less well on the tests I was able to give some qualitative feedback about what their relative strengths and weaknesses might be and offered them some ideas about simple memory strategies that might be of use.

Dissemination of research findings

It has been argued that research results should be offered to participants as an ethical imperative to avoid treating individuals who volunteer for research solely as a means to an end (Fernandez, Kodish & Weijer, 2003). Even in the case that there is

no direct benefit to the disclosure the feelings of self worth through altruism and pleasure in knowing one has contributed to scientific knowledge should not be underestimated. This practice also offers potential benefits to the research community including raising the public awareness of research and its impact on clinical practice, and enhancing trust and interest in the research process through transparency (Fernandez et al., 2003).

The participants in this study were enthusiastic and interested in my research. I wanted to nurture this interest and also to acknowledge the collective outcome of their participation. As such I promised to disseminate an outline of the results. A lay summary is currently in preparation which will remind the participants of the aims and objectives of the study, the nature of their involvement and outline the overall results and plans for dissemination of the findings.

I am unsure how the participants will use this information. It may be that they will benefit from the reassurance that they are doing well as an educated, active, socially engaged cohort. They may value it as evidence that issues relating to ageing are seen as important within the research community and promote this within their peer group. I recently received an enthusiastic email from one of my participants requesting an article for the newsletter of a local educational group from where some participants were recruited. Participants might have hopes and ideas about the potential for this research to effect change in clinical practice. Unfortunately due in part to the conceptual difficulties mentioned earlier the results of the research are not clear enough to suggest any feasible changes in clinical practice. These results might however be of interest academically which could open up the possibility for more focused research in this area in the future.

How this research might have been done differently

Keeping a research journal throughout the process of data collection allowed me to reflect on the impact of various choices I made in relation to carrying out this research. Some ideas subsequently arose about other ways I might have approached the study.

Given the potential importance of offering some kind of direct benefit to participants, advertising the research differently, for example framing it as a free memory test, may have increased the response rate. It may have yielded a different type of participant, possibly those who were worried about their memory and thus may have been more susceptible to the threat induction. This might have created other difficulties however, potentially opening up a forum for people with real memory problems to attend thereby increasing the likelihood of participants naturally meeting the diagnostic criteria for dementia. A more thorough telephone screening procedure would have been needed to take this into consideration.

Recruiting in person may have been another option as it would have removed the need for individuals to take on the responsibility of contacting me about the research. This may have enabled participation from those who were ambiguous about the study or who had some anxieties regarding their potential involvement. I expect that recruiting in this way would have made the sample more heterogeneous in terms of confidence in memory ability.

Another important change to the method of data collection might have been to offer appointments in local areas rather than asking people to travel to central London. This would have benefitted those who were unable to make the journey due to time constraints or physical difficulties. Although this is likely to have increased the response rate it would have impacted significantly on my time and would not

have been the most efficient way to collect data. One possible solution might have been to offer specific dates in different areas, perhaps timed to coincide with the educational courses these participants attended.

References

Andreoletti, C., & Lachman, M. (2004). Susceptibility and resilience to memory ageing stereotypes: Education matters more than age. *Experimental Ageing Research*, *30*, 129–148.

Chasteen, A. L., Bhattacharyya, S., Horhota, M., Tam, R., & Hasher, L. (2005). How feelings of stereotype threat influence older adults' memory performance. *Experimental Ageing Research*, *31*, 235–260.

Craik, F. I. M. & McDowd, J. M. (1987). Age differences in recall and recognition. Journal of Experimental Psychology. Learning Memory and Cognition, *13*(3), 474-479.

Fernandez, C. V., Kodish, E. & Weijer, C. (2003). Informing Study Participants of Research Results: An Ethical Imperative. *Ethics and Human Research, 25*(3), 12-19.

Kaye, J. M., Lawton P. & Kaye, D. (1990). Attitudes of elderly people about clinical research on ageing. *The Gerontologist, 30*(1), 100-106.

Suhr, J. A. & Gunstad, J. (2002). "Diagnosis threat": The effect of negative expectations on cognitive performance in head injury. *Journal of Clinical and Experimental Neuropsychology, 24*(4), 448-457.

Suhr, J. A., & Gunstad, J. (2005). Further exploration of the effect of 'diagnosis threat' on cognitive performance in individuals with mild head injury. *Journal of the International Neuropsychological Society, 11*(1), 23-29.

Appendix A

Participant Information Sheet

Study Title: An investigation into stereotyping and memory

Researcher: Victoria Page
Email:
Telephone:

You have been invited to take part in a research study. Before you decide you need to understand what the research would involve for you. Please take time to read the information carefully and feel free to ask me questions or request further information.

The aim of this study is to examine individual differences in performance on a range of tests for attention, memory, and language skills. Research has shown that people vary a great deal in these abilities and that performance on these tests is unrelated to age. For example the strategies we use for learning lists of words tend to stay stable throughtout the lifespan. The tests I will carry out today will help me to find out a bit more about this.

OR

The aim of this study is to examine performance on a range of tests for attention, memory, and language skills. Research has shown that the normal ageing process can have a negative impact on these abilities. For example as we get older, from around age 65, we find it more difficult to remember details of recent events or to learn new information. The tests I will carry out today will help me to find out a bit more about this.

OR

The aim of this study is to examine performance on a range of tests for attention, memory, and language skills. Research has shown that as people age they perform worse on such tests. One possible reason for this could be age-related illnesses such as dementia. The term dementia refers to a collection of symptoms including memory loss and cognitive decline e.g. a reduction in reasoning and communication skills. According to the Alzheimer's Society there are about 750,000 people in the UK with dementia. It most commonly occurs in people aged 65 or over. The tests I will carry out today will help me to find out a bit more about this.

You have been invited to participate in this study as a healthy adult aged 65-70. In total there will be 60 participants involved in the study.

What your participation will involve
In this study you will be asked to complete two self-report questionnaires:

1. The Positive and Negative Affect Schedule (PANAS)
2. The State-Trait Anxiety Inventory (STAI)

You will also be given two tests requiring mostly verbal and some written responses.

1. The Auditory-Verbal Learning Test (AVLT)
2. The Addenbrooke's Cognitive Examination – Revised (ACE-R)

Testing will take around 1 hour.

There are no foreseeable risks or discomforts; neither are there any perceived benefits associated with your participation in this study.

<u>Your rights as a participant</u>

This study has been approved by the UCL Research Ethics Committee (Project ID Number: 2997/001). Your participation is entirely voluntary and your rights as a participant, including the right to withdraw at any point without penalty, are ensured. You can leave the study, or request a break, at any time.

The results of this study will be written up as part of my doctoral dissertation. Your participation in the study is confidential and all results will be made anonymous meaning it will not be possible for anyone to identify individual participant's data. Your data will not be made available to any other person. The data will be collected and stored in accordance with the Data Protection Act 1998. It will be retained for up to 18 months at which time it will be disposed of in a secure manner.

On completion of the testing the researcher will be available to answer any questions you may have about your participation in the study. At this point you will be provided with further information about the study. You will be also given the option of requesting a copy of the completed study.

If, at a later date you have any questions or concerns relating to your participation in this study please contact me using the details on this information sheet.

If you are happy to proceed please read and complete the consent form.
Thank you.

Appendix B

Consent Form

Title of study: An investigation into stereotyping and memory

Name of Researcher: Victoria Page

1. I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. I understand that I am also free to ask more questions and request further information at a later date.

2. I understand that my participation is voluntary and that I am free to withdraw at any time without penalty and without giving any reason.

3. I understand that my participation is entirely confidential and that all data collected in this study will be made anonymous.

4. I agree to take part in the above study

_____   _____   _____

Name of participant   Signature   Date

_____   _____   _____

Name of researcher   Signature   Date

Appendix C

Participant Debriefing Information Sheet

Researcher: Victoria Page
Email:
Telephone:

Thank you for your participation in this study. Now that you have completed the tests I would like to tell you a bit more about the background of the study and what the results may show.

In this study I carried out the same set of tests with all participants, however I gave people different information prior to testing. One third of participants were told that these were tests of individual differences in ability; one third were told they were tests of age related differences in ability and one third were told they were tests of abilities which could be affected by illnesses such as dementia. Participants were assigned randomly to one of these conditions. I am hoping to show that there will be overall differences in performance between the three conditions.

Research has shown that people's performance or behaviour can be affected by commonly held beliefs or stereotypes about their social group. When these stereotypes are negative we experience something known as 'stereotype threat' which means that we become aware that we are at risk of confirming, as self-characteristic, this negative stereotype about our social group.

This idea was first introduced in the 90s when researchers in America found that Black college students performed more poorly on tests than White students when their race was emphasized. When race was not emphasized, however, Black students performed the same as or better than White students. The results showed that academic performance can be hindered by the awareness that one's performance might be viewed through the lens of racial stereotypes.

We believe that a similar effect may be present with older adults and tests of memory based on the stereotype that ageing leads to deterioration in memory skills. Several studies have found that activating negative stereotypes of ageing can lead to reduced performance on memory tasks.

Memory tasks are an integral part of clinical testing for dementia. An understanding of the impact of stereotype threat on performance on these tasks has important implications for the way we present the tests. If we can thus reduce stereotype threat we are more likely to obtain an accurate picture of an individual's ability.

The tests I used in this study are routinely used in dementia testing however they cannot be used as standalone tests to make a diagnosis. I will be able to provide feedback on your test performance today however I am not able to use these results to make or rule out a diagnosis of dementia. If you have any concerns about your memory please visit your GP. If I have concerns about your performance on these tests I will discuss these with you and advise you to seek help from your GP.

If after leaving here today you find that you have any questions or concerns about your participation in this study please contact me either by phone or email and I will endeavour to get back to you as soon as possible. If you would like to receive a copy of the study once it is completed next year I will be able to provide this electronically. Please email me on the address above to request a copy.

Appendix D

Ethical Approval

UCL RESEARCH ETHICS COMMITTEE
GRADUATE SCHOOL OFFICE

Dr John King
Division of Psychology and Language Sciences
Room TP440,
1-19 Torrington Place
UCL

5 April 2011

Dear Dr King

Notification of Ethical Approval
Ethics Application: 2997/001: An investigation into stereotyping and memory

I am pleased to confirm that in my capacity as Chair of the UCL Research Ethics Committee I have approved your project for the duration of the study (i.e. until September 2012).

Approval is subject to the following conditions:

1. You must seek Chair's approval for proposed amendments to the research for which this approval has been given. Ethical approval is specific to this project and must not be treated as applicable to research of a similar nature. Each research project is reviewed separately and if there are significant changes to the research protocol you should seek confirmation of continued ethical approval by completing the 'Amendment Approval Request Form'.

The form identified above can be accessed by logging on to the ethics website homepage: http://www.grad.ucl.ac.uk/ethics/ and clicking on the button marked 'Key Responsibilities of the Researcher Following Approval'.

2. It is your responsibility to report to the Committee any unanticipated problems or adverse events involving risks to participants or others. Both non-serious and serious adverse events must be reported.

Reporting Non-Serious Adverse Events
For non-serious adverse events you will need to inform Helen Dougal, Ethics Committee Administrator (ethics@ucl.ac.uk), within ten days of an adverse incident occurring and provide a full written report that should include any amendments to the participant information sheet and study protocol. The Chair or Vice-Chair of the Ethics Committee will confirm that the incident is non-serious and report to the Committee at the next meeting. The final view of the Committee will be communicated to you.

Reporting Serious Adverse Events
The Ethics Committee should be notified of all serious adverse events via the Ethics Committee Administrator immediately the incident occurs. Where the adverse incident is unexpected and serious, the Chair or Vice-Chair will decide whether the study should be terminated pending the opinion of an independent expert. The adverse event will be considered at the next Committee meeting and a decision will be made on the need to change the information leaflet and/or study protocol.

On completion of the research you must submit a brief report (a maximum of two sides of A4) of your findings/concluding comments to the Committee, which includes in particular issues relating to the ethical implications of the research.

Best wishes for the research.

Yours sincerely
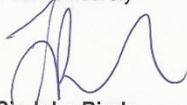
Sir John Birch
Chair of the UCL Research Ethics Committee

Cc: Victoria Hanrahan, Division of Psychology and Language Sciences, UCL