

Validating Self-Reports For Use in Contact Research

**Melanie Sarah Sharp
Hertford College**



Thesis submitted for the degree of Doctor of Philosophy

**Department of Experimental Psychology
University of Oxford**

March, 2009

ACKNOWLEDGEMENTS

I owe *enormous* thanks to the following people:

For getting me through this, Miles Hewstone.

For paying for it, the Economic and Social Research Council.

For invaluable help with analysis, Charles Judd.

For extremely helpful suggestions, Tessa West and Alberto Voci.

For being indispensable in so many ways, Penny Tarrant

For always being lovely, Gwenda Simons.

For sharing my pain, Claudia Marinetti.

For providing companionship (and eye-candy), Keon West.

For taking me in when I ran out of money, Christopher Sharp.

For listening to me rant, and resisting the urge to murder me, Sarah Sharp.

For making it all worthwhile, Mike McCabe.

CONTENTS

Acknowledgements	ii
Contents	iii
List of Tables	vii
List of Figures	viii
Abstract	ix
Detailed Abstract	x
 CHAPTER ONE: Introduction	 1
Prejudice: Fading or Faking?	6
Explicit and Implicit Prejudice	8
Prejudice reduction	10
The Contact Hypothesis	11
Allport's conditions - Cooperation	11
Allport's conditions - Common goals	12
Allport's conditions - Equal status	12
Allport's conditions - Social sanction	13
Later work	13
Cross group friendship	14
Indirect cross group friendship	15
Methods used in contact research	16
Observational studies	18
Experimental studies	19
Correlational studies	20
Weaknesses of the self-report methodology	21
Lack of adequate self-knowledge	21
Sources of error in self-insight – Overconfidence	23
Sources of error in self-insight – Biased information processing	26
Context effects	28
Response biases	31
Socially desirable responding	32
Non-dyadic methodology	33
Thesis aim	35
Validating self-reports	35
Validity	35
Existing methods	37
Observer-reports	39
Choice of 'observer'	40
Overview of the Thesis	43
 CHAPTER TWO: Resistance of contact self-reports' to context effects	 45
Choice of outgroup	47
Study 1	48
Hypotheses	49
Method	49
Participants	49
Procedure	49
Materials	50

Recall items.....	50
Contact/Attitude items	51
Results and Discussion	52
Study 2	55
Hypotheses.....	56
Method	56
Participants	56
Procedure	56
Materials	58
Results and Discussion	60
General Discussion	62
CHAPTER THREE: Validation of contact self-reports using the observer-reports of family members	65
Study 3	65
Factors affecting observer-reports	66
Non-independence	67
Dyadic analyses	70
Pseudo-couple analysis	70
Hypotheses.....	71
Method	72
Participants	72
Procedure	72
Materials	72
Results and Discussion	76
Study 4	86
Hypotheses.....	87
Method	88
Participants	88
Procedure	88
Materials	89
Results and Discussion	92
Study 5	97
Hypotheses.....	98
Method	99
Participants	99
Procedure	99
Materials	99
Results and Discussion	100
General Discussion	109
CHAPTER FOUR: Validation of Contact Self-Reports using Multiple Observer-ratings.....	112
Non-independence	114
Study 6	114
Hypotheses.....	115
Method	116
Participants	116
Procedure	116
Materials	117

Results and Discussion	119
Study 7	127
Hypotheses.....	128
Method	129
Participants	129
Procedure	129
Materials	129
Results and Discussion	130
General Discussion	137
Utility and limitations of self-reports of contact.....	138
Discriminant validity	140
CHAPTER FIVE: Validation of contact self-reports using online communities	143
The online community	145
Online contact	147
Study 8	149
Hypotheses.....	150
Method	151
Participants	151
Procedure	151
Materials	152
Results and Discussion	153
Study 9	156
Implicit Association Test.....	158
Hypotheses.....	159
Method	159
Participants	159
Procedure	159
Materials	160
IAT.....	161
Results and Discussion	163
General Discussion	166
CHAPTER SIX: Summary and Conclusions.....	169
Thesis Aims	169
Resistance to context effects	174
Agreement with observer-reports	174
Criterion-related validity.....	179
Theoretical implications of the research	180
Validity of self-reports for measuring contact	180
Attitude self-reports	181
Low contact.....	181
Potential limitations of the research.....	182
Potential directions for future research.....	184
Summary	187
REFERENCES	188

APPENDICES	225
Appendix 1: Full matrix of correlations among target effects and self-ratings – Asian and gay contact and liking (Study 7)	225
Appendix 2: Facebook Friends Page screenprint.....	226
Appendix 3: Outgroup contact and attitude questions	227
Quantity of contact with outgroup – Included in Studies 2-6, 8, and 9	227
Quality of contact with outgroup – Included in Studies 6 and 8	227
Direct cross-group friendship with outgroup - Included in all studies	227
Indirect cross-group friendship with outgroup - Included in Studies 2-6, 8, and 9	228
Attitude towards outgroup – Subsection of these included in all studies	228
Appendix 4: Closeness measures	229
Studies 3, 4 and 5	229
Study 6	229
Studies 6 and 7	229
Appendix 5: Self-disclosure measures used in Studies 3, 4, 5 and 7	230
Appendix 6: Measures of ‘visible’ traits, for providing a baseline of comparison for self-observer agreement	231
Extroversion - used in Studies 3, 4, 5, & 6	231
Political views – used in Study 6	231

LIST OF TABLES

Table 1.1	Representation of ethnic and gender groups in the United States Congress	4
Table 1.2	Representation of ethnic and gender groups in the UK Parliament	5
Table 2.1	Mean attitude and contact as a function of presence and valence of prime (Study 1)	53
Table 2.2	Mean attitude and contact as a function of presence and valence of prime (Study 2)	61
Table 3.1	Mean contact, attitudes, extroversion, closeness and demographic variables (Study 3)	78
Table 3.2	Similarity in married couples: Real discrepancy and discrepancy expected by chance between male and female self-reports (Study 3)	80
Table 3.3	Discrepancy between self-reports and observer-reports: Real discrepancy and discrepancy expected by chance (Study 3)	81
Table 3.4	Mean contact at home and university, attitudes, extroversion, self-disclosure and demographic variables (Study 4)	93
Table 3.5	Similarity in parent-child dyads: Real discrepancy and discrepancy expected by chance between student and parent self-reports (Study 4)	94
Table 3.6	Discrepancy between self-reports and observer-reports: Real discrepancy and discrepancy expected by chance (Study 4)	95
Table 3.7	Reliability statistics for self- and observer-reported extroversion, attitudes and contact at Time 1 and Time 2 (Study 5)	101
Table 3.8	Mean self- and observer-reported levels of extroversion, attitudes and contact at Time 1 and Time 2 (Study 5)	102
Table 3.9	Students' similarity to their parents before and after attending university: Discrepancy and discrepancy expected by chance (Study 5)	105
Table 3.10	Self-observer agreement before and after attending university: Discrepancy and discrepancy expected by chance (Study 5)	106
Table 3.11	Discriminant validity: Cross-matched discrepancy and cross-matched discrepancy expected by chance (Study 5)	108
Table 4.1	Mean ratings of self and others (Study 6)	121
Table 4.2	Relative variance partitioning – SOREMO analysis (Study 6)	123
Table 4.3	Self-rater and self-target correlations (Study 6)	124
Table 4.4	Mean ratings of self and others (Study 7)	131
Table 4.5	Relative variance partitioning – SOREMO analysis (Study 7)	132
Table 4.6	Self-rater and self-target correlations (Study 7)	133
Table 4.7	Correlations among target effects and self-ratings: Asian and gay contact and liking (Study 7)	135
Table 5.1	F tests of self- and observer-reported outgroup contact and attitudes between those with recent online contact and those without (Study 8)	154
Table 5.2	Means & standard deviations (Study 9)	165
Table 6.1	Summary of studies in thesis	170

LIST OF FIGURES

Figure 1.1	Percentage of respondents who would/would not vote for a well-qualified black candidate for President of the United States.	2
Figure 1.2	Methods used in contact research, as reported by Pettigrew and Tropp, 2006	17
Figure 2.1	Passages used in Study 2	57
Figure 3.1	Moderating effect of length of relationship on the relationship between self-disclosure and discrepancy between male observer-reports and female self-reports of direct cross-group friendship.	85
Figure 5.1	Schematic description and illustration of the IAT as used in Study 9	162

Validating Self-Reports for use in Contact Research

Melanie Sarah Sharp
Hertford College

D.Phil in Experimental Psychology
Submitted Hilary Term, 2009

ABSTRACT

The aim of this thesis was to investigate whether self-report measures of contact are valid for use in research testing the ‘contact hypothesis’. The vast majority of contact research has relied on the assumed validity of self-report methods of data collection (Pettigrew & Tropp, 2006), even though the potential weaknesses of self-report methodology generally have been well documented. This reliance is necessary, as self-reports remain the only practical method so far developed of measuring certain of the facilitating conditions developed by Allport (1954/1979), and particularly of direct and indirect cross-group friendship (Pettigrew, 1998; Wright, Aron, McLaughlin-Volpe, & Ropp, 1997). However, if self-reports are not a valid method for measuring contact, the derived implications of a large portion of the research effort are potentially flawed. This thesis attempted to address this important oversight, using a variety of methods to investigate whether the use of self-reports in future research on intergroup contact is appropriate. Studies 1 and 2 demonstrated that self-reports of contact show considerable resistance to context effects, particularly in comparison with self-reports of the more subjective construct of attitudes. Studies 3-5 demonstrated that self-reports of contact agree with the observer-reports of a single observer who knows the target intimately – the spouse or parent. Studies 6 and 7 replicate this agreement through the consensually supported observer-reports of three close friends of the target, thereby reducing any variance due to individual response biases. Finally, studies 8 and 9 demonstrate the concurrent criterion-related validity of self-reports of contact, in that they are able to predict contact on a very large online network called Facebook, on which real-world rather than purely online friendships are primarily represented. These findings offer considerable support for the validity of self-reports as a suitable method for measuring contact. As self-reports remain the only method which has thus far proven suitable for the measurement of those aspects of contact which are essential for exploration of the contact hypothesis, this thesis presents a very heartening and optimistic conclusion and supports the continued use of self-reports in contact research.

DETAILED ABSTRACT

Research into the issues surrounding Allport's (1954/1979) proposal that, under certain conditions, contact is an effective means for reducing prejudice, has relied, for the most part, on self-reporting methods of measuring contact. As the issues to be explored include such aspects as the quality and nature of the contact experienced, as well as the intimacy and potential for friendship, there appears to be no alternative but to use self-reporting methods. More recently, research has centred on the prejudice-reducing benefits of individuals forming direct friendships with outgroup members (for a review, see Turner, Hewstone, Voci, Paolini, & Christ, 2008), as well as on the potential of simply being aware that other ingroup members have outgroup friends (Wright, Aron, McLaughlin-Volpe, & Ropp, 1997), for prejudice reduction. These aspects too are difficult to measure efficiently and practically without the use of self-reports.

However, researchers in both personality and social psychology have long questioned the validity of self-reports. Individuals do not always have the required self-knowledge to provide accurate self-reports (Haddock, Rothman, Reber, & Schwarz, 1999), and what knowledge they have is often flawed due to overconfidence or biased information processing (Dunning, 2005). Even assuming individuals have access to the constructs they are being asked to report on, self-reports have been shown to be context-dependent (Schwarz, 1999), with preceding questions influencing the answers to subsequent ones (Schwarz & Sudman, 1992). Most importantly, self-reports are very susceptible to response biases (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003) such as socially desirable responding (SDR), acquiescence or its opposite, and extreme or overly moderate responding (Paulhus, 1991). This thesis attempted to assess the validity of self-reports for the measurement of contact, in order to determine whether

the methodology on which a large portion of the intergroup contact research effort hinges does in fact hold water.

Chapter One provides an introduction to the topics examined in the thesis. I first provide an overview of contact research, concentrating on the evidence surrounding Allport's (1954/1979) contact hypothesis as well as introducing the concepts of direct and indirect cross-group friendship. I then discuss the self-report methodology, its benefits as well as its weaknesses, and how they have been addressed in other areas. The issue of validity is then covered, and I evaluate the methods which have been used to date to assess the validity of self-reports. Finally, I provide an overview describing how this thesis will address the question of the validity of self-reported contact.

In Chapter Two, two studies demonstrate that self-reports of contact (although not self-reports of attitudes) are resistant to context effects. Study 1 primed participants to consider a positive or negative encounter with a member of the outgroup, either before or after completing self-report questions about their contact and attitudes. While no effects were found when the prime came after the self-report questions, when the prime came first attitudes were affected in the direction of valence of contact while contact self-reports remained unaffected. Study 2 provided participants with a 'newspaper article' to read which described either a positive or negative outgroup encounter, while control groups read the same story with mention of the outgroup removed. Again, attitudes were affected in the direction of the valence of the prime while contact self-reports remained unaffected.

Three studies are described in the third chapter, all making use of the method most commonly suggested in personality research for the validation of self-reports (Piedmont, McCrae, Riemann & Angleitner, 2000). Observer-ratings are arguably the best method developed to date for validating self-reports (Howard, 1994). The person

best equipped to make such ratings is the spouse (McCrae, 1993), and Study 3 shows that self-ratings of contact agree with spouse-ratings. Studies 4 and 5 address the issue that many people have no spouse, and that this is particularly true of the population most commonly used in social-psychological research, namely that of students. These studies showed that parents' ratings of students' contact agree strongly with the students' self-ratings. These three studies used a technique appropriate to the analysis of non-independent, dyadic data. The moderating influence of self-disclosure is also tested and discussed.

Chapter Four corrected a potential weakness of the studies reported in Chapter Three. Individual response styles will have contributed error variance in studies which gathered observer-reports from a single observer. Studies 6 and 7 made use of the round-robin method to examine agreement between self-reports and the observer-reports of three close friends. As it is unlikely that the three observers would share a single response style, the consensual portion of their observer-ratings should include no error variance and should constitute an ideal criterion against which to compare self-reports. Studies 6 and 7 showed that self-reports of contact do indeed agree with consensual observer-reports. Study 7 also demonstrated the discriminant validity of self-reports of contact, by showing that participants distinguish between contact with two different outgroups rather than rating a target's contact with outgroups in general. Both these studies used a method appropriate for analysing round-robin data, in which every member of a group rates every other member.

Chapter Five examined predictive and concurrent criterion-related validity by demonstrating a strong relationship between self-reports of contact and real observed contact in a large online community which reflects real-world friends rather than merely 'e-relationships'. Study 8 showed that self-reported contact predicted whether,

at a later date, participants would have experienced a recent instance of online contact with an outgroup, while Study 9 found that participants' self-reported contact agreed strongly with the number of real-world outgroup friends who interact on the 'Facebook' network.

The final chapter summarises the main findings of the thesis and their implications. The contributions of this thesis to the field of intergroup contact are discussed, and I point out that these chapters present the first, and thus far the only, studies undertaken to investigate the validity of self-reports of intergroup contact. I discuss the potential weaknesses of the research reported here, and provide some suggestions for the direction that future research in this area could take. I conclude that this thesis has demonstrated the validity of self-reports of contact, supporting their continued use in intergroup contact research.

CHAPTER ONE:

Introduction

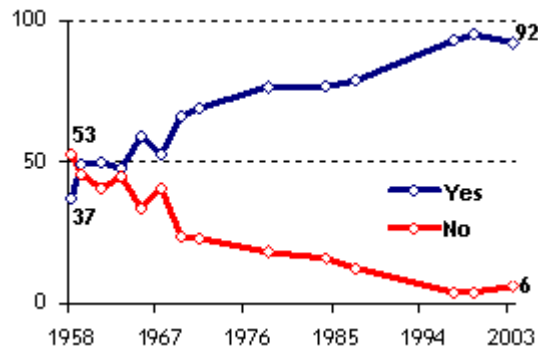
“Does this mean that America is now beyond black and white? No, but we are much further ahead than we were a week ago. Obama’s victory, no matter what one’s politics, is a redemptive moment in the life of a nation for which race has been called, simply and starkly, ‘the American dilemma.’”

Meacham (2008, p. 1)

On November 4th, 2008, the United States elected its first ever black president. Newspapers the world over have proclaimed this a major step forward for racial equality. Polls conducted by the Gallup Organisation have been asking Americans since 1958 whether they are willing to vote for a qualified black candidate in a presidential election (Clayton, 2007). Figure 1.1 demonstrates that, in 1958, 53% responded that they would not be willing. In 1984 this figure had decreased to 18%, and in 2003, only 6% gave the same response. In 2006, a Newsweek poll cited by Keeter and Samaranayake (2007) found that only 3% of Americans now admit to being unwilling to vote for a qualified black candidate. And, in 2008, during a time when Muslims are increasingly hated and feared, a black candidate whose father was raised a Muslim has been decisively voted to the position of what most newspapers describe as ‘leader of the free world’.

Figure 1.1

Percentage of respondents who would/would not vote for a well-qualified black candidate for President of the United States.¹



These findings do indeed promote optimism. But the election of Democratic candidate Barack Obama to the presidency of the United States does not by any means signal the end of racism. Obama's campaign encountered a number of obstacles as a result of explicit racial prejudice. Dowd (2008) and Merida (2008) report that Democratic canvassers received such responses as “I’ll never vote for a black person” (Muncie, Indiana), “I wouldn’t vote for him because he is black” (Pittsburgh, Pennsylvania) and even “I couldn’t possibly vote for Obama. Hang that darky from a tree!” (Susquehanna County, Pennsylvania, where 98% of residents are white). Bomb threats were made to three Obama campaign offices in Indiana, and plans were even drawn up for a Neo-Nazi magazine cover to feature a picture of Obama speaking to a crowd with a rifle’s cross-hairs superimposed on his head, and the headline “Kill This Nigger?” (Potok, 2008). Even Democratic voters have shown prejudice towards Obama; a poll taken in May 2008 in West Virginia found that 2 in 10 white Democratic voters said that race was important in how they voted, and more than 80% of those chose to vote to nominate Hillary Rodham Clinton as presidential candidate,

¹ Source: Keeter & Samaranayake (2007)

rather than Obama, as a result. An Obama volunteer working in Lackawanna, Pennsylvania, reported that a fellow democrat volunteer, working on the Clinton campaign, said to her “He’s a half breed and he’s a Muslim. How can you trust that?”

There is no doubt that racial prejudice was a salient issue in the 2008 United States Presidential Election. Babington (2008) predicted that Barack Obama could lose up to six percent of the votes simply as a result of being black. The 2000 and 2004 elections were decided by much smaller margins than six percent, suggesting that Obama could easily have lost the election because of racial prejudice. Babington quoted Stanford University political scientist Paul Sniderman as saying that, for a minority candidate, there is “a penalty for prejudice, and it's not trivial” and that racial prejudice “might be enough to tip the election.”

Although racial prejudice was not enough to cost Obama the election in this instance, it remains harder for blacks to get elected to positions of leadership (Terkildsen, 1993), and to gain employment (Bertrand & Mullainathan, 2003). Bertrand and Mullainathan (2003) sent fictitious resumes to help-wanted newspaper ads, providing either a very white-sounding name (such as Emily Walsh or Greg Baker) or a very black-sounding name (such as Lakisha Washington or Jamal Jones) with the result that -- all other things remaining equal -- white names received 50% more calls to interview than black names. This finding held true across different industries and types of occupation, and did not decrease among self-proclaimed 'equal opportunity employers'. Bertrand and Mullainathan suggest that their results show that affirmative action programs have not led, as feared by some, to 'reverse discrimination', and that racial discrimination is still very much a prominent feature of the labour market.

Whites dominate public offices in the United States and the United Kingdom, as can be seen in Tables 1.1 and 1.2. Non-whites and women are under-represented in both countries. Twenty-five states have never had a black Congressman (Yourish, 2008) and, at the time of writing, Piyush 'Bobby' Jindal, Governor of Louisiana, is the first and only South Asian elected government official at the national level. Britain has never had a non-white prime minister, and the most liberal of the main British political parties, the Liberal Democrats, has no black or Asian Members of Parliament. At the present rate of progress it will take until 2080 before ethnic minorities are properly represented in the House of Commons (McSmith, 2007).

Table 1.1

Representation of ethnic and gender groups in the United States Congress

	White	Non-White	Women	Total
Senate	93 (93)	7 (7)	16 (16)	100 (100)
House of Representatives	361 (83)	74 (17)	74 (17)	435 (100)
Total in US Congress	454 (85)	81 (15)	90 (17)	535 (100)
Population	221,331,507 (73)	80,289,650 (27)	152,962,259 (51)	301,621,157 (100)

Note. Percentages in parentheses. Table compiled partially using data from Center on congress (2008) and U.S. Census Bureau (2006, 2007).

Table 1.2

Representation of ethnic and gender groups in the UK Parliament

	White	Non-Whites	Women	Total
Liberal Democrat MPs	62 (100)	0 (0)	10 (16)	62 (100)
Labour MPs	342 (96)	13 (0)	98 (28)	355 (100)
Conservative MPs	196 (99)	2 (0)	17 (9)	198 (100)
Other MPs	31 (100)	0 (0)	3(10)	31 (100)
Total MPs	631 (98)	15 (2)	121 (19)	646 (100)
Population	54,153,890 (92)	4,635,296 (8)	30,488,622 (52)	58,789,194 (100)

Note. Percentages in parentheses. Table compiled partly using data from Cracknell (2005) and Office for National Statistics (2001).

The final report in September 2007 of the British Commission for Racial Equality described how there have been great improvements in racial prejudice and discrimination since the Commission for Racial Equality was established in 1976 by the Race Relations Act, but that the goal, to eradicate such prejudice and discrimination from British life entirely, is still a long way off. The report, titled 'A Lot Done, a Lot to Do', stated how the victims of inner-city youth violence are mostly black, and that an ethnic minority baby born in Britain today is still more likely to receive poor quality education and standard of living than a white baby. The employment rate is 75% overall, but only 60% among ethnic minorities (Commission for Racial Equality, 2007). Although ethnic minorities comprise almost 8% of Britain's population, only 3.5% of police officers, 4% of prison officers and 7% of court staff are from ethnic minorities. The top 10% of all professions is still almost entirely white, and employment practices still exist that exploit non-white labour (Commission for Racial Equality, 2007). Finally, a new public anxiety has arisen surrounding Muslims, resulting from events such as the 7th July suicide bombings in

London. The Commission for Racial Equality has not been able to stop rising hostility towards immigrants and asylum seekers, and there seems to be a tendency to blame white racism on terrorist Muslims and immigrants (McSmith, 2007). Even among the younger generation who have grown up in a society where it is socially unacceptable to express prejudice, bias still exists. There are fewer cross-race than same-race friendships among students in the United States and United Kingdom (Boulton & Smith, 1996; Braha & Rutter, 1980; Clark & Ayers, 1992; Davey, 1983; Denscombe, Szulc, Patrick, & Wood, 1986; Graham & Cohen, 1997; Hallinan & Teixeira, 1987a; Howes & Wu, 1990; Shrum, Cheek, & Hunter, 1988 cited in Aboud, Mendelson & Purdy, 2003), and even in integrated schools in the United States, students have three times as many same-race friends as cross-race friends (DuBois & Hirsch, 1990; Hallinan & Teixeira, 1987a). This is true of both black and white students, so the finding does not result from a difference in the numbers of students of various ethnicities, it is simply a same-race preference. It is quite apparent that we have a long way to go before we can truly say we live in a fair and just society.

Prejudice: Fading or Faking?²

The finding that since 1958 the percentage of Americans unwilling to vote for a well-qualified black candidate for president has decreased from 53% to 3% (Keeter & Samaranayake, 2007) seems at odds with the inequality still apparent in society today. This decrease may not reflect real attitude change, but simply a change in willingness to admit to what increasingly amounts to a socially unacceptable prejudice.

People *seem* to be willing to express support for a black candidate for election to public office (Terkildsen, 1993) or for president (Schuman, Steeh & Bobo, 1985)

² From Sigall & Page (1971)

but actual votes frequently tell a very different story. Pre-election estimates of support for black candidates do not correlate with the final vote tallies in bi-racial contests (Bergholz 1982; Clymer 1989; Finkel, Guterbock & Borg, 1991; Pettigrew & Alston 1988; Sussman 1985). This phenomenon, known as the 'Bradley effect', results from the tendency of a significant number of voters to report to pollsters an intention to vote for a black candidate in a bi-racial election, but to opt for the white candidate once safely and anonymously hidden in the polling booth.

The term originates in the 1982 gubernatorial race in California, where African American Tom Bradley ran against European American George Deukmejian. Bradley was consistently ahead by anywhere from 9 to 22 points (a 100 point lead reflecting a monopoly) in the polls leading up to the election, but narrowly lost the election itself. Later bi-racial contests have demonstrated the same phenomenon, with pre-election polls over-predicting eventual leads by 10.6 points for Douglas Wilder for Governor of Virginia in 1989, and by 16 points for David Dinkins for Mayor of New York City in the same year, to give two examples (Keeter & Samaranayake, 2007). This discrepancy does not reflect a change of heart; even after they have cast their ballot, voters seem reticent to admit that they chose the white candidate, for fear of seeming prejudiced (Cose, 2006). In the Douglas Wilder example, an exit poll (conducted as voters leave the booths on Election Day) showed Wilder winning by 10 points, when he actually won by 0.4, a discrepancy that was no doubt exacerbated by the fact that this exit poll³ used interviewers and face-to-face questioning, rather than the anonymous written ballot used by most exit polls (Keeter & Samaranayake, 2007).

³Mason-Dixon poll commissioned by WJLA-TV of Washington. Brad Coker (President of Mason-Dixon) said "We're pretty sure white voters were telling us they voted for Wilder when they really voted for Coleman. . . . It was a racial issue, and because it was racial, they probably didn't want to admit it to a pollster standing in the parking lot of the elementary school where their kids go." Rosenthal (1989).

Keeter and Samaranayake point out that the other two, non-bi-racial, state-wide races were accurately predicted by the same exit poll. As it is no longer socially acceptable to express prejudice explicitly, it is tempting to assume it is a thing of the past. But it is apparent that explicit prejudice may be reduced without an equal reduction in implicit prejudice.

Explicit and Implicit Prejudice

“There is a certain amount of racism that exists in the United States - whether it's conscious or not it's true”

Geraldine Ferraro, the first woman to run as a major-party vice presidential candidate, cited in Nagourney (2006, p. 1).

Explicit prejudice operates at a conscious level and is expressed overtly (Wilson, Lindsey, & Schooler, 2000). It is this type of prejudice which is normally picked up by self-report techniques in a questionnaire or survey. Implicit prejudice, on the other hand, is thought to be unintentionally and automatically activated (Greenwald & Banaji, 1995), often by the mere presence of a target individual or attitude object (Chen & Bargh, 1997). This prejudice can operate at an unconscious level so that a person is unaware that their actions may be being affected by prejudice. This type of prejudice can only be measured implicitly, by, for example, the *evaluative priming technique* (see Fazio, Jackson, Dunton & Williams, 1995), or the *Implicit Association Test* (Greenwald, McGhee, & Schwartz, 1998). In the evaluative priming technique (Fazio et al., 1995), a prejudice-relevant target is presented for such a short period of time that the participant is barely aware that he or she has seen the target and that any prejudices associated with that target have been activated. The participant then makes various judgements concerning whether a presented evaluation

term is positive or negative. When primed with a target about which the individual holds negative implicit attitudes, the 'positive' judgement will take longer than the 'negative', revealing the unconscious prejudices triggered by the target. This approach can utilise subliminal priming (Wittenbrink, Judd, & Park, 1997) or supraliminal priming (Fazio et al., 1995), the latter of which leads to a greater likelihood the purpose of the experiment will be transparent. The Implicit Association Test (IAT, Greenwald et al., 1998) does not attempt subliminality, instead openly assessing the strength of an association between an attitude object and an evaluation by comparing response times for categorising attitude objects when the target categories are paired with the positive versus the negative responses on a keyboard. This technique is described in greater detail in Chapter Five.

Because the expression of explicit prejudice can be controlled to a far greater degree than the expression of implicit prejudice (Dambrun & Guimond, 2004), measures of one do not necessarily correlate with measures of the other (Blair, 2001; Dovidio, Kawakami, & Beach, 2001; Wilson, Lindsey, & Schooler, 2000). The two types of prejudice are particularly unlikely to correlate in situations when it would be considered socially unacceptable to express prejudice (Dovidio & Fazio, 1992). Given that explicit and implicit prejudice both draw on the same underlying knowledge structure (Dambrun & Guimond, 2004), it appears likely that the lack of correlation between the two is a result of individuals controlling their explicitly expressed prejudice as a result of social pressure, but being unable to control their true prejudice, as occurs in the example of the Bradley Effect, which might more accurately be described as covert racism (Citrin, Green & Sears, 1990).

This apparent lack of correlation between real attitudes and expressed attitudes highlights an important flaw in research using methodologies whose validity relies on

participants providing honest answers. While this flaw has been addressed to some extent in prejudice research by the capacity of the aforementioned implicit measures to assess an individual's true level of prejudice, such implicit measures are unsuitable for examining many influential factors which are relevant in research into the reduction of prejudice. This thesis attempts in some part to address this issue.

Prejudice reduction

Prejudice has been a research focus for almost a century (Fiske, 1998). The mechanisms underlying prejudice have been explored, as have the ways in which prejudice can be expressed, and the consequences of this expression. A pioneer in the field, Gordon Allport (1954) proposed that, in order to reduce cognitive load in processing information about individuals, humans tend to categorise and stereotype, to group people together and attribute traits to a group rather than to an individual. This in turn can lead to ingroup favouritism and outgroup discrimination: prejudice. Many methods exist with the potential to reduce such prejudice, such as the provision of stereotype-disconfirming information (Weber & Crocker, 1983), the use of social influence or peer pressure (Tan, Tan, Avdeyeva, Crandall, Fukushi, Nyandwi, Chin, & Wu, 2001), or the increase of perspective taking to encourage empathy towards the outgroup (Galinsky & Moskowitz, 2000). This thesis, however, will focus on research resulting from Allport's (1954) claim that exposure to members of other groups can dispel the notion that individuals from one group all share the traits attributed to the group, and can improve attitudes towards that group. This notion, the 'contact hypothesis', remains one of the most influential ideas existing today for the reduction of intergroup prejudice (Oskamp & Jones, 2000).

The Contact Hypothesis

Contact between groups has, for a long time, been considered possibly the most effective means of reducing prejudice, and more than 50 years of research have supported its continued use (see Pettigrew & Tropp, 2006, for a meta-analytic review). Although research into the idea that contact can reduce prejudice began in the 1930s, Allport's (1954) work has been credited with formulating the theory in the most influential manner. The contact hypothesis specifies certain optimal conditions, under which contact is hypothesised to improve attitudes. Allport underlined that contact alone was not enough, indeed, under the wrong conditions, contact could *increase* prejudice (Baker, 1934; Brookes, 1975; Meer & Freedman, 1966). I will now set out the conditions proposed by Allport (1954/1979) under which contact is hypothesised to reduce prejudice.

Allport's conditions - Cooperation

Importantly, contact must be co-operative; groups must work together to achieve a common goal or goals. Much work has supported the need for groups to co-operate towards something beneficial for both, in contrast with competition for limited resources, a situation often leading to intergroup hostility (Realistic Group Conflict Theory, Campbell, 1965; Sherif & Sherif, 1953; 1966). In World War II, black and white soldiers, although officially segregated, relied on each other in battle, and white soldiers who had experienced integrated combat subsequently reported improved racial attitudes (Singer, 1948; Stouffer, Lumsdaine, Lumsdaine, Williams, Smith, Janis, Star, & Cottrell, 1949b). Students who experienced interdependent learning, requiring cooperation between students of various ethnicities to teach each other small parts of the whole to be learnt, showed greater liking of their outgroup peers (Jigsaw Classroom, Aronson & Bridgemen, 1979)

Allport's conditions - Common goals

This condition is inseparable from the last; groups must want the same thing, but rather than competing for it, a spirit of cooperation should encourage groups to work together to achieve what benefits them both. The work most quoted in the exploration of the combination of these first two conditions is that of Sherif, Harvey, White, Hood, and Sherif (1961). Sherif et al. grouped boys at a summer camp, which alone was sufficient to instigate intergroup tension (Gaertner, Dovidio, Banker, Houlette, Johnson, & McGlynn, 2000). However, when Sherif et al. introduced competition between the groups for a 'limited resource' of, for example, a prize of penknives for the winning team, full-scale conflict emerged, with name-calling and even armed raids on outgroup territory (Gaertner et al., 2000). Contact alone did not serve to reduce this conflict, and in fact exacerbated it, but once the groups were required to co-operate to achieve a series of common goals, relations between the groups improved. To avoid perceptions of being entitled to a larger portion of the profit than another group, groups should contribute equally to obtain that profit (Slavin, 1978). However, if groups cooperate on the same task, a threat may result to the distinctiveness of group members' social identity (Tajfel & Turner, 1979) which may then lead to increased intergroup bias (Brown & Wade, 1987; Hewstone & Brown, 1986). A fine line must therefore be trodden; groups must cooperate, but on distinctly different tasks, which must be of equal value, in order to achieve their common goal.

Allport's conditions - Equal status

Those involved in the encounter must be of equal status. Sims and Patrick (1936) reported that contact *increased* prejudice among white college students who moved from the North to the South of the United States to study, due to the fact that

they encountered only lower status black men and women, and all the students and faculty were white. The effect was strengthened by the fact that the social norm in this all-white environment was predictably one of somewhat negative attitudes towards blacks. Contact with professional or skilled black men was found to lead to improved attitudes among army veterans (Mackenzie, 1948). However, even if equal status is achieved within the contact encounter, attitudes in general may not improve if there is a large difference in status in society in general (Minard, 1952), as societal differences in status tend to cause tension during the experimental contact situation (Moaz, 2000). It is unfortunate that those groups for whom improved intergroup relations would be most beneficial tend to be those who suffer from low status in society, thereby decreasing the effectiveness of contact with higher-status majority groups.

Allport's conditions - Social sanction

As in the previous example, if society in general does not support the contact, it is difficult to create a situation where contact can improve attitudes. Efforts to desegregate housing can be beneficial due to their backing by authorities (Deutsch & Collins, 1951), and when school children are encouraged by teachers to work together with black students, it can lead to increased liking between black and white children (Walker & Crogan, 1998). Indeed, in some situations, the existence of social approval towards the outgroup, even in the absence of contact, can lead to improved attitudes (e.g., Tan et. al., 2001).

Later work

In the years since Allport's pioneering work, considerable effort has gone into clarifying, testing, and expounding on the contact hypothesis, and a massive amount of research evidence has been accumulated to support the reliable effect of intergroup contact in decreasing prejudice under the right conditions. A number of theoretical

models and reformulations of the contact hypothesis have appeared (e.g., Brewer & Miller, 1984; Brown & Hewstone, 2005; Hewstone & Brown, 1986; Pettigrew, 1998) and research continues today to expand knowledge about when, and how, contact works to improve intergroup attitudes. Additional 'optimal conditions' were proposed for extending the theory. Cook (1962) suggested that contact should be between similar individuals, who disconfirmed stereotypes, and should involve close physical proximity with the potential for forming acquaintances. Blanchard, Adelman, and Cook (1975) added the requirement that the interaction have positive outcomes and that an equal number of ingroup and outgroup members should be present. Contact should be voluntary, and take place between groups with favourable prior relations (Tajfel, 1978), against a prosperous economic background in order to reduce potential intergroup competition (Wagner & Machleit, 1986), and should be between young, higher class, well-educated individuals (Williams, 1964), who do not already hold extremely negative attitudes (Ben-Ari & Amir, 1986) and are low in authoritarianism and high in self-esteem (Stephan & Rosenfield, 1978). In other words, 'optimum' contact was almost impossible to achieve (Hewstone, 1996), making the contact hypothesis essentially unfalsifiable. Researchers have subsequently suggested that no condition should be deemed essential, but that each condition, if present, would facilitate improved attitudes between groups (Brown & Hewstone, 2005; Pettigrew, 1998).

Cross group friendship

One idea which did take hold was that of the importance of intimate, rather than casual, contact (Amir, 1969). Pettigrew (1998) suggested that a condition which should be regarded as essential to prejudice reduction is that of friendship potential. Cross-group friendship, or friendship with an outgroup member, usually involves

many of the suggested facilitating conditions of contact, such as intimacy, acquaintance potential, voluntary contact, close physical proximity, contact between similar individuals, and positive outcomes (in that contact with a friend is pleasant). In addition, friendships involve empathy, sharing of resources, and pride in the other's achievements (Aron, Aron, Tudor, & Nelson, 1991). Friendships necessarily involve long term contact, and enable individuals to discover stereotype-disconfirming information about each other (Rothbart & John, 1985). Of course, friendships involve reciprocal affection (Rubin, Bukowski & Parker, 1998), and the contact is personalized, collaborative, of equal status and trusting (Pettigrew, 1998). Friends also typically show loyalty and tend to contribute to self-validation, and this gives cross-group friendship the potential to increase self-esteem and perceived socioemotional support (Hartup & Stevens, 1997). Much evidence supports the prejudice-reducing benefits of individuals forming direct friendships with outgroup members (e.g., Khmelkov & Hallinan, 1999; Schofield, 1995; Stephan, 1999; for a full review, see Turner, Hewstone, Voci, Paolini, & Christ, 2008). An extensive meta-analysis performed by Pettigrew and Tropp (2006) showed that the 154 studies of intergroup contact which included a measure of cross-group friendship as a predictor variable had a significantly higher effect size ($r = -.27$) than the 1,211 which did not ($r = -.18, p < .001$).

Indirect cross group friendship

Recently, research has shown that a similar effect can be obtained indirectly, through individuals simply being aware that other ingroup members have friends in the outgroup ('indirect cross-group friendship' or 'extended contact'; Wright, Aron, McLaughlin-Volpe, & Ropp, 1997). This vicarious contact can have an effect on attitudes, above and beyond the effect of direct cross-group friendship, as a result of

referent informational influence (Hogg & Turner, 1987). Referent informational influence is the way in which individuals gather information about social norms through watching other ingroup members and use that information to discover how to behave. It is probable that Allport's 'social sanction' condition facilitates prejudice as a result of this reliance on social norms. Indeed, recent research by Turner, Hewstone, Voci, and Vonofakou (2008) confirmed that perceived norms are a key mediator of the impact of indirect contact on outgroup attitudes. When individuals perceive that another ingroup member is friends with an outgroup member, they can be influenced to see it as socially acceptable to be friends with outgroup members in general, and show more positive attitudes to that outgroup (Aron, Melinat, Aron, Vallone, & Bator, 1997; Liebkind & McAllister, 1999; Paolini, Hewstone, Cairns, & Voci, 2004; Turner, Hewstone, & Voci, 2007; Wright et. al, 1997).

I will now give consideration to the methodology of contact research. First I will describe the range of possible paradigms and their associated weaknesses. I will then discuss self-reports, and their weaknesses. Finally I will move on to describe proposed methods for validating self-reports of contact, and outline how this thesis provides a detailed examination of the validity of self-report data using a variety of methods.

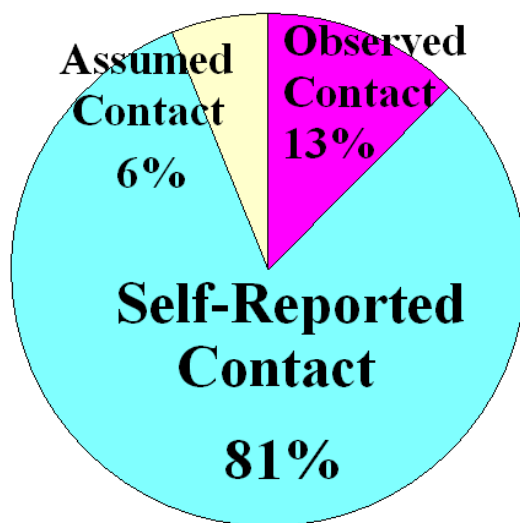
Methods used in contact research

There are, potentially, many paradigms that could be used in the study of intergroup contact. However, the most common by far has been the self-report method, which requires research participants to answer questions about themselves, and report their own level of the variable of interest. The vast majority of intergroup contact research has relied on self-report measures. As is illustrated in Figure 1.2, Pettigrew and Tropp's (2006) meta-analysis of approximately 200, 000 participants

found that 81% provided self-report measures of contact. Disturbingly, Pettigrew and Tropp found that self-report studies yielded significantly smaller effect sizes than samples with directly-observed contact ($r = -.246$ vs. $r = .210$). However, it remains the case that alternative paradigms such as *observational studies* (e.g., Cummings & Lambert 1997; Espenshade 1997; Fossett & Kiecolt, 1989; Giles, 1977, Hood & Morris 1997; Schofield & Sagar, 1977) and *experimental studies* (e.g., Wilder, 1984), although valuable, cannot provide all the information that can be obtained using self-reports.

Figure 1.2

Methods used in contact research, as reported by Pettigrew and Tropp, 2006



Observational studies

Observational approaches to assessing contact have commonly involved examining the percentage, in the respondents' area of residence, of non-ingroup members (Cummings & Lambert 1997; Espenshade 1997; Fossett & Kiecolt, 1989; Giles, 1977) or the percentage of members of various minority groups (Hood & Morris 1997). These methods, however, only measure *opportunity* for contact (Wagner, Hewstone, & Machleit, 1989), and do not ascertain whether or not that opportunity is taken up. As Pettigrew and Tropp (2006) have pointed out, one cannot infer contact from opportunity for contact, and mere proximity is not sufficient to ensure social contact (Festinger & Kelley, 1951). Taylor, Dubé and Bellerose (1986) defined this phenomenon as "illusory contact", ostensible integration which does not in reality lead to intergroup interaction, due to individuals continuing to favour ingroup members when choosing interaction partners. Dixon and Durrheim (2003), observing a desegregated beach in post-apartheid South Africa, found clear patterns of racial isolation still appearing despite the presence of members of multiple different racial groups on the beach. Furthermore, these 'macro-level' observational measures of contact do not take the *quality* of the contact into account. It must be ascertained whether, or to what extent, the relevant conditions for beneficial contact (Allport, 1954) are met.

'Micro-level' observational studies (e.g., Campbell, Kruskal, & Wallace, 1966; Clack, Dixon & Tredoux, 2005; McCauley, Plummer, Moskalenko, & Mordkoff, 2001; Schofield & Sagar, 1977) record *actual* contact in a certain situation such as seating preferences in a cafeteria. As the situation (physical, temporal, political, historical etc) is known, this method permits the researcher to ascertain to some extent the presence of Allport's conditions. For example, Schofield's research (Schofield,

1979; Schofield & Sagar, 1977) was conducted in a mixed school chosen specifically for its adherence to Allport's (1954) criteria. The level of integration in the chosen location can be estimated by measuring the number of mixed-race adjacencies (Dixon, Tredoux, Durrheim, Finchilescu & Clack, 2008; Campbell et al., 1966; Schofield & Sagar, 1977) or social clusters (McCauley et al., 2001). However, this method of contact measurement is only of use with a view to examining the contact existing in general in a location or organization (for example, for evaluating the success of a desegregation intervention), and has not been used as an individual difference measure. Attitudes of those for whom contact is being observed have not been measured. Because these studies have not related individual-level measures of contact to individual-level measures of attitude for the same participants their ability to assess the effectiveness of contact is limited. It might theoretically be possible to record the seating preferences of specific individuals over a period of time, in the manner employed by observational studies, but it would be extremely time-consuming and inefficient as a measure of an individual's level of contact. Furthermore, these studies are limited to measuring contact in only one situation, and the contact observed may not continue beyond the observed situation.

Experimental studies

A more robust method for research on the contact hypothesis may be that employed in *experimental studies* which allow us to test the causal effect of contact on attitudes. Bettencourt, Brewer, Croak and Miller (1992) employed the minimal group paradigm (in which groups are created randomly, based ostensibly on some irrelevant personality trait, Tajfel, 1970) and engineered conditions of intergroup cooperation or competition. Intergroup cooperation eliminated ingroup favouritism and biased ratings of members of other teams, in comparison with ingroup

competition. Wilder (1984) manipulated contact by providing a specific contact experience with a confederate to half the participants, and comparing them with a control condition. This approach provides strong evidence for the causal effects of contact because participants are randomly assigned to conditions. However, the experimental method cannot examine the various nuanced forms that contact can and does take outside the experimental location. Such contact may be of a very different nature to the contact manipulated and studied, and may still contribute a great deal to the attitudes of the participants.

It is necessary to explore all contact experienced by a person, and to ascertain to what extent this contact fulfils Allport's 'key conditions', in order to obtain the information relevant to the contact hypothesis. It seems indeed implausible that experimental manipulations could effectively examine all of Allport's 'key conditions' for contact. Additionally, experimental manipulations are necessarily limited in duration and scope. Thus it is difficult to reach conclusions from such manipulations about the cumulative impact of extended contact with many diverse individuals. Finally, it would not be feasible to manipulate true cross-group friendship in a laboratory, and as this appears to be an important form of contact, the experimental method is limited in potential.

Correlational studies

Correlational studies obtain data on two or more variables in order to discover a relationship between them. Data can be gathered through observation, as described above, but is more commonly collected through the use of self-report questionnaires. Self-reports, as I have described, involve asking the participant to report on their own levels of the dependent variables in question. This method makes it possible to ascertain the quality, as well as the quantity, of contact, and the extent of friendship

existing with outgroup members. This is essential in order to collect the information necessary to elucidate the effects of cross-group friendship on attitudes. Participants can also report on the amount of contact their friends have, which gains the researcher a measure of a participant's indirect contact, something which would be very difficult to measure without self-reporting techniques.

Weaknesses of the self-report methodology

While self-reports remain the most efficient and useful method of eliciting the information necessary to support many hypotheses regarding the contact hypothesis, it must be acknowledged that the self-report methodology itself suffers from a number of weaknesses. Researchers in both personality and social psychology have long questioned the validity of self-reports. Individuals cannot always be relied upon to report accurately about themselves, and many things can influence how people respond to questions on questionnaires. I will discuss lack of adequate self-knowledge for answering self-report questions accurately, factors which may bias self-report responses, and finally problems with the one-way nature of self-reports for examining constructs which are not one-way in nature.

Lack of adequate self-knowledge

It is likely that many participants might simply not have access to the attitudes or feelings they are asked to judge (Haddock, Rothman, Reber, & Schwarz, 1999). For instance, individuals who have little experience with an issue or have seldom thought about it may have less accessible information as to the content, strength and certainty of their attitudes concerning that issue (Basilli, 1993, 1996a, 1996b). Participants often find it difficult to recall the details of what often amounts to a set of low-salience behaviours (Price & Zaller, 1993), and responses are therefore often based on guesswork. Even more problematically, because participants are

automatically inclined to search their memories for examples of the behaviour about which they must report, their guesses tend to be systematically biased towards over-reporting their levels of this behaviour (Bechtel, Achelpohl, & Akers, 1972). The ability to engage in introspection in the manner required for self-reporting can prove somewhat limited as people tend to suppress undesirable thoughts, feelings or memories (Wilson & Dunn, 2004). In fact, introspection can lead to illusionary self-knowledge, or false confidence in self-reports (Wilson & LaFleur, 1995). Wilson and LaFleur (1995) found that deeper, as compared to more shallow, introspection can even lead to distortion of self-reports, and therefore to reduced validity.

Even where self-knowledge exists, therefore, this may not be as accurate as the respondent may assume. A great deal of work exists to support the suggestion that people tend to have inaccurate self-knowledge. When evaluating their own competence, for example, the correlation between a person's perception and the reality is "often modest to meagre, if it appears at all" (Dunning, 2005, p. 4). Individuals have been shown to be poor at judging their own competence in many areas: intelligence, self-ratings of which correlated no better with a standardised test than did the ratings made by strangers after watching the individual reading a standard weather report for 90 seconds (Borkenau & Liebler, 1993), ability to tell when someone is lying (no correlation, DePaulo, Charlton, Cooper, Lindsay, & Muhlenbruck, 1997), or ability to express emotion to others successfully (no correlation, Riggio, Widaman, & Friedman, 1985). Adolescent boys do not accurately judge their own level of knowledge about condom use (Crosby & Yarber, 2001), and naval officers' self-ratings of leadership ability do not predict how often a superior officer recommends them for promotion (Bass & Yammarino, 1991). Even purported experts show a surprising lack of perception of their own level of competence;

surgical residents' self-ratings of their surgical skill did not predict their score on an objective test (Risucci, Tortolani, & Ward, 1989), nurses' self-rated proficiency at basic life support tasks did not correlate at all with evaluations by their seniors (Marteau, Johnston, Wynne, & Evans, 1989), and physicians' self-rated understanding of thyroid disorders showed no correlation with external evaluations (Tracey, Arroll, Richmond, & Barham, 1997).

Sources of error in self-insight – Overconfidence

In general, errors in self-knowledge tend towards overconfidence. A common pattern found in research into the accuracy of self-knowledge is that more than half the participants indicate that their skill/knowledge/moral character/social status is above average (Dunning, 2005, p. 6). This is, of course, a statistical impossibility! In a large survey of high school seniors (College Board, 1976-1977) 60% rated themselves as above average in athletic ability, while only 6% rated themselves below average, relative to other high school seniors. In leadership ability, 70% said they were above average, and 2% said they were below. Regarding 'ability to get along with others', almost all said they were at least average, 60% said they were in the top 10% of their peers, and 25% said they were in the top 1%! Ninety-four percent of college professors believe they do above average work, another statistical impossibility (Cross, 1977). People think they are more likely than others to encounter positive life events, and less likely than others to encounter negative ones (Weinstein, 1980). People think they are less susceptible to the flu (Larwood, 1978), and better at driving (Marottoli & Richardson, 1998) than the average person. Motorcyclists think they are less likely than most bikers to cause an accident (Rutter, Quine, & Albery, 1998), bungee jumpers (but not their families!) think they are less likely to meet harm than the typical bungee jumper (Middleton, Harris, & Surman, 1996), and business

executives think their firm is more likely to succeed than the typical firm in their business sector (Cooper, Woo, & Dunkelberg, 1988; Larwood & Whittaker, 1977).

It could be argued that this pattern reflects a tendency to underestimate others rather than to overestimate oneself, but a great deal of research using objective measures for comparison has found that this is not the case. Lawyers overestimate their chances of winning a particular case (Loftus & Wagenaar, 1988), negotiators overestimate the likelihood that the other side will accept their final offer (Neale & Bazerman, 1983), stock pickers are overconfident that they have chosen winners when more often than not the stocks they buy tend to do less well than stocks they sold to buy them (Odean, 1998), and surgical trainees overestimate the accuracy of their diagnoses using x-rays of possible fractures (Oksam, Kingma, & Klasen, 2000). Even in the field of psychology (where we ought to know better) clinical psychologists are overconfident in their inferences about a client after reading case study materials (Oskamp, 1965). Even when people say they are 100% sure, they are still wrong 20% of the time (Fischhoff, Slovic, & Lichtenstein, 1977). Physicians who were 88% sure that their patient has pneumonia, turned out to be wrong 80% of the time (Christensen-Szalanski & Bushyhead, 1981). And finally, in terms of self-knowledge about self-knowledge, most people will say they are uniquely able to avoid the overconfidence and bias that other people demonstrate so pervasively (Friedrich 1996; Pronin, Lin, & Ross, 2002).

Why does such pervasive overconfidence occur? Among other researchers, Taylor and Brown (1988) argue that this bias is essential for motivation, good moods, creativity, generosity, and a good life. This view is not shared by all, however (see, e.g., Colvin & Block, 1994). Dunning (2005) suggests that the incompetent are unable to judge themselves accurately, as they do not know what competence looks like. To

judge oneself accurately at a skill like athletic achievement, all that is required is good enough eyesight to know the level the best are able to achieve. Self-ratings of athletic achievement are therefore often accurate ($r = .47$, Mabe & West, 1982; see also Felson, 1981). In areas which require expertise such as performing heart transplants or fixing a carburettor, lack of ability to produce even a plausible response will be sufficient to inform an individual of his or her incompetence (Dunning, 2005). However, accurate judgements of the self are only possible if the judge knows what constitutes the maximum possible level of the attribute in question (Kruger & Dunning, 1999). If given a score of 97, one might assume one had done well until one found out the maximum score was 700. In many areas, the judge's perception of what knowledge is possible is likely to be little more than what constitutes the judge's own knowledge, or in other words, they cannot know what they do not know, and this failure to see accurately the gaps in their knowledge results in overestimation of that knowledge. In the context of contact research, this may imply that those with low outgroup contact will assume their level of contact is normal, and therefore overestimate it.

The cause of inaccurate self-knowledge in the highly competent is usually a tendency to *underestimate* their own competence, assuming that their own knowledge is shared by a high proportion of people (the false consensus effect; Marks & Miller, 1987; Ross, Greene, & House, 1977). The highly competent tend to overestimate their peers by 26% (Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2004). Once something is known, it feels obvious (Fischhoff, 1975). Once an event takes place, hindsight makes it appear predictable (Christensen-Szalanski & Willham, 1991). People who know an answer assume that a higher-than-accurate percentage of others know it also (Fussell & Kraus, 1991, 1992; Nickerson, Baddeley, & Freeman, 1987).

This may lead those who have high contact with the outgroup to overestimate the extent to which this is normal, and therefore to under-estimate their own contact. The two effects combined would produce a closer-than-realistic similarity between the self-reported contact of high- and low-contact individuals, a relevant problem in contact research.

Sources of error in self-insight – Biased information processing

Individuals tend to selectively process information, ignoring feedback that challenges their positive self-conceptions (Brown & Dutton, 1995; Sedikides & Strube, 1997; Taylor & Brown, 1988). This motivation to maintain or enhance a positive self-concept has been comprehensively explored by Sedikides, Green and colleagues, who conclude that information processing occurs more deeply for self-affirming information than for self-threatening information (Green & Sedikides, 2004; Green, Sedikides & Gregg, 2008; Sedikides & Green, 2000, 2004). Information processing begins in the same way for both positive and negative personal feedback, but while positive feedback leads to a search for confirmatory evidence in episodic autobiographical memory, negative feedback does not, and is therefore less likely to be recalled at a later date, or incorporated into the self-view (Sedikides & Green, 2000).

Relatedly, once a suggestion is made, people tend to try to confirm that suggestion (Dunning, 2005, p. 47). Evidence that confirms the hypothesis is emphasised and evidence that disconfirms it is de-emphasised. The answer, and the information sought (or distorted) to confirm it, depend on the original hypothesis (Brownstein, 2003), which can come from the way a question is posed (Dunning & Parpal, 1989; Kunda et. al., 1993), or even from an expectation, a perception of what

the experimenter expects, or what the participant wants to be the truth (Pyszczynski & Greenberg, 1987).

This is particularly relevant to the self-report methodology, and relates to a response bias referred to as ‘acquiescence’ (Paulhus, 1991), discussed in more detail below. If an item on a questionnaire asks ‘are you outgoing?’, the respondent would then try to think of examples when he or she acted in an outgoing manner, and this would exaggerate his or her response. If asked ‘are you shy?’, instead, he or she would be more likely to emphasize shyness. The impression a person has of themselves after a question framed in one way is different from the impression they hold after a question framed in the opposite way (Kunda, Fong, Sanitioso, & Reber, 1993; Snyder & Cantor, 1979). For this reason, questions of this sort are often asked from both points of view, with some questions ‘reverse-scored’. However, in the context of contact research, there is no way to ask a participant how much contact he or she has with outgroup members, without causing him or her to think about times when contact occurred, and therefore probably to emphasize them. Even if participants could be asked how much *ingroup* contact they had, so that this item could be reverse scored, the reverse score would not provide information on one particular outgroup, but all outgroups. One solution might be to reverse-score a question regarding contact *avoidance*, but this question is likely to elicit social desirability effects. This propensity of individuals towards attempting to confirm the first suggestion they get is strongly related to another weakness of self-reports relevant to contact research; that of potential *context effects*.

Context effects

When a person is asked a question about the city 'Sydney', and then several questions later is asked 'What is the capital of Australia?' that person is likely to erroneously answer 'Sydney', as a result of its being recently activated in the brain (Kelley & Lindsay, 1993). This example highlights the effect context can have on self-reports. It has long been apparent that attitude measurement, for example, is extremely context-dependent (Schwarz, 1999). Preceding questions have frequently been shown to influence the answers to subsequent ones (Schwarz & Sudman, 1992). Schwarz and Bless's (1992a) inclusion-exclusion model outlines how question-order affects responses. They suggest that an evaluative judgement requires both a target stimulus and a standard by which to evaluate it. A mental representation of either the target stimulus or the standard of comparison cannot realistically be based on all available information and must necessarily be affected most by the most easily accessible information.

Two important determinants of accessibility are the frequency with which information is accessed and how recently it has been accessed (for a review see Schwarz & Strack, 1999). 'Chronically accessible information', the most frequently accessed, and 'temporarily accessible information', for example that information most recently accessed in order to answer a previous question, are therefore most likely to affect responses. But how responses will be affected depends on whether the information forms part of the target stimulus or of the standard of comparison. If it forms part of the target stimulus, it results in 'assimilation effects', or in other words, positive information results in positive judgements. Salancik and Conway (1975), for example, used 'priming' questions to manipulate the salience of either pro- or anti-religious behaviours. When they primed pro-religious behaviours they found

participants subsequently reported more positive attitudes towards religion generally than when anti-religious behaviours were primed. If the information forms part of the standard of comparison, however, this effect is reversed, and the information results in 'contrast effects' (i.e., positive information results in negative judgements). Strack, Schwarz and Gschneidinger (1985) found that judgements of current well-being were influenced by preceding questions concerning recall of past and current life events. Recalling current life events generally influenced responses in the direction of the valence of the event. Recall of a positive event from the present led to more positive reports of current well-being. If the event was in the past, however, the effect was reversed, as the event was used as a standard of comparison. Recall of a past positive event would cause participants to evaluate their current well-being comparatively negatively.

Schwarz, Bless, Strack, Klumpp, Rittenauer-Schatka and Simons (1991) found another way in which information can be used. If information is difficult to recall, it can lead participants to believe it should be weighted less heavily in subsequent evaluative judgements. Schwarz et al. found that if participants were asked to recall 12, rather than 6, examples of assertive behaviours, they subsequently described themselves as less assertive. Haddock et al. (1999) similarly asked participants to think of three (easy) or seven (hard) arguments either supporting or opposing their own views on doctor-assisted suicide. Responses of participants who held extreme attitudes were found to be resistant to context effects, but participants who held moderate attitudes prior to the task subsequently reported their attitudes to be more intense, more personally important and more certain when they had been asked to think of either three supporting arguments or seven opposing arguments, but not in the other conditions.

It is apparent that context affects attitudinal responses, at least when there is a degree of subjectivity in the questions, but it remains to be seen whether contact measures are also affected by such context effects. Contact is very rarely studied in isolation and contact measures are therefore in most cases presented along with other measures. It is therefore important to test the susceptibility of self-report measures of contact to context effects. If self-report measures of contact are not resistant to context effects, this would call into question the validity of much research on contact which has not taken context effects into account when presenting contact measures after or alongside other measures.

In theory, whereas attitudes are quintessentially subjective and evaluative responses, and hence one would expect them to be subject to context effects, self-reports of contact should be factual and objective (at least when assessing dimensions of contact such as its quantity, duration, or number of friends). Yet, it is plausible that recalling encounters with an outgroup might increase the accessibility of such experiences and thereby affect self-reports of contact (Schwarz & Strack, 1999). It is equally plausible that if recalling such an event is difficult (Haddock et al., 1999; Schwarz et al., 1991), or if more events can be recalled from some time ago than from recent times (Strack et al., 1985), participants will be led to assume that their own contact is relatively low, and as such reported contact will be deflated. If found, either of these effects would significantly undermine contact research which has relied on self-reports and has not taken context effects into account. Chapter Two in this thesis examines the effect of context on self-reports of contact.

Response biases

Method variance, variance attributable to the measurement method rather than the construct of interest (Fiske, 1982), can confound empirical results, leading to potentially misleading conclusions (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). This variance, or measurement error, can take a *systematic* or an *unsystematic* form. The main form that error variance can take is that of a response bias (Podsakoff et al., 2003). Paulhus (1991) points out three primary *systematic* response biases: *socially desirable responding* (SDR), *acquiescence*, and *extreme responding*. *Socially desirable responding* involves participants presenting an idealised version of themselves. I will consider this important response bias in detail below. The second primary response bias identified by Paulhus (1991) is *acquiescence* (or, more rarely, its opposite); people have a tendency to agree (or disagree) with whatever statements are made in the questions (Dicken, 1963; Lentz, 1938), although this can be controlled for to a certain extent by using negative wording in some questions. Thirdly, Paulhus claims, people use scales in different ways; some respondents are prone to an *extreme response bias*, tending to exaggerate their responses and choose the extremes on the scale to reflect even their more moderate answers, and some do the opposite, revealing a *moderacy response bias*, and tending to avoid the extremes of the scale even for their strongest opinions.

Many other factors can also affect the validity of self-reports. Responses can be influenced by *unsystematic* biases such as random responding, carelessness and misunderstanding (Piedmont, McCrae, Riemann & Angleitner, 2000; Tellegen, 1988). Lack of adequate self-insight has already been discussed, and the resulting uncertainty over how to respond can exacerbate other response biases as, without a definite answer to give, participants may be forced to rely more strongly on such strategies as

acquiescence or SDR. Norm-based heuristics, for example, are particularly likely to influence self-ratings on items for which clear social norms exist, but about which the individual lacks clear self-knowledge (Konstabel, 2006).

Socially desirable responding

Socially desirable responding is the response bias that has received most research attention. Participants tend to present themselves in an unrealistically positive light, by enhancing positively perceived characteristics (Phillips & Clancy, 1972) or reducing negatively perceived characteristics (Knudson, Pope, & Irish, 1967; Sigall & Page, 1971; Wiseman, 1972). This can occur either because participants deceive themselves into believing that this is how they really are (self-deception; Paulhus & John, 1998), or because they wish to control how they appear to others (impression management; Barrick & Mount, 1996; Paulhus, 1984, 1986). Impression management may be a purposeful distortion to avoid social disapproval (Crowne, 1979), and may therefore be susceptible to situational influences in that it might be exacerbated in situations in which motivation exists to present a desirable image. In contrast, self-deception is a more stable trait, linked to optimism (Scheier & Carver, 1985), self-esteem (Winters & Neale, 1985), and self-confidence (Paulhus, 1988), whereby participants simply see the best in themselves at all times, irrespective of the situation. This response bias was identified over half a century ago (Edwards, 1953) and still remains a problem for research today. Particularly when self-reports assess a socially sensitive issue, such as attitudes towards other races, results can be distorted (Sigall & Page, 1971).

These response biases have been pinpointed in the context of personality research, but there is no reason why they should not also be present in contact research. Although, as I have reported, the vast majority of research on the contact

hypothesis has used self-report measures, there has been no research evaluating the validity of self-reports of contact (although Vonofakou, 2008, found that SDR was not influential enough to remove the relationship between direct and indirect cross-group friendship and attitudes). Yet it can be argued that SDR is even more of a problem in contact research than in the field of personality generally, as many personality traits such as extroversion or talkativeness have no negative or positive connotations, whereas items measuring intergroup contact and, especially, expression of prejudice (which, today, is socially unacceptable), are almost universally valenced in one direction or another, and the issues in question are more sensitive, with stricter norms.

Non-dyadic methodology

A more complex problem that tends to typify studies which use self-reports has been pointed out by Aboud, Mendelson, and Purdy (2003). They state that existing research examining cross-group friendship in particular has for the most part relied on one-way friend nominations. Although this is an efficient and convenient way to collect self-reports, it provides a limited index of close relationships as it does not necessarily demonstrate a friendship, or indeed actual contact (Rubin, Bukowski, & Parker, 1998). Although one person may nominate another as a friend, the nominee may not share the perception, revealing that no real friendship exists as friendship is very much a two-way phenomenon. The one-way self-report is vulnerable to the effects of individuals claiming that they have friends in the outgroup who are really nothing more than casual acquaintances or indeed people the individual has met but briefly. Claiming those people as friends would show no more than a desire to befriend outgroup members, and a link between this desire and a positive evaluation of outgroup members is not surprising and does not actually provide any data of

theoretical interest for the purposes of examining the relationship between cross-group friendship and attitudes (Aboud et al., 2003, p. 165).

Kenny, Kashy, and Cook (2006) point out that most social science research is focused on the study of individuals. They explain that we live in an individualistic culture (Smith & Bond, 1994), and that psychology emphasizes the individual before higher levels of analysis (Bond & Kenny, 2002). They point out that social science research relies for the most part on standard statistical methods which were developed for the analysis of individuals, and which therefore assume that data are independent - that once variation caused by the independent variable is controlled for, the data from each individual in a study will be unrelated to the data from every other individual in the study (Kenny et al., 2006). They argue, however, that a good deal of intergroup and interpersonal research that has been analysed in a unilateral manner is in fact dyadic in nature, and has therefore been analysed inappropriately. Any self-report which asks one person about their feelings towards, attitudes about, or contact with another person, or a group, is dyadic in nature, and ought to be analysed in an appropriate manner. Researchers have commonly analysed such self-reports as if they originated solely from the person providing the response, rather than as if they reflected something about both parties in the dyad. This error is referred to as pseudo-unilaterality (Duncan, Kanki, Mokros, & Fiske, 1984).

Attitudes and contact, the variables of interest in the contact hypothesis, do not occur in isolation, and yet researchers have tended to analyse them as if they tell us only about the respondent (Kenny et al., 2006). In order to examine these variables properly, data should, ideally, be collected from both members of the dyad. Of course, it would be impossible to ask, for example, 'the elderly' to take part in an experiment as one member of a dyad, with the original participant as the other member. But 'the

elderly' would still be providing a source of variance; they would still be affecting the responses of the participant answering questions about their attitudes towards the elderly. This highlights a difficulty for contact research which it is beyond the scope of this thesis to address, but those experiments which do specify a single person as the opposite member of a dyad, for example by asking participants about a specific outgroup friend, should ideally collect data from both members of the dyad so that data can be analysed appropriately and the underlying sources of variance properly examined. Where this thesis examines dyadic data (see chapters three and four), the appropriate dyadic analyses have been performed in an effort to address this difficulty.

Thesis aim

As so much research on the contact hypothesis relies on self-report measures of contact (particularly with regards to the quality of that contact, and to direct and indirect cross-group friendships) this highlights an important potential threat to the validity of contact research. As illustrated in Figure 1.2, a huge portion of the whole research endeavour on the contact hypothesis relies on the assumed validity of self-report methods of data collection, but if these are not valid, then the derived implications of the research are potentially flawed. This thesis attempts to address this important oversight, using a variety of methods to investigate whether the use of self-reports in future research on intergroup contact is appropriate.

Validating self-reports

Validity

There are four types of validity commonly examined in social research. *Internal validity* requires that two variables be related, and that this relationship be a causal one rather than an incidental one. Threats to internal validity are covered

exhaustively elsewhere (e.g., Reis & Judd, 2000) and the inclusion of a suitable control group in order to minimise these is a common practice. *External validity* refers to the generalisability of results to the population in general; this can be maximised by gathering data from a sample which is representative of the general population, and can be tested by replicating studies with different samples. *Construct validity*, as defined by Carmines and Zeller (1979), is the extent to which any measuring instrument measures what it is intended to measure, with minimal contributions from constructs of disinterest (such as the tendency to respond in a socially desirable manner) and random error. This can be tested by examining whether the measure is related to measures of other variables which are related to the construct being measured.

Nunnally (1978, p. 87) suggests that, in order to demonstrate *criterion-related validity*, the instrument must predict some criterion external to the instrument itself but which can be expected to arise from having a certain score on the test. This criterion can be available for measurement at the same time as the test is taken (*concurrent validity*) or at some later time (*predictive validity*). This would suggest that self-reports of contact should predict amount of real observed contact with the outgroup, both current and future. As it is difficult and time-consuming to accurately observe even one individual's real contact with the outgroup over an adequate period of time, this cannot be seen as a practical and efficient method of validating contact self-reports, and no studies have yet demonstrated criterion-related validity of self-reports of contact using this method. A method which could be employed is that of examining mutually-nominated cross-group friendships (e.g., Lee, Howes & Chamberlain, 2007), but this requires asking every individual in a closed group to nominate all of their friends. In such groups, for example among children in a

classroom, or workers in an office, the number of potential friendships is limited by the number of people, and any cross-group friendships existing outside the group will not be taken into account. Chapter Five examines large online communities which include almost every person in a large real community (in this case, Oxford University) and in which mutually-nominated friendships reflect real-life friendships, in an effort to overcome this problem and demonstrate criterion-related validity in self-reports of contact.

All these are examples of convergent validity, which requires that scores from two tests which purport to measure the same construct be similar to each other.

Discriminant validity is demonstrated when two scores that supposedly reflect unrelated constructs do *not* agree. Chapters Three and Four demonstrate discriminant validity.

Existing methods

Response biases (as discussed above) threaten *construct* validity, and methods exist which purport to diminish such threats. I will focus on Socially Desirable Responding (SDR). There are a number of methods of coping with SDR (for a review, see Nederhof, 1985). Perhaps the most popular method in contemporary social psychology is the development of implicit measures, which are ideally beyond the control of the respondent (see Greenwald, Banaji, Rudman, Farnham, Nosek & Mellott, 2002); however, no one has yet suggested an implicit measure of contact. More realistic preventative countermeasures include the inclusion of filler items to disguise the researcher's interest, providing assurances of confidentiality and anonymity (Becker & Bakal, 1970; Boruch & Cecil, 1979), and alleviating concerns by informing participants that there are no 'wrong' answers. All these strategies have

been found to be partially successful in reducing SDR, but they represent only a partial corrective (Bradburn, 1983).

Scales have been developed to measure Socially Desirable Responding so that it can be controlled for, thereby increasing the proportion of the measure which reflects the construct of interest. The degree to which participants are ‘faking’ can be tested (Arbisi & Ben-Porath, 1995), and many scales exist which purport to test the degree to which participants tend to respond in a socially desirable manner, usually by measuring the degree to which participants claim improbable virtues and deny common human frailties. Such scales include the Marlowe-Crowne scale (Crowne & Marlowe, 1960), Paulhus’ (1984, 1986) Balanced Inventory of Desirable Responding (BIDR), and, more specifically relevant to intergroup contact research, Dunton and Fazio’s (1997) Motivation to Control Prejudiced Responding scale (MCPR). An example of the use of such scales in contact research is provided by Vonofakou (2008), whose study controlled for scores on the Balanced Inventory of Desirable Responding while demonstrating that intergroup anxiety mediates the relationship between direct and indirect cross-group friendship and attitudes.

However, the utility of these validity scales has recently been questioned. Much research has found that controlling for scores on such scales either did not affect, or in fact *reduced*, the predictive validity of self-ratings (Ones, Viswesvaran & Reiss, 1995; Piedmont, McCrae, Riemann & Angleitner, 2000). Other studies found that groups of participants scoring ‘low’ or ‘high’ on SDR did not differ in terms of the predictive validity of their self-ratings (Barrick & Mount, 1996; Borkenau & Ostendorf, 1992; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990). In fact, Diener, Sandvik, Pavot, and Gallagher (1991) found evidence in support of the theory that “social desirability is a substantive personality characteristic which enhances well-

being, rather than being a response artefact and source of error variance” (p. 35). If a person’s score on an SDR scale reflects a stable facet of personality, partialling it out would be detrimental rather than beneficial, and this helps to explain why such scales have often been found to *reduce* predictive validity (Ones et al., 1995; Piedmont et al., 2000). I therefore argue that the validity of self-reports cannot be assured solely by the use of the scales currently existing for that purpose, and this underlines the need for alternative methods, in this case with specific reference to research on intergroup contact, with which to verify self-reports as a data collection technique.

Observer-reports

The method suggested by most researchers who have failed to find support for validity scales is to use observer-ratings (Piedmont et al., 2000), and these are now an accepted method for validation of self-reports in personality research and elsewhere. For example, using observer-reports, West, French, Kemp & Elander (1993) showed that self-reports were valid for assessing certain aspects of a person’s driving style, but invalid for other aspects. Self-reports and observer-reports are unlikely to share the same systematic biases. As a result, the correlation between the two should largely reflect true score variance rather than shared error variance, thus permitting validity assessments of the self-reports (Borkenau & Ostendorf, 1992; Piedmont et al., 2000). Observer-ratings are also less susceptible to SDR than self-ratings (Krahé, 2008), which increases their value, especially in a field which involves such sensitive issues as racial prejudice. Observer-ratings have also been used as an alternative to self-report; Lefkowitz, Eron, Walder, and Huesmann (1977), for example, found that aggression in 8 year old boys, as reported by their peers, was related to the level of violence occurring in their favourite TV programs. However, their potential to replace

self-reports is limited, as it is impossible to perform mediation analyses without self-report data on the variables of interest.

Cronbach (1971, p. 447) points out that one does not validate a test but rather an “interpretation of data arising from a specified procedure”. This implies that a test can be valid for measuring one phenomenon but invalid for another, and that the validity of a test must therefore be checked for each phenomenon it is used to measure. Even if the validity of self-reports in examining various aspects of personality had been conclusively demonstrated, this has not yet been attempted in the measurement of outgroup contact.

Choice of ‘observer’

The effectiveness of observer-ratings for verifying self-reports depends, however, on many factors, including the closeness of the relationship between observer and target, the amount of time for which they have been acquainted, and the amount of self-disclosure in the relationship. Norman and Goldberg (1966) found, for example, that higher familiarity between target and observer was associated with higher correlation between self- and observer-ratings, and that there was essentially zero correlation when the target was not known to the observer. On the other hand, other work has shown that there is evidence for significant agreement between one’s self judgment and how one is consensually judged by observers, even when observer judgements are based on only very short controlled samples of unacquainted targets’ behaviour (Borkenau & Liebler, 1995; Borkenau, Mauer, Riemann, Spinath & Angleitner, 2004; Funder, 1999).

Simply observing behaviour, therefore, even without any prior acquaintance, can be sufficient when inferring personality from the observation of behaviour, but concern exists that observer-ratings may only be effective for validating self-reports

on items concerning publicly observable behaviour, and will therefore be less useful for validating measures of attitudes or more invisible personality traits (Cronbach, Glaser, Nanda, & Raiaratnam, 1972; Sudman & Bradman, 1974). In the case of attributes such as racial attitudes and experiences of intergroup contact, there are probably very few cues that are publicly available to observers who are unacquainted with the person they are rating. Hence, in the case of these more covert attitudes it makes sense that observers, if they are to provide consensual judgments that might be used to validate self-reports, should be well acquainted with the target individual they are judging. McCrae (1993) found that the best possible observer was the target's spouse; correlations between self- and spouse-ratings were consistently higher than those between self- and peer-ratings. Spouse-ratings of personality also appear to be reliable and stable over long periods of time (Costa & McCrae, 1988). Chapter Three begins by comparing the self-reported contact and attitudes of each member of a married couple dyad, with their contact and attitudes as rated by their spouse or legal partner. As previously discussed, dyadic analyses are appropriate here. However, as a large proportion of social psychological research is conducted with university students as participants, and as most university students are unmarried, an alternative 'observer' is necessary for validation of students' self-reports. Chapter Three therefore continues by examining dyads composed of a student and one of their parents.

Once an individual has left home, whether to go to university or to start employment, a parent's ability to provide worthwhile observer-ratings of contact and attitudes will depend on how much their son or daughter is prepared to self-disclose to them. An observer who is in close proximity to the participant would be more preferable. As mentioned previously, not many students are married, and even in a

non-student population, not everyone has a spouse or legal partner; participants should not be excluded based on a meaningful factor such as their marital status. Even ignoring the meaningful exclusion of non-married participants, each target necessarily has only one spouse and personality may be perceived differently in different relationships (Casselden & Hampson, 1990). Although a spouse will most likely know the target better than anyone else, they may have their own response styles which contribute error variance, as might participants' parents, and it has been found beneficial to average the observer-ratings of more than one observer to reduce these effects. Kolar, Funder and Colvin (1996) found that simply increasing the number of observers increases the validity of ratings, as it controls for response biases in each individual observer which presumably will not be shared by all observers. Importantly, observers tend to agree when rating a target; McCrae and Costa's (1989) review of concurrent inter-judge agreement studies found consensus correlations of between 0.4 and 0.6. If multiple observers show consensus in their judgments, then that part of their judgments that is consensual becomes the criterion for validating self-reports (Kenny, 1994).

It would appear therefore that the optimal situation for validating self-ratings with observer-ratings would involve ratings by at least three peers with whom the target is well acquainted. Robins, Mendelsohn, Connell and Kwan (2004) and Emler, Tarry and St. James (2007) made use of a 'round-robin' design (Kenny, 1994) wherein participants took part in groups of four acquaintances, with each member rating not only themselves but each other member on each relevant trait or attribute. Neither of these studies used the round-robin design to evaluate measures relevant to the contact hypothesis, and Chapter Four makes use of this method to validate self-reports through the use of observer-reports.

Overview of the Thesis

Chapter Two demonstrates that self-reports of contact (but not self-reports of attitudes) are resistant to context effects. Two studies primed participants to consider positive or negative contact with an outgroup (control groups did not consider such contact), and found that attitudes were affected in the direction of valence of contact while contact self-reports remained unaffected.

Chapter Three describes three studies using observer-reports to validate self-reports of contact. Dyads composed of married couples, or of a student and one of their parents, reported on their own and each others' contact and attitudes. Self-reports of contact were found to be largely valid (at least, as assessed by other's reports), although parents' reports were found to agree more strongly with students' reports of contact at home, where parents can observe directly, than with students' reports of contact at university, where parents must rely on what students tell them.

As individual response styles will have contributed error variance in the studies reported in Chapter Three, the fourth chapter makes use of the round-robin method to examine agreement between self-reports and the observer-reports of three close friends, who cannot be expected to share a single response style.

To examine concurrent criterion-related validity in a direct fashion, Chapter Five describes the relationship between self-reports of contact and real observed contact in a large online community which reflects real-world friends rather than merely 'e-relationships'. A clear link is demonstrated between the number of real-world outgroup friends who interact on the 'Facebook' network, and participants' self-reported contact.

Chapter Six discusses the main findings of the thesis and their implications, and points out that these chapters are the first, and thus far the only, studies

undertaken to investigate the validity of self-reports of intergroup contact. Taken together, these chapters (and the nine original empirical studies they report) demonstrate the validity of self-reports of contact, and support their continued use in research which aims to reduce prejudice through optimal contact with outgroup members.

CHAPTER TWO:

Resistance of contact self-reports' to context effects

The current study aims to assess the validity of contact scales in terms of their resilience to context effects. As has been already discussed (see page 28 of Chapter One), attitude measurement is extremely context-dependent (Schwarz, 1999). Preceding questions can influence the answers to subsequent ones (Schwarz & Sudman, 1992), either by being assimilated into the target stimulus (Salancik & Conway, 1975) or by forming part of the standard of comparison (Strack, Schwarz & Gschneidinger, 1985). Recalling current life events, for example, generally influences responses in the direction of the valence of the event, while recalling past events has the reverse effect, whereby recalling a pleasant event in the past would cause participants to evaluate their current well-being comparatively negatively (Strack et al., 1985). Finding it difficult to recall an event also leads participants to evaluate their current state as including such an event infrequently, independent of whether this is in fact the case (Haddock, Rothman, Reber & Schwarz, 1999; Schwarz et al., 1991). The autobiographical memory literature shows that negative life events are poorly recalled in comparison with positive ones (Linton, 1982; Skowronski, Betz, Thompson, & Shannon, 1991; Wagenaar, 1986; Walker, Skowronski, & Thompson, 2003), due to the affect linked to negative events fading faster than the affect linked to positive events, as immediate steps are taken in the brain to reduce negative emotion right after it occurs. The resulting difficulty in recalling a negative event from the past might lead to an inflated sense that things were more positive in the past than they are now.

This empirical chapter is dedicated to discovering whether contact self-reports are resistant to such effects, due to their comparative objectivity. When completing a

self-report of amount of contact with an outgroup, it could be argued that a participant is drawing on quantitative fact, the amount of contact they have actually had.

However, answering such questions ‘cold’, having not considered the matter at all, usually, until presented with the question, may produce a different answer than answering after being asked to recall specific encounters. Recalling such experiences may increase the accessibility of memories of outgroup contact, and increase the amount of contact reported (Schwarz & Strack, 1999). Difficulty recalling such experiences is likely to have the opposite effect (Haddock et al., 1999; Schwarz et al., 1991), and only being able to recall such events from a long time ago is likely to lead participants to believe they have relatively little contact in comparison with past periods in their life (Strack et al., 1985).

Contact research most often presents questions concerning amount of contact experienced by a participant either before or after questions concerning attitudes towards the outgroup of interest. These measures are often similar in nature, and may therefore be at risk from common method variance (Podsakoff et al., 2003). Common method variance is that variance which results from the similarity between methods for measuring two constructs, rather than from a meaningful relationship between the constructs themselves (Fiske, 1982). Common method variance can exaggerate an association between two constructs measured with the same method (Wiggins, 1973). Context effects are related to common method variance in that both are elicited when subsequent responses automatically trigger recall of prior, similar, measures, with the tendency to, consciously or unconsciously, refer back to those measures. Podsakoff et al. (2003) suggest that the problem can be reduced by decreasing the perceived relevance of the previously-recalled information in short-term memory, through the employment of various techniques. Reducing the degree to which measurement of the

predictor variable is similar to, or connected with, the measurement of the criterion variable diminishes the respondents' tendency to use prior responses in answering subsequent questions, thus reducing consistency motives (tendencies to answer similar questions similarly as a result of the method rather than the construct being measured).

However, reducing common method variance does not eliminate the effects of context. Thoughts or memories triggered automatically by prior questions may affect subsequent responses even if these are presented in a different format. Many studies counterbalance the order of contact and attitude measures, and test for order effects, in order to control for the effects of thinking about one's attitude towards the outgroup before answering questions about contact with that outgroup, or *vice versa*. However, this precaution alone cannot account for the different effects that may be elicited depending on the participant's prior experiences and the varying ways in which people process information and reflect on past experience when answering self-report questions.

No research has yet focused on isolating context effects in self-reports of contact. This chapter uses both real recalled experiences and experimentally-derived experiences to demonstrate that, while self-reported attitudes are affected by such priming techniques, self-reported contact remains robust across all conditions.

Choice of outgroup

The census reported by the Office for National Statistics in April 2001 revealed South Asians to be the largest minority group in the United Kingdom, with half of all British minorities describing themselves as belonging to various subgroups of the South Asian ethnicity such as Indian, Pakistani, or Bangladeshi (Office for National Statistics, 2001). This means that the minority group most likely to be

encountered frequently by the majority group in the United Kingdom is the South Asian minority group. In addition, the Muslim faith has become a highly salient outgroup in recent years. The bombings in London in 2005 prompted a fivefold rise in faith hate crimes directed at Muslims (Dodd, 2005). In the fortnight following the bombings, 200 faith hate crimes predominantly directed at British Muslims were recorded by the Metropolitan police where in the same fortnight the previous year only 30 were reported. Arguably, the Muslim and South Asian minority groups cannot be separated, as many majority group members associate the one with the other. For these reasons, South Asians would appear to be the optimal outgroup for the purposes of this research, as they constitute a large (and therefore frequently encountered) and topically interesting minority group.

Study 1

Study 1 required participants to recall an event from their life in which they had experienced contact with the outgroup. Half were asked to recall a positive encounter and the other half were asked to recall an occasion of negative contact. This method was used by Wojciszke (1994, Study 2, 2005) in his studies of competence and morality. Half the participants were subsequently presented with self-report measures of contact, and half received the measures before they were asked to recall the event; thus half the participants were ‘primed’ and half were not. This method yielded a 2 (valence of recalled experience: positive vs. negative) x 2 (order: primed vs. unprimed) factorial design. Participants were also asked how long ago the event had taken place, because previous research suggests, as discussed above, that if the event took place some time ago, the information it provided might be used as a standard of comparison rather than to provide information on current levels of contact, and might therefore have an opposite effect to that intended (Strack et al., 1985).

Similarly, participants were asked how easy the event was to recall, in order to control for availability effects (Schwarz et al., 1991). Negative events were expected to be harder to recall, as this would replicate previous research on autobiographical memory (Linton, 1982; Skowronski et al., 1991; Wagenaar, 1986; Walker et al., 2003).

Hypotheses

1. When controlling for ease of recall and how long ago the event took place, attitudes towards the outgroup would be more positive in the positive-prime than in the negative-prime condition, compared with the unprimed condition. This would replicate a good deal of previous research (e.g., Salancik & Conway, 1975).

2. If self-reports of contact are, as argued, more objective responses, there would be no difference in self-reported contact in the positive and negative conditions, whether primed or unprimed, compared with the control condition; in other words, contact self-reports would prove to be resistant to context effects.

3. Negative events would be harder to recall, replicating previous research (Linton, 1982; Skowronski et al., 1991; Wagenaar, 1986; Walker et al., 2003).

Method

Participants

A total of 165 white students (94 female and 67 male, 4 undisclosed, mean age = 20.15, $SD = 1.94$) took part in the research.

Procedure

The questionnaire consisted of 'recall' items, measures of contact with and attitude towards Asians, and standard demographic questions. Following the demographic questions, half the participants received the recall questions followed by the attitude and contact measures (primed condition), and half received questions in

the reverse order (unprimed condition). Participants filled in the questionnaire in separate rooms or in separate parts of a large room.

Materials

Following demographic questions, participants were given the instruction: “You may now move on to the main part of the study. Please complete each page before turning to the next and do not return to pages you have completed previously once you have moved on. For the following questions, please note that ‘Asians’ refers to people living in the United Kingdom who are of Indian, Pakistani or Bangladeshi origin, rather than Chinese/Japanese/Korean etc.” It was important that participants move only forwards through the questionnaire, to make certain that primed participants were primed and non-primed participants were not.

Recall items

Half the participants were asked to recall a positive and half a negative encounter with a member of the outgroup. The following instructions were given: “We would like you to recall a situation in which you had a [positive/negative] encounter with an Asian person. This may or may not have been someone you knew, and the fact that the person was Asian need not have been relevant to the event. Please could you describe what happened to you? What exactly took place? What was said and done? What were the consequences? For example: ‘I was sitting on a bench when she came over and said...’” Participants were then given 20 lines on which to describe their experience. These descriptions were not analysed, being included simply to ensure participants spent some time considering the situation in question.

Subjective recency of the recalled event was measured by presenting two tick boxes beside the options *relatively recently?* and *some time ago?*, following the

question “Do you feel this event happened?”. Participants were also asked to estimate how many months ago the event happened.

Ease of recall was measured with two items: “How difficult or easy did you find it to recall this [positive/negative] experience?” and “How difficult or easy did you find it to write about this [positive/negative] experience?”, with answers given on seven-point scales anchored at ‘*difficult*’ (1) and ‘*easy*’ (7). These two items ($r = .41$, $N = 165$, $p < .001$) formed a composite measure in which higher scores indicated higher ease of recall.

Contact/Attitude items

Contact was measured with four items, two referring to contact in participants’ home towns, and two referring to contact while at university. Participants were asked ‘About how many of your friends [at home/at university] are Asian?’, with answers given on a seven-point scale labelled with ‘*None*’ (1), ‘*A few*’ (2), ‘*Less than half*’ (3), ‘*About half*’ (4), ‘*More than half*’ (5), ‘*Most*’ (6) and ‘*All*’ (7). Participants were also asked ‘How often do you spend time with friends [at home/at university] who are Asians?’, and answers were given on a seven-point scale labelled with ‘*Never*’ (1), ‘*Rarely*’ (2), ‘*Sometimes*’ (3), ‘*Quite a bit*’ (4), ‘*Often*’ (5), ‘*Very often*’ (6) and ‘*All the time*’ (7). The four items were averaged together to produce a reliable index of contact with Asians (Cronbach’s $\alpha = .80$) on which a high score indicated more contact.

Attitudes were measured using a ‘feeling thermometer’ and a set of five bi-polar adjectives on seven-point scales preceded by the instruction: ‘Based on your experience rate the extent to which you have each of the following feelings about Asians in general. (Please circle one number on each scale).’ Bi-polar adjectives were Negative-Positive, Friendly-Hostile, Suspicious-Trusting, Respect-Contempt and

Admiration-Disgust (Wright et al., 1997). One of Wright et al.'s items, Warm-Cold, was not included as it was rendered redundant by the feeling thermometer. After reversing items Hostile, Contempt and Disgust and transforming the feeling thermometer score from a 100 to a seven-point scale, all six items were averaged together to produce a reliable index of attitude to Asians (Cronbach's $\alpha = .82$) on which a high score indicated a more positive attitude.

The order of the Contact and Attitude measures was counterbalanced (there were no systematic effects of order).

Results and Discussion

The main results are displayed in Table 2.1. Before performing analyses to discover whether attitude judgements and reported contact would be affected by recalling positive or negative outgroup encounters, before or after the contact self-reports, it was necessary to conduct some preliminary analyses. The first of these was to ascertain how recency and ease of recall affected the data. Previous research suggests that events which happened a long time ago or are hard to recall would be used as a standard of comparison and have an opposite priming effect than that hypothesised (Schwarz et al., 1991; Strack et al., 1985). It was only necessary to examine the primed data, as the unprimed could not be affected in any way by either the recency of the event or the ease of recall. A one-way analysis of variance (ANOVA) was therefore computed on just the primed data, in order to discover any effects of the prime on 'ease of recall'. If, for example, negative events were deemed harder to recall than positive events, such as would be expected in accordance with autobiographical memory literature (Linton, 1982; Skowronski et al., 1991; Wagenaar, 1986; Walker et al., 2003), this would complicate the effect of the prime on attitude and contact judgements, and ease of recall would have to be controlled for

in the main analysis. However, contrary to the hypothesis, it was found that ease of recall was the same whether the prime was positive or negative; $F(1, 74) = .07$, ns. Negative events were no harder to recall than were positive ones. One possible explanation for this deviation from the prediction is based on the work of Sedikides and colleagues, who found that negative *self*-referent events are poorly recalled while negative *other*-referent events are not (e.g., Sedikides & Green, 2000). It is probable that an event in which an Asian person behaved unpleasantly towards a participant would be categorised as other- rather than self-referent, and would not therefore trigger the self-protecting instinct that leads to shallow processing and poor recall (Sedikides & Green, 2000).

Table 2.1

Mean attitude and contact as a function of presence and valence of prime (Study 1)

	Primed (Recall then judgement)		Unprimed (Judgement then recall)	
Valence of prime	Negative (<i>N</i> = 31)	Positive (<i>N</i> = 42)	Negative (<i>N</i> = 41)	Positive (<i>N</i> = 38)
<i>Mean</i> attitude	3.53 (1.35)	4.68 (1.25)	3.91 (1.37)	3.97 (1.49)
<i>Mean</i> contact	2.72 (1.25)	1.90 (1.19)	1.93 (0.97)	1.86 (1.21)

Note. Standard deviations in parentheses.

Second, a comparison was made between how long ago the recalled positive and negative events occurred. Responses to the item asking participants to estimate how many months ago the event took place were extremely skewed and could not be normalized by transformations. Frequencies were therefore compared for the dichotomous variable, relatively recently/some time ago. Positive events were far more likely to be reported as occurring recently, and negative events as not recently,

than would be expected by chance, $\chi^2 = 16.18$ $p < .001$. The dichotomous variable ‘recent or not’ was therefore used as a covariate in the main analysis, below.⁴

The primed data were analysed separately from the unprimed, as there was no way for the prime to affect self-reported contact and attitudes when it was presented after the contact and attitudes measures. An analysis of covariance (ANCOVA) was computed on the primed data, with valence of prime as the independent variable, ‘recent or not’ as a covariate, and attitudes as the dependent measure. This ANCOVA revealed that attitudes were more positive after a positive prime than after a negative prime, $F(1, 71) = 5.88$, $p = .02$. As hypothesised, recalling an outgroup encounter immediately prior to reporting attitudes affected reported attitudes in a manner consistent with the valence of the encounter (see Table 2.1). There was no effect of the covariate, recency, on attitudes. As expected, no effects were found of valence of prime on the unprimed data.

The same ANCOVA with contact as the dependent variable revealed no effects of valence of prime on the primed data, $F(1, 70) = 1.91$, ns, indicating that contact self-reports are resistant to priming effects. Recalling an outgroup encounter immediately prior to reporting level of contact did not affect those reports. However, the covariate, recent or not, did have an effect on contact, $F(1, 70) = 7.49$, $p < .01$. Participants who reported subjectively recent events, whether positive or negative, reported significantly lower outgroup contact ($M_{\text{recent}} = 1.72$; $M_{\text{not recent}} = 2.76$). Again, type of prime did not have any effect on contact in the unprimed data.

⁴ The median scores for the skewed variable of estimated months were consistent with the dichotomous variable. Positive events occurred fewer months ago (median = 6 mths) than did negative events (median = 36 mths).

Study 2

The second study improved on the first in a number of ways. Study 1 gave participants the freedom to pick any event from their lives to describe, which, although eliciting realistic events, did not of course allow control over the type of event chosen to serve as the prime, or over its historical recency. Indeed, effects of recency were found (although not of ease of recall). Because participants were able to choose an event from any time in their life, it was only possible to control statistically for these effects, *post hoc*. An improvement would be to prevent them occurring.

In order to exercise control over these aspects, Study 2 dispensed with the method of asking participants to recall an event from their own lives. Instead, participants read about an encounter between a member of their own ingroup and a member of the outgroup. As such it was possible to exert complete control over the priming event, its valence, and its time of occurrence, hence to ensure its suitability for the purposes of this study.

The target outgroup was also changed in the second study. The first study used Asians as an outgroup. It is possible participants controlled their self-reports of a prior experience with an ethnic outgroup, for reasons of social desirability or political correctness. As it is more socially acceptable to express negative attitudes towards a rival university than to an ethnic outgroup, the second study used students from another university in the same town as the participants' university.

For Study 2 a one-factor design was used, with three levels of the independent variable (positive outgroup experience vs. negative outgroup experience vs. control). Finally, to the measures of contact (proportion of outgroup friends and the amount of time spent with them) were added measures of quantity of contact, and also indirect

cross-group friendship, to yield a more sophisticated and comprehensive measure of participants' prior contact.

Hypotheses

1. As in Study 1, attitudes towards the outgroup would be more positive in the positive-prime than in the negative-prime condition, compared with the unprimed condition.

2. There would be no difference in self-reported contact (quantity, or direct or indirect cross-group friendship) in the positive and negative conditions, whether primed or unprimed, compared with the control condition. Replicating Study 1, contact self-reports would prove to be resistant to context effects.

Method

Participants

A total of 118 white students (63 female and 55 male, *mean age* = 20.31) took part in the research.

Procedure

Participants read a passage which was ostensibly taken from a student newspaper, which described one of three events depending on the condition. Figure 2.1 shows the three passages as they appeared in the experiment. The 'positive prime' condition involved a passage describing an event wherein an ingroup member (student at the participant's university) was attacked and injured by an unknown assailant, and was taken to hospital by an outgroup member (student of another university in the same town). The 'negative prime' condition involved a passage describing an event in which the same ingroup member was attacked and injured by an outgroup member, and taken to hospital by an unidentified individual. The passage read in the control

Figure 2.1

Passages used in Study 2

Condition: Positive indirect intergroup encounter

Please read this short newspaper extract.

Last Thursday night, May 22nd, at Filth nightclub the police were called after an Oxford University student was physically assaulted by an unknown attacker. He was knocked unconscious, after a violent physical assault. His collarbone was broken, and he sustained heavy bruising and lacerations to his face. An ambulance was called by an Oxford Brookes student, who accompanied him to the John Radcliffe Hospital, and stayed with him until he regained consciousness. The police have yet to identify the attacker.

Condition: Negative indirect intergroup encounter

Please read this short newspaper extract.

Last Thursday night, May 22nd, at Filth nightclub the police were called after an Oxford University student was physically assaulted by an Oxford Brookes student. He was knocked unconscious, after a violent physical assault. His collarbone was broken, and he sustained heavy bruising and lacerations to his face. An ambulance was called by an anonymous witness, who accompanied him to the John Radcliffe Hospital, and stayed with him until he regained consciousness.

Condition: No indirect intergroup encounter

Please read this short newspaper extract.

Last Thursday night, May 22nd, at Filth nightclub the police were called after a man was physically assaulted by an unknown attacker. He was knocked unconscious, after a violent physical assault. His collarbone was broken, and he sustained heavy bruising and lacerations to his face. An ambulance was called by an anonymous witness, who accompanied him to the John Radcliffe Hospital, and stayed with him until he regained consciousness. The police have yet to identify the attacker.

condition simply involved an unidentified victim being attacked by an unknown assailant and taken to hospital by ‘a person’ who was not described further.

Participants then filled in a questionnaire, in separate rooms or in separate parts of a large room. The questionnaire included questions which ostensibly examined comprehension of the passage, but which actually functioned as a manipulation check.

Materials

The questionnaire consisted of ‘manipulation check’ items, measures of quantity of contact, direct and indirect cross-group friendship, attitude towards Asians, and standard demographic questions.

The manipulation check items comprised three questions, asking for a description of the three people who had filled the roles of assailant, victim, and rescuer. All participants correctly identified the ingroup and outgroup when group membership was specified in the ‘newspaper article’.

The order of the contact and attitude measures was counterbalanced (there were no reliable effects of order). Direct cross-group friendship measures were ‘About how many of your friends are students from [the other university]?’ and ‘How often do you spend time with friends who are students from [the other university]?’; answers were given on a seven-point scale labelled with ‘None’ (1), ‘A few’ (2), ‘Less than half’ (3), ‘About half’ (4), ‘More than half’ (5), ‘Most’ (6) and ‘All’ (7)). These two items ($r = .74$, $N = 118$, $p < .001$) formed a composite measure in which higher scores indicated a higher degree of direct cross-group friendship.

Indirect cross-group friendship was measured with the question ‘About how many of your [ingroup university] friends do you think have friends who are students

from [the other university]?’ and answers were given on the same seven-point scale as for direct cross-group friendship measures. Multiple-item scales tend to be more reliable than single-item measures, because random measurement error is cancelled out when aggregating across multiple items (Robins, Hendin, & Trzesniewski, 2001). However, measurement error is not always random, and systematic errors such as socially desirable responding are compounded by scales involving redundancy (Robins et al., 2001). Redundancy can also lead to fatigue, which can cause participants to respond randomly or skip questions. One benefit of single-item measures is that they eliminate item redundancy⁵ (Burisch, 1984b). Multiple-item scales are necessary to ensure content validity for multifaceted constructs (Robins et al., 2001) but as indirect cross-group friendship is conceptually unidimensional, I suggest that a single item measure may still possess adequate content validity.

Quantity of contact was measured with two questions: ‘How often do you meet students from [the other university] in your everyday life?’ and ‘How often do you spend time with people who are students [at the other university]?’. Answers were given on a seven-point scale labelled with ‘*Never*’ (1), ‘*Rarely*’ (2), ‘*Sometimes*’ (3), ‘*Quite a bit*’ (4), ‘*Often*’ (5), ‘*Very often*’ (6) and ‘*All the time*’ (7). These two items were highly correlated ($r = .89$, $N = 118$, $p < .001$) and averaged to form a measure in which higher scores indicated higher quantity of contact.

Attitudes were measured using a ‘feeling thermometer’ and the same five bipolar adjectives, rated on seven-point scales, used in Study 1 (after Wright et al., 1997). After again reversing three items (‘*Hostile*’, ‘*Contempt*’ and ‘*Disgust*’) and

⁵ Single-item measures have been used appropriately and to great effect in a wide range of research areas (e.g., Aron, Aron, & Danny, 1992; Campbell, Converse, & Rodgers, 1976; Diener, 1984; Hazan & Shaver, 1987; Russell, Weiss, & Mendelsohn, 1989; Sandvik, Diener, & Seidlitz, 1993)

transforming the feeling thermometer score from a 100 to a seven-point scale, all six items were averaged together to produce a reliable index of attitude to the student outgroup (Cronbach's $\alpha = .88$) on which a high score indicated a more positive attitude.

Results and Discussion

Main results are shown in Table 2.2. A one-way ANOVA with 'condition' as the single, three-level factor, and attitudes towards the outgroup as the dependent variable, revealed a significant main effect of priming on attitudes, $F(2,106) = 13.03$, $p < .001$. Post-hoc Scheffé's tests revealed that attitudes in the positive condition ($M = 5.15$) were significantly more positive than attitudes in the control condition ($M = 4.69$, $p = .046$). Attitude in the negative condition ($M = 4.22$) were also significantly less positive than attitudes in the control condition ($p = .041$). Finally, attitudes in the negative condition were also significantly less positive than attitudes in the positive condition ($p < .001$).

A second one-way ANOVA on contact as the dependent variable revealed no effect of 'condition' on quantity of contact, $F(2,106) = .84$, ns, direct cross-group friendship, $F(2,106) = .65$, ns, or indirect cross-group friendship, $F(2,106) = 1.84$, ns.

Table 2.2

Mean attitude and contact as a function of presence and valence of prime (Study 2)

	Negative (<i>N</i> = 40)	Positive (<i>N</i> = 39)	Control (<i>N</i> = 39)
<i>Mean</i> attitude	4.39 (.76)	5.07 (.17)	4.73 (.15)
<i>Mean</i> contact (Composite measure)	2.37 (.15)	2.31 (.14)	2.62 (.20)
<i>Mean</i> quantity of contact	2.44 (1.18)	2.41 (.90)	2.72 (1.30)
<i>Mean</i> direct cross-group friendship	2.04 (.85)	2.19 (.95)	2.33 (1.38)
<i>Mean</i> indirect cross-group friendship	2.26 (1.10)	2.47 (1.16)	2.94 (1.56)

Note. Standard deviations in parentheses.

Thus those participants who read an account of a positive intergroup interaction subsequently reported more positive attitudes, and those who read of a negative intergroup interaction reported more negative attitudes. However, their self-reports of contact were not influenced in the same manner; participants did not exaggerate their intergroup contact as a result of the positive experience, nor did they underreport their contact as a result of reading about a negative interaction. Any temporary shift in attitudes experienced by participants as a result of the prime did not affect their reported contact. This speaks to the robustness of self-reports as a measure of contact, if they fail to be influenced by a prime which affects attitudes in a predictable manner.

General Discussion

Two studies supported the hypothesis that self-reports of contact, being of a more objective nature, are resistant to context effects, while self-reported attitudes, being a good deal more subjective, are not. I discuss these findings in terms of the interrelations between subjective recency of events and degree of intergroup contact, priming effects on attitudes, and implications for the conclusions drawn from research on the contact hypothesis.

Contact was unaffected in either study by the valence of the intergroup encounter considered. However, in the first study, one variable, recency, was found to have a relationship with contact. Participants who reported the recalled encounter to be subjectively recent tended to report lower contact than participants who reported the recalled encounter to have occurred a long time ago. Prior research (Strack et al., 1985) would suggest that the opposite effect might be found, that those participants who cannot recall a recent event would consider themselves to have low contact. An important point, however, is that recency was a dependent, not an independent

variable; participants were not required to try to think of a recent event. The instructions simply stated that participants should try to recall ‘an event’, and this phrasing, arguably, accounts for this counter-intuitive finding. Participants were free to recall any event, from any time in their past, and those who recalled events from a long time ago did not necessarily ‘fail’ to find a recent event. I suggest that participants who have a comparatively large number of Asian friends, and therefore a comparatively large number of events from which to choose, would not necessarily choose the most recent event in their memory (‘temporarily accessible information’, Schwarz & Bless, 1992a). Participants with close Asian friends, with whom they share a history of encounters, may discuss or ‘access’ such encounters on a regular basis, leading to their constituting ‘chronically accessible information’ (Schwarz & Bless, 1992a), and increasing the likelihood of such non-recent events being chosen for consideration in the first study. Indeed, I suggest that participants with relatively few Asian friends would have to draw on superficial events, which would be unlikely to be recalled a long time later (Craig & Lockhart, 1972). Low-contact participants would therefore be more likely to recall a recent, superficial event, while high-contact participants could choose from a greater number of non-superficial events, recent or otherwise. Thus in addition to the analysed effect of recency on self-reported contact, I suggest that quality and quantity of intergroup contact can also have an effect on the subjective recency of recalled intergroup encounters.

As concerns attitude self-reports, both studies found that when participants were primed to consider a positive or negative intergroup encounter, their attitudes were in accordance with the valence of the encounter, in comparison to controls. The first study revealed that this effect was not reversed when the recalled event was seen as hard to recall, or as occurring a long time ago, as might be expected from prior

research suggesting that such events might be used as a standard of comparison rather than as part of the target stimulus (Schwarz et al., 1991; Strack et al., 1985). In particular, one might have expected to discover that when participants found it hard to think of a negative event, they might have experienced an increase in their positive attitudes, and, *vice versa*, if they found it hard to think of a positive experience they might have felt more negatively towards the outgroup. This finding is explained, however, by an additional analysis on both primed and unprimed data with ease of recall as the dependent variable, which revealed a just-significant effect of valence of to-be-recalled event, showing that, in accordance with the hypothesis, positive events *did* tend to be slightly easier to recall than negative events.⁶ Therefore ease of recall may have amplified the effects of valence of prime on attitudes. Although marginal, this finding suggests a potential prejudice intervention paradigm, which could be explored by future research (see Chapter Six).

⁶ $M_{\text{positive}} = 4.06$, $SD = 1.03$, $M_{\text{negative}} = 3.74$, $SD = 1.00$, $F(1,156) = 3.92$, $p = .05$

CHAPTER THREE:

Validation of contact self-reports using the observer-reports of family members

Having found that self-reports of contact are resistant to context effects, it remains to be discovered whether self-reports of contact actually measure what they purport to measure. As described earlier, (see Chapter One), the method of validation suggested by most researchers is to use observer-ratings (Piedmont et al., 2000), because self-reports and observer-reports are unlikely to share the same systematic biases. The correlation between the two should therefore reflect true score variance rather than shared error variance (Borkenau & Ostendorf, 1992; Piedmont et al., 2000). The decreased susceptibility of observer-ratings to socially-desirable responding (SDR) is also an advantage (Krahé, 2008). Observer-ratings are arguably the best method developed to date for validating self-reports (Howard, 1994). Study 3 examines the validity of self-reports of outgroup contact by comparing them with the reports of the observer who is, arguably, best equipped to judge - the spouse (McCrae, 1993).

Study 3

Higher familiarity between observer and target is associated with higher correlation between self- and observer-ratings (Norman & Goldberg, 1966), and observers who know the target well are particularly necessary when validating measures of attitudes or more invisible personality traits (Cronbach et al., 1972; Sudman & Bradman, 1974). While it is arguable that outgroup contact is not necessarily invisible, in the sense that an intergroup encounter can of course be observed by others, it is also the case that many encounters are *not* observed, and must be communicated to another person in order to affect an observer-report. As race is a

socially sensitive issue, it is arguable that a participant in an intergroup interaction is unlikely to communicate the racial group membership of the other people in the interaction to anyone other than those he or she is closest to. Correlations between self-ratings and ratings made by a participant's spouse have been found to be consistently higher than correlations between self-ratings and ratings made by peers (McCrae, 1993). Spouse-ratings of personality also appear to be reliable and stable over long periods of time (Costa & McCrae, 1988).

Factors affecting observer-reports

It must be noted that observer-reports, even those of a spouse, do not provide a perfect measure of the target's level of a given variable. While less prone to SDR (Krahé, 2008), observers can still rely on response strategies such as extreme responding or moderacy bias, norm-based heuristics, or self-projection (Cronbach, 1955). Self-projection refers to the tendency to exaggerate similarities between the self and others, and norm-based responding refers to the tendency to base responses on the perception of the average person rather than on the target themselves. There is cause for optimism in this case, in that these strategies are more likely to guide responses when the respondent is unsure of how to respond (Konstabel, 2006), or when the target's beliefs and behaviours are ambiguous (Lenton, Bryan, Hastie & Fischer, 2007). Self-projection, for example, is often used to compensate for a lack of knowledge about the person being rated (Hodges, Johnsen & Scott, 2002), which is more likely to occur when observers are rating a less familiar target, rather than their spouse. However, these factors must still be taken into consideration when comparing observer-reports with self-reports, as they are likely to inflate self-observer agreement (Cronbach, 1955). Self-projection will inflate self-observer agreement to the extent that the target is similar to the judge, and norm-based heuristics will inflate self-

observer agreement to the extent that the target is similar to the norm. Self-observer agreement resulting from the similarity of the target to the norm is referred to as *stereotype accuracy* (Cronbach, 1955). The accepted method for controlling for stereotype accuracy and self-projection when computing self-observer agreement is to perform a partial correlation (Vogt & Colvin, 1999). However, this is only appropriate when data are independent.

Non-independence

The problem with using observers who are familiar with the target, is that the closer the relationship between the two, the greater the degree of non-independence. As Kenny et al. (2006) point out, social science research has relied mostly on standard statistical methods which were developed for the analysis of individuals, and which therefore assume that data are independent. The data reported in this chapter are not independent, in that even when controlling for any variation caused by the independent variable, the data from one member of each married couple are likely to be strongly affected by the other member of the couple. As outlined by Kenny et al. (2006), non-independence between two persons can result from voluntary linkage, such as between friends or lovers, kinship linkage, such as between siblings or between a parent and child, yoked linkage, such as between two people who have shared environmental stimuli, for example having gone to the same university, or from any combination of these. Most voluntary or kinship linkage also involves yoked linkage, for example, and of course married couples involve all three types at once.

The sources of non-independence are listed by Kenny et al. (2006) as *compositional*, *partner influence*, *mutual influence*, and *common fate*. *Compositional* non-independence refers to similarities pre-dating the dyad's creation. Similarities in political beliefs, for example, might contribute to the formation of a friendship or even

a marriage (assortative mating, Epstein & Guttman, 1984). A married couple are likely to show similarities in education, age, socioeconomic status and religion, even though all these things might have pre-dated the marriage. A more obvious cause of non-independence in dyad members is *partner influence*, which describes the effect one member has on the other. A husband might be introduced to skiing by his wife, and subsequently go on to see this pursuit as just an important part of his identity as his wife sees it as part of hers. Relatedly, *mutual influence* refers to the influence both members might have on each other. Reciprocal liking is an example of non-independence caused by influence from both members of the dyad. Finally, both members of a dyad are likely to be exposed to similar environmental stimuli following the formation of the dyad, leading to *common fate* non-independence. A married couple usually live in the same house, and take care of the same children – similarities resulting from the dyad's formation and which lead to an irretrievable loss of dependence in data gathered from the dyad.

To illustrate this with an example relevant to contact research, Sophie and Ian met while working as physiotherapists at the same practice, and both tend to meet a relatively high number of disabled people through this work (compositional non-independence). Ian introduces Sophie to his friend Erica who has recently suffered a back injury leaving her partly paralysed (partner influence). Following discussion, Sophie and Ian decide to give a class together to teach disabled people and their carers how to practise massage therapy at home (mutual influence and common fate). Sophie's self-reported contact with disabled people cannot be considered separately from Ian's contact, as there is so much non-independence existing between the two. If both participate in a study relating contact to attitudes towards disabled people, the relationship between contact and attitudes would be inflated if Ian's data were entered

as if sourced from one individual and Sophie's from another, as much of the same contact is contributing to both Sophie's and Ian's self-reports.

Structural Equation Modelling (SEM), as recommended by Kenny and Acitelli (2001) for dyadic analysis, determines the degree of variance uniquely attributable to self-observer agreement, while controlling for self-projection, or assumed similarity. However, and most relevantly in this case, when Ian provides an observer-report of Sophie's contact, he will necessarily be considering his own contact as well as hers, because the two are intertwined. It would be very difficult therefore to separate the effect of non-independence from the effect of self-projection. Ian's observer-reports are likely to agree with Sophie's self-reports even if they are based purely on self-projection. Controlling for the effect of Ian's self-reported contact on his ratings of Sophie's contact is likely to prove counter-productive, as Sophie and Ian share a lot of the same contact. Controlling for self-projection and stereotype accuracy is only beneficial to the extent that the target is unlike the stereotype, and target and judge are dissimilar (Cline, 1964).

Funder (2001) suggests that correcting for stereotype accuracy can also do more harm than good, for example if person A predicts person B's responses perfectly, correcting for stereotype accuracy will make person A's predictions less, rather than more accurate. However, Kenny et al. (2006) counter this argument by pointing out that while person A's accuracy is lowered, the accuracy of many other scores will be raised. It remains difficult, however, to ensure that only artifactual agreement is being partialled out, and no substantive agreement. However, a method exists to control for stereotype accuracy without partialling out meaningful agreement resulting from similarities between spouses. This method is *pseudo-couple analysis* (Corsini, 1956) and will be discussed below.

Dyadic analyses

It can be seen that unilateral analyses such as analysis of variance and multiple regression are inappropriate here, and that data must be examined using a technique which can account for non-independence. The analysis suggested by Kenny et al. (2006) for a reciprocal design, in which data from both members of a dyad are measured, is to create a dyadic index. This is a measure of correspondence between the two members of the dyad on the variables of interest, and has been used by many of the most important studies of couples (e.g., Levinger & Breedlove, 1966; Murstein & Beck, 1972). Acitelli, Kenny and Weiner (2001) used dyadic indexes to analyse the predictions of husbands and wives concerning their partners' responses to a marital values questionnaire. As Study 3 involves a similar design, this analysis was deemed to be appropriate here. Computing a dyadic index implies an *idiographic* rather than a *nomothetic* analysis, as a similarity or dissimilarity measure is computed for each dyad separately rather than across all dyads.

Kenny et al. (2006) suggest that when the *level* of a variable is of interest rather than the shape or spread of the scale items making up that variable, the optimal method is to average the item scores and use a discrepancy analysis. As the variable of interest in this case is the *level* of contact experienced with an outgroup, and any differences in the spread of the items are likely to be due to an extreme responding or moderacy bias, a discrepancy measure was computed for each dyad, on each variable.

Pseudo-couple analysis

Once real discrepancy scores have been computed between married couples, each male should be paired with a female at random who is not his wife, and further discrepancy scores should be computed. The same should be done for all females, and an average taken between male and female pseudo-couple discrepancy scores. This

average should reflect the degree of discrepancy that would be expected by chance, thereby providing a baseline for comparison of the real discrepancy scores. If no difference is found to exist between the real and pseudo discrepancy scores, this would suggest that any agreement between observer- and self-reports is artifactual, in that it reflects knowledge of people in general rather than of the target in question, and is not sufficient to validate self-reports as a measure of contact. A test for a significant difference between the real and pseudo discrepancy scores should provide a fairly conservative evaluation of whether observer-reports reflect, to a significant degree, the observers' judgements purely of the targets rather than of people in general. Any significant self-observer agreement should speak to the validity of self-reports as a measure of contact.

Study 3 compared the self-reported contact and attitudes of each member of a married couple dyad with their contact and attitudes as rated by their spouse or legal partner, and taking into account the level of agreement which would be expected by chance. Extroversion was included as an attribute to be rated as it is one of the more visible traits of personality (Borkenau & Liebler, 1995) which is well-predicted by observer-ratings (Borkenau & Ostendorf, 1992; Costa & McCrae, 1988; Emler et al., 2007; Ones et al., 1995; Piedmont et al., 2000). Thus it was used as a baseline, against which to compare agreement of self-other measures of intergroup contact.

Hypotheses

1. Self-reports of extroversion would agree strongly with observer-reports, replicating previous research.
2. Self-reports of direct cross-group friendship and quantity of contact would be validated by observer-reports.
3. Attitudes, being more susceptible to SDR, would not be validated.

4. Indirect cross-group friendship would prove very difficult for an observer to rate, as the observer would have to know about the contact of their spouse's friends rather than the contact of their spouse in order to make an informed estimate.

5. Amount of self-disclosure would moderate self-observer agreement, in that couples who do not self-disclose very much would make poor observers for the purposes of validating self-reports.

Method

Participants

Sixty-two heterosexual white British couples took part voluntarily in the study. They comprised 62 females aged between 23 and 83 (mean age = 49) and 62 males aged between 24 and 81 (mean age = 52). Couples had been together between 1 and 55 years, with a mean relationship length of 25 years.

Procedure

Each member of a couple filled in the questionnaire at the same time, in separate rooms, without conferring. Participants answered questions about themselves and then about their spouse, in a reciprocal dyadic design (Kenny, 1988). The ostensible purpose of the questionnaire was to discover whether spouses knew more about each other on socially acceptable topics than on topics where social norms prevent easy conversation. The questionnaire appeared to start at topics that would be easy to discuss and increasingly probe socially less acceptable topics.

Materials

The questionnaire consisted of measures of contact with Asians (quantity of contact and direct and indirect cross-group friendship), attitude towards Asians and extroversion.

Following demographic questions, the questionnaire began with a section on preferences for various leisure activities; this was to fit the ostensible purpose of the questionnaire; data from these questions were not analysed.

Extroversion was then measured using the four-item extroversion measure (seven-point bi-polar scales anchored at '*Restrained- Candid*', '*Active- Passive*', '*Silent- Talkative*', '*Aloof- Gregarious*') from the NEO-Five Factor Inventory (Costa & McCrae, 1988) as well as an additional single seven-point scale anchored at '*Extroverted*' and '*Introverted*'. After reversing the '*Extroverted - Introverted*' and '*Active - Passive*' measures, the composite scale, on which higher scores reflected higher levels of extroversion, yielded a reliable index of female self-reports ($\alpha = .84$), female observer-reports ($\alpha = .77$), male self-reports ($\alpha = .80$), and male observer-reports ($\alpha = .84$).

The order of the contact and attitude measures was counterbalanced, and these were preceded by the note used in Study 1, specifying the ethnic group the questions referred to: 'For the following questions, please note that 'Asians' refers to people living in the United Kingdom who are of Indian, Pakistani or Bangladeshi origin, rather than Chinese/Japanese/Korean etc'.

For the 'contact' measures, participants were provided with the instruction 'These questions refer to contact you may have with Asians in your home town.'

Quantity of contact was measured by two items: 'How often do you meet Asians in your everyday life at home?' and 'How often do you *talk with* people who are Asians?'. Answers were given on seven-point scales labelled with '*Never*' (1), '*Rarely*' (2), '*Sometimes*' (3), '*Quite a bit*' (4), '*Often*' (5), '*Very often*' (6) and '*All the time*' (7). An adequate inter-item correlation was obtained for female self-reports, $r = .82$, $N = 62$, $p < .001$, female observer-reports, $r = .78$, $N = 62$, $p < .001$, male

self-reports, $r = .62$, $N = 62$, $p < .001$, and male observer-reports $r = .79$, $N = 62$, $p < .001$.

Direct cross-group friendship was measured with two items: 'About how many of your friends at home are Asian?' (with answers given on a seven-point scale labelled with 'None' (1), 'A few' (2), 'Less than half' (3), 'About half' (4), 'More than half' (5), 'Most' (6) and 'All' (7), and 'How often do you spend time with friends at home who are Asians?' (answers on the same seven-point scale as that presented with the 'Quantity' items). An adequate inter-item correlation was obtained for female self-reports, $r = .67$, $N = 62$, $p < .001$, female observer-reports, $r = .64$, $N = 62$, $p < .001$, male self-reports, $r = .62$, $N = 62$, $p < .001$, and male observer-reports $r = .76$, $N = 62$, $p < .001$.

Indirect cross-group friendship was measured with a single item: 'About how many of your *white friends* at home have friends who are Asian?', with answers given on a seven-point scale labelled with 'None' (1), 'A few' (2), 'Less than half' (3), 'About half' (4), 'More than half' (5), 'Most' (6) and 'All' (7).

Outgroup attitudes were measured using a 'feeling thermometer' and a set of five bi-polar adjectives on seven-point scales preceded by the instruction: 'Based on your experience rate the extent to which you have each of the following feelings about Asians in general. (Please circle one number on each scale).' Bi-polar adjectives were 'Negative - Positive', 'Friendly - Hostile', 'Suspicious - Trusting', 'Respect - Contempt' and 'Admiration - Disgust'. These five adjectives were taken from Wright et al.'s (1997) six-item scale; the sixth item, 'Warm - Cold', was not included, as it was made redundant by the feeling thermometer. After transforming the feeling thermometer scores to put them on the same scale as the other attitude scores (by dividing all scores by 100 and multiplying by seven)

and reversing '*Friendly - Hostile*', '*Respect - Contempt*' and '*Admiration - Disgust*' items, it was found that there were strongly inconsistent correlations with contact across the 'Attitude to Asians' items, and only the three consistent items ('*Friendly*', '*Respect*', and the transformed feeling thermometer) were aggregated into a single 'Attitude' measure, on which a high score indicated a more positive attitude. Female self-reports on this scale yielded a Cronbach's alpha of .86, female observer-reports .92, male self-reports .88, and male observer-reports .88.

After these 'self-ratings', participants were presented with the following paragraph: 'Now we would like to know about the person who is participating with you (your spouse). The rest of this booklet contains questions similar to those you just answered, but they will relate to your spouse; we will refer to this person as 'X'. We understand that many of the following questions may be difficult, but do try to answer every question to the best of your ability!'

Participants were asked to provide demographic information about their spouse, and to state how long they had been together. Participants also answered questions concerning the closeness of their relationship with their spouse - a seven-point scale anchored at '*not at all close*' (1) and '*very close*' (7) - and the degree of self-disclosure between the two. Intended self-disclosure between the couple was measured with four items (from Turner, 2005): 'Imagine you had a problem that was worrying you. How likely is it that you would tell X?', 'Imagine you had an important secret. How likely is it that you would choose X to tell?', 'If X was worried about something, how likely is it that they would tell you about the problem?' and 'If X had an exciting secret, how likely is it that they would tell you about the secret?'. These items were accompanied with seven-point scales anchored at '*Not at all likely*' (1) and '*Very likely*' (7). Actual self-disclosure was measured with four items (also from

Turner, 2005): ‘How often do you talk about how you are feeling to X?’, ‘How often do you discuss intimate or personal issues with X?’, ‘How often does X talk to you about how they are feeling?’ and ‘How often does X discuss intimate or personal issues with you?’. Seven-point scales were presented anchored at ‘*Never*’ (1) and ‘*All the time*’ (7), with ‘*Sometimes*’ at the midpoint. A principal components analysis performed on the 16 self-disclosure items (i.e., both self- and observer-reported self-disclosure of both males and females), using Quartimax rotation, revealed only one factor. It was therefore deemed appropriate to measure the self-disclosure of each couple rather than of each individual, or in other words to include self-disclosure as a *between-dyads* rather than an *within-dyads* variable (Kenny et al., 2006). For this reason, the 16 self-disclosure items were combined into one scale where higher scores indicated more self-disclosure between the couple ($\alpha = .95$).

Participants were then given all the ‘self-rating’ questions again, but this time the questions were reworded so that they asked the participant to rate their spouse.

Results and Discussion

Means and standard deviations are reported in Table 3.1. The next step was to calculate three discrepancy scores for each dyad on each variable, followed by three pseudo-couple discrepancy scores. Thus, for each of the 62 dyads participating, the following scores had to be computed, for quantity of contact, direct and indirect cross-group friendship, and attitudes, and also extroversion (age was also included in the analysis as it demonstrates the technique in the clearest fashion):

- Discrepancy between male self-report and female self-report
- Discrepancy between male self-report and female observer-report
- Discrepancy between female self-report and male observer-report

- Discrepancy between male self-report and a random female self-report
- Discrepancy between male observer-report and a random female self-report
- Discrepancy between female observer-report and a random male self-report

The average discrepancy between male and female self-reports constitutes a dyadic index of the average *dissimilarity* of married couples on the variable of interest. The average discrepancy between pseudo-couples constitutes the dyadic index of dissimilarity expected by chance. These discrepancy scores are displayed in Table 3.2. Some mean scores are lower than the associated standard deviations, reflecting a non-normal distribution in the data. Skewed data are to be expected when using discrepancy scores, as when there is reason to expect similarity between the pairs of scores, discrepancy will tend towards zero, causing a skewed distribution. Pseudo-discrepancy scores will show a skewed distribution to the extent that all participants are similar on the constructs measured, while real discrepancy scores will be skewed to the extent that individuals are similar to their spouse. A paired-samples *t*-test was performed to test the differences between real and pseudo discrepancy levels for each dyad. The violation of the normality assumption would tend to decrease the power of the paired *t*-test, so the results are a fairly conservative test of the validity of self-reports of contact.

Table 3.1

Mean contact, attitudes, extroversion, closeness and demographic variables (Study 3)

	Female self-report	Female observer-report	Male self-report	Male observer-report
Age	49.32 (10.96)	51.73 (12.04)	51.55 (11.76)	49.34 (10.96)
Relationship length	25.26 (11.43)		25.07 (11.61)	
Closeness	6.32 (1.15)		6.52 (0.90)	
Extroversion	4.58 (1.10)	4.37 (1.08)	4.39 (1.06)	4.69 (1.09)
Attitudes	5.18 (1.03)	4.92 (1.12)	4.92 (1.12)	5.01 (1.12)
Quantity of contact	3.23 (1.42)	2.83 (1.13)	3.38 (1.46)	2.97 (1.55)
Direct cross-group friendship	1.92 (0.89)	1.90 (0.94)	1.87 (0.86)	2.01 (1.01)
Indirect cross-group friendship	2.31 (1.21)	2.13 (1.01)	2.36 (1.29)	2.30 (1.29)

Note: Standard deviations in parentheses

The results of these paired-samples *t*-tests are reported in Table 3.2. It can be seen, for example, that in pseudo-couples (a male and a female paired at random), age difference will be 12 years on average, while among married couples the average difference in age is less than four years. The *t*-test shows this difference to be significant. Despite the similarity expected due to assortative mating, couples are no more similar in their levels of extroversion than would be expected by chance, and show no more similar attitudes towards Asians than would exist between pseudo-couples. Couples are no more similar to each other in their levels of quantity of contact than they would be to a random person of the opposite sex, but they are more similar in terms of their direct and indirect cross-group friendship than would be expected by chance. This result is likely to be affected by couples sharing many of the same friends.

It can be seen that couples tend to be similar in terms of their direct and indirect Asian cross-group friendship, but that any other similarities found between spouses may be primarily due to the similarities that exist between people in general. However, the discrepancy of interest is that between self-reports of one member of the dyad, and observer-reports of the other member. Table 3.3 displays the discrepancies between self- and observer-reports, and those expected by chance, for males and females.

Table 3.2

Similarity in married couples: Real discrepancy and discrepancy expected by chance between male and female self-reports (Study 3)

	<i>Mean</i>	<i>Mean expected by chance</i>	<i>t (61)</i>
Age	3.81 (4.12)	12.19 (11.20)	-5.95 ****
Extroversion	1.21 (1.01)	1.29 (0.98)	-0.51
Attitude	1.04 (0.82)	1.27 (0.84)	-1.61
Quantity of contact	1.21 (1.14)	1.58 (1.26)	-1.62
Direct cross-group friendship	0.59 (0.71)	0.94 (0.77)	-2.57 *
Indirect cross-group friendship	0.89 (1.13)	1.33 (1.31)	-2.36 *

Note: Standard deviations in parentheses

**** $p < .001$

*** $p < .005$

** $p < .01$

* $p < .05$

Table 3.3

Discrepancy between self-reports and observer-reports: Real discrepancy and discrepancy expected by chance (Study 3)

	Discrepancy between male self-reports and female observer-reports			Discrepancy between female self-reports and male observer-reports		
	<i>Mean</i>	<i>Mean</i> expected by chance	<i>t</i> (61)	<i>Mean</i>	<i>Mean</i> expected by chance	<i>t</i> (61)
Age	0.18 (0.98)	12.44 (10.17)	-9.41 ****	0.05 (0.22)	11.44 (9.77)	-9.17 ****
Extroversion	0.63 (0.62)	1.11 (0.77)	-4.26 ****	0.74 (0.66)	1.23 (0.91)	-3.38 ***
Attitude	0.82 (0.79)	1.23 (0.85)	-3.10 ***	0.91 (0.77)	1.32 (0.86)	-2.73***
Quantity of contact	1.23 (1.10)	1.52 (1.12)	-1.51	0.95 (1.01)	1.66 (1.33)	-3.36 ***
Direct cross-group friendship	0.47 (0.57)	0.93 (0.78)	-3.48 ****	0.54 (0.64)	0.98 (0.90)	-3.12 ***
Indirect cross-group friendship	0.73 (1.09)	1.11 (1.22)	-2.05 *	0.94 (1.10)	1.35 (1.43)	-2.16 *

Note: Standard deviations in parentheses

**** $p < .001$

*** $p < .005$

** $p < .01$

* $p < .05$

Unsurprisingly, both men's and women's estimates of their partners' ages were significantly closer to their partners' actual ages than to the age of a partner picked at random. Self-reports of the very visible trait of extroversion also agreed with observer-reports significantly more than would be expected were the reports based purely on knowledge of people in general as opposed to knowledge of the individual in question. Contrary to the hypothesis, attitudes showed the same result; observer-ratings were not based on knowledge of the average person's attitudes to Asians, but on knowledge of the individual person being rated. The same was true of quantity of contact (female self-reports only⁷) and of direct cross-group friendship and, surprisingly, of indirect cross-group friendship. It would appear that observers have specific knowledge allowing them to rate their spouse on all measures, even that of indirect cross-group friendship. Considering that rating indirect cross-group friendship would require knowledge of the contact of the target's friends rather than of the target him or herself, it would not be surprising to find that observer-ratings on this measure were based on an assumed norm, reflecting stereotype accuracy rather than specific knowledge. However, it is possible that spouses share a great deal of the same friends, and that they therefore have just as much knowledge about their spouse's indirect cross-group friendship as about their own. These findings strongly support the validity of self-reports of contact, as observer-ratings agreed with self-ratings to a significant degree, in that they agreed significantly more than would be expected by chance.

The next stage of the analysis was to determine whether the level of self-disclosure ($M = 5.80$, $SD = 1.01$) between the couples influenced the (dis)agreement between self- and observer-reports. The scores reflecting the discrepancy between

⁷ This could be due to the reduction in power of the paired t -test caused by the violation of the assumption of normality.

self- and observer-reports were regressed onto self-disclosure. Each variable in turn was subjected to this analysis: extroversion, attitudes, quantity of contact, and direct and indirect cross-group friendship. No effects approached significance, suggesting that self-disclosure does not affect self-observer agreement in married couples. This may be due to a ceiling effect, as the mean level of self-disclosure was high. Indeed, 75% of the self-disclosure scores fell above 5.38 on a seven-point scale. It is unsurprising that married couples self-disclose to a high degree. It was therefore considered that a more appropriate analysis might consider length of relationship. It might be expected that couples who have been married longer would have had more opportunity to discuss, observe, or reflect upon the issues of interest.

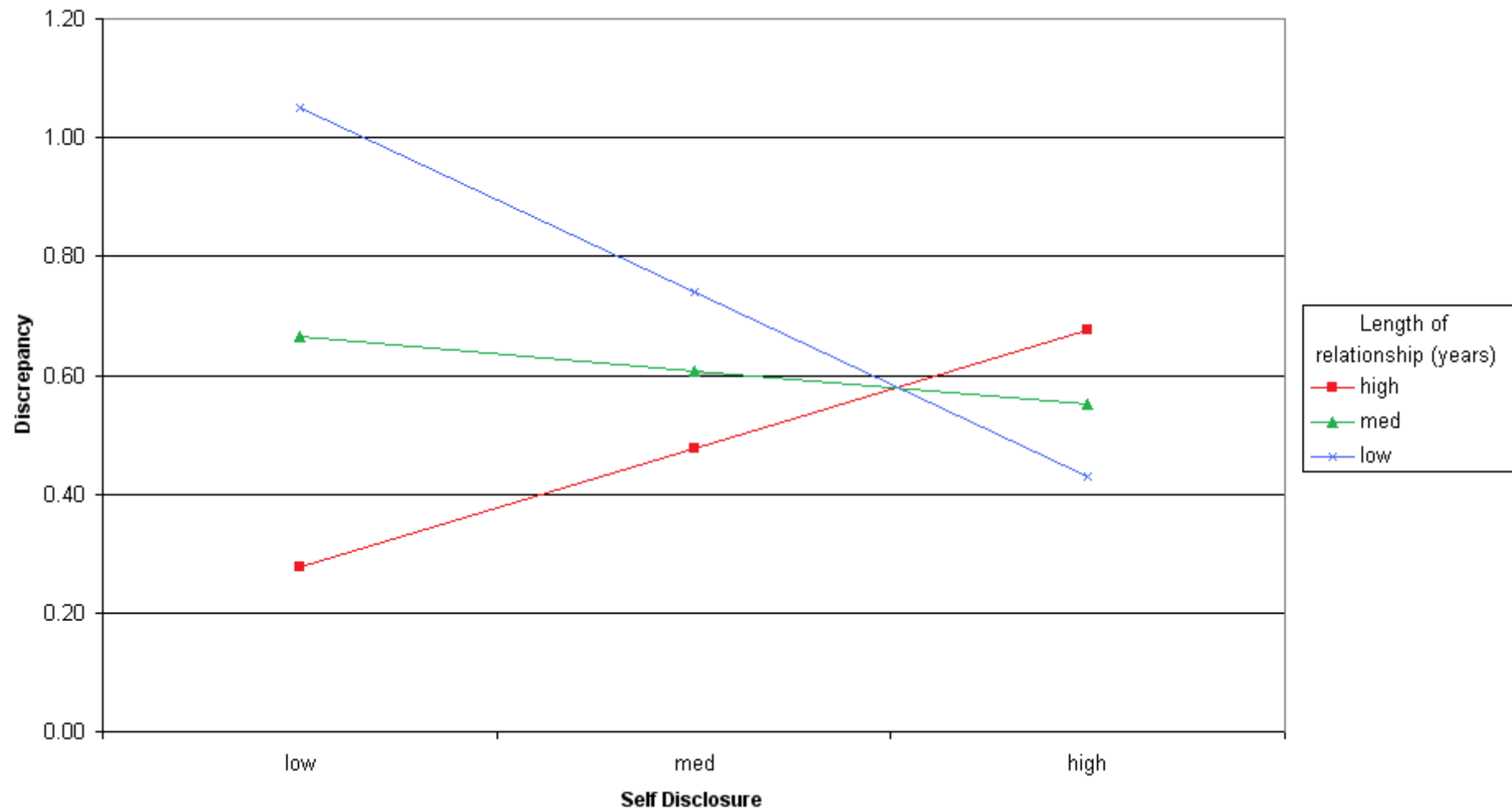
An average was taken of the length of relationship as estimated by each member of a couple, and self-observer discrepancy was regressed onto this variable. Again, no effects reached significance, although significance was approached for direct cross-group friendship. Length of relationship marginally negatively predicted discrepancy between male self-reports and female observer-reports of direct cross-group friendship, $\beta = -.24$, $p = .06$, and between female self-reports and male observer-reports, $\beta = -.22$, $p = .09$. This suggests that discrepancy was marginally higher for those couples who had been together a shorter time. This finding supports previous work which has shown that increased acquaintance leads to an increase in accuracy, but not to the extent that might be intuitively expected (Kenny, 2000). Kenny (2000) found that most of the increase in accuracy happened in the early stages of a relationship. It is probable that a ceiling effect is reached, beyond which members of a couple simply seek to confirm what they already know about a person, or what they wish to be true (Brownstein, 2003; Dunning, 2005; Ickes & Simpson, 1997; Murray, Holmes & Griffin, 1996; Sillars & Scott, 1983).

Finally, a moderated regression analysis was performed in which self-observer discrepancy concerning direct cross-group friendship was regressed onto self-disclosure, length of relationship, and the product of these two variables. The model explained a significant portion of the variance, $R^2 = .40$, $F(3,58) = 3.73$, $p = .02$. Self-disclosure predicted discrepancy between female self-reports and male observer-reports of direct cross-group friendship $\beta = -.95$, $p = .007$, as did length of relationship $\beta = -2.50$, $p = .005$, and the product was also highly significant $\beta = -2.50$, $p = .009$. Figure 3.1 illustrates the finding that self-disclosure increased self-observer agreement (i.e., decreased discrepancy) between female self-reports and male observer-reports of direct cross-group friendship for those couples who had been together a shorter time, but not for those who had been together a longer time.

Self-disclosure may fulfil different functions at different stages of a relationship. In the early stages of a relationship, it may be more important to disclose information about the self which is deeply personal, than in later stages, when such information has been already established. An alternative interpretation is that people in different stages of a relationship may hold different criteria concerning what constitutes 'intimate or personal issues'. A couple who have been together little more than a year may consider that talking about a certain topic might constitute an intimate or personal issue, whereas a couple who have already learned the most intimate personal details about each other may consider the same topic to be rather commonplace, and fail to include discussions about this topic in their self-reports concerning self-disclosure.

Figure 3.1

Moderating effect of length of relationship on the relationship between self-disclosure and discrepancy between male observer-reports and female self-reports of direct cross-group friendship



Study 4

Study 3 compared the self-reported contact and attitudes of each member of a married couple dyad with their contact and attitudes as rated by their spouse or legal partner. Self-reports of contact, attitudes, and extroversion were clearly validated by spouse-ratings. However, a large proportion of social-psychological research is conducted with university students as participants. Most university students are unmarried, and this means that the population for whom Study 3 validated self-reports of contact does not include many members of the population whose self-reports are most commonly used in contact research. It is important that the self-reports of this population are validated, using an alternative and more appropriate group of observers. Study 4 therefore examined dyads composed of a student and one of their parents.

Undergraduate students participated as targets, and each recruited one of their parents to participate as an observer. Because students experience two environments while they are at university, contact measures were included in Study 4 to assess the level of contact students experienced at university *and* in their home town. Casselden and Hampson (1990) point out that the situation in which the observer knows the target must be similar to the situation in which behaviour (or in this case, contact) is being measured, as targets may be perceived differently in different situations. It was therefore hypothesised that self-observer agreement would be higher for students' contact in their home town, as this contact would be more visible to parents.

Parents' racial attitudes tend to influence the racial attitudes of their children (Sinclair, Dunn & Lowery, 2005). It was therefore hypothesised that students' attitudes would be similar to the attitudes of their parents. However, students were expected to experience greater contact at university than at home, due to the fact that

Oxford University is a very cosmopolitan environment. Indeed, some participants in Study 3 wrote on their questionnaires that the explanation for what they saw as a low level of contact was that there were very few Asians living in their home town. Students' contact at university was therefore hypothesised to be greater than the contact of their parents. This increased contact in students might lead to a corresponding increase in the positivity of their attitudes towards Asians, when compared to the attitudes of their parents.

Study 4 was expected to replicate the results of Study 3, in that students' self-reports of direct cross-group friendship, and possibly attitudes, would be validated by observer-reports provided by parents. However, despite the failure of Study 3 to find a moderating effect of self-disclosure, it was predicted that such an effect would be found in Study 4, as a ceiling effect was not expected. The degree to which parents are able to predict students' contact, particularly at university, is likely to depend to some extent on what a student chooses to communicate to their parent. For this reason, self-disclosure was expected to moderate self-observer agreement on contact, especially at university.

Extroversion was again included as a visible trait to be rated, in order that self-observer agreement on this trait might be used as a baseline by which to compare self-observer agreement on other variables.

Hypotheses

1. Students' self-reports of direct cross-group friendship would be validated by observer-reports provided by parents
2. Self-observer agreement would be higher for students' contact in their home town, as this contact would be more visible to parents.

3. Self-observer agreement would be greater for visible constructs such as extroversion and direct cross-group friendship than for attitudes.
4. Self-disclosure would moderate self-observer agreement on contact, especially at university.
5. Students' attitudes would be similar to the attitudes of their parents
6. Students would report greater contact at university than at home, due to the fact that Oxford University is a very cosmopolitan environment.
7. Students' contact at university would therefore be greater than the contact of their parents.

Method

Participants

Seventy-one white British undergraduate students (50 female and 21 male, mean age 19 years) took part in the study, each in a dyad with one of their parents. Gender of the chosen parent was not prescribed, where students had more than one parent they made their own choice as to which parent would participate with them. Parents comprised 49 female and 22 male white British adults (mean age 52 years). Thus the dataset comprised 71 dyads composed of one undergraduate student still living at home in university holidays and one of their parents.

Procedure

Students recruited a parent to take part in the study and administered it during university vacation. Each member of a parent-child dyad filled in the questionnaire at the same time, in separate rooms, without conferring. Each member of a dyad answered questions about themselves and then about their parent/child, in a reciprocal dyadic design (Kenny, 1988).

Materials

The questionnaire consisted of measures of contact at home and contact at university (quantity of contact, and direct and indirect cross-group friendship), attitude towards Asians and extroversion. Participants completed self-ratings and then rated the other member of their dyad on all the same questions.

Following demographic questions, the questionnaire began with a section on preferences for various leisure activities; this was to fit the ostensible purpose of the questionnaire (to explore intra-familial knowledge); data from these questions were not analysed.

Extroversion was then measured using the same items as in Study 3. After reversing the '*Extroverted - Introverted*' and '*Active - Passive*' measures, the composite scale yielded a reliable index of student self-reports ($\alpha = .71$) student observer-reports ($\alpha = .82$), parent self-reports ($\alpha = .86$) and parent observer-reports ($\alpha = .82$), on which higher scores reflected higher levels of extroversion.

The order of the Contact measures and the Attitude measures was counterbalanced (no order effects were detected), and these were again preceded by the same note specifying the ethnic group concerned.

For the 'contact at home' measures, participants were provided with the instruction: 'These questions refer to contact you may have with Asians in your *home town*. If you are a student, try to think of your life before you went away to university.' For the 'contact at university' measures, participants were provided with the instruction 'These questions refer to contact you may have with Asians *at university*. If you are a parent, please skip this section.'

Quantity of contact was measured by two items: 'How often do you meet Asians in your everyday life at Home[or university]?' and 'How often do you *talk*

with people who are Asians?'. Answers were given on seven-point scales labelled with 'Never' (1), 'Rarely' (2), 'Sometimes' (3), 'Quite a bit' (4), 'Often' (5), 'Very often' (6) and 'All the time' (7). Inter-item correlation for student self-reports at home was $r = .82$, $N = 71$, $p < .001$, for student observer-reports $r = .92$, $N = 71$, $p < .001$, for parent self-reports $r = .84$, $N = 71$, $p < .001$, and for parent observer-reports at home $r = .86$, $N = 71$, $p < .001$. Inter-item correlation for student self-reports at university was $r = .80$, $N = 71$, $p < .001$, and for parent observer-reports at university $r = .81$, $N = 71$, $p < .001$.

Direct cross-group friendship was measured with two items: 'About how many of your friends at home [or university] are Asian?', with answers given on a seven-point scale labelled with 'None' (1), 'A few' (2), 'Less than half' (3), 'About half' (4), 'More than half' (5), 'Most' (6) and 'All' (7), and 'How often do you spend time with friends at home [or university] who are Asians?', with answers on the same seven-point scale as that presented with the 'Quantity' items. Inter-item correlation for student self-reports at home was $r = .76$, $N = 71$, $p < .001$, for student observer-reports $r = .68$, $N = 71$, $p < .001$, for parent self-reports $r = .85$, $N = 71$, $p < .001$, and for parent observer-reports at home $r = .73$, $N = 71$, $p < .001$. Inter-item correlation for student self-reports at university was $r = .62$, $N = 71$, $p < .001$, and for parent observer-reports at university $r = .52$, $N = 71$, $p < .001$.

Indirect cross-group friendship was measured with the same single item as in Study 3.

Attitudes were measured using the same measures as in Study 3. After transforming the feeling thermometer scores to put them on the same scale as the other attitude scores (by dividing all scores by 100 and multiplying by seven) and reversing 'Friendly - Hostile', 'Respect - Contempt' and 'Admiration - Disgust'

items, the items were collapsed to form reliable indices of student self-reports ($\alpha = .83$), student observer-reports ($\alpha = .78$), parent self-reports ($\alpha = .86$) and parent observer-reports ($\alpha = .87$) on which higher scores indicated more positive attitudes towards Asians.

After these 'self-ratings', participants were presented with the following paragraph: 'Now we would like to know about the person who is participating with you (your parent/child). The rest of this booklet contains questions similar to those you just answered, but they will relate to your parent/child; we will refer to this person as 'X'.' Participants were asked to provide demographic information about their parent/child, and to state their relationship to them (mother/father/son/daughter etc.). Participants also answered questions concerning the closeness of their relationship with their parent/child (a seven-point scale anchored at '*not at all close*' (1) and '*very close*' (7)) and the degree of self-disclosure between the two. Self-disclosure was measured with two items: 'How often do you talk about how you are feeling to X?', 'How often do you discuss intimate or personal issues with X?'; and the participant then answered questions about the other member of the dyad's self-disclosure: 'How often does X talk to you about how they are feeling?' and 'How often does X discuss intimate or personal issues with you?' Seven-point scales were presented, anchored at '*Never*' (1) and '*All the time*' (7), with '*Sometimes*' at the midpoint. In Study 3, an average of all self-disclosure scores was computed for use as a between-dyads variable, but it was expected that self-disclosure would act as a within-dyads variable in Study 4. For this reason, average scores were computed for student self-reports, $r = .69$, $N = 71$, $p < .001$, student observer-reports, $r = .63$, $N = 71$, $p < .001$, parent self-reports, $r = .68$, $N = 71$, $p < .001$, and parent observer-reports, $r = .70$, $N = 71$, $p < .001$.

Participants were then given the ‘self-rating’ questions again, but this time the questions were reworded so that they asked the participant to rate the other member of their dyad.

Results and Discussion

Means and standard deviations are reported in Table 3.4. Preliminary nomothetic analyses revealed that students tended to report a higher degree of direct, $t(70) = 2.99, p = .004$, and indirect cross-group friendship at home, $t(70) = 5.25, p < 0.001$, than parents did, and were more extroverted, $t(70) = 2.42, p = .02$. As hypothesised, students were found to experience greater contact at university than at home, in terms of their quantity of contact, $t(70) = 3.97, p < 0.001$, their direct cross-group friendship, $t(70) = 3.91, p < 0.001$, and their indirect cross-group friendship, $t(70) = 2.82, p = .006$. It was hypothesised that students’ attitudes would be similar to those of their parents (Sinclair et al., 2005). Mean attitudes did not differ between students and parents, in support of this hypothesis, $t(70) = .50, ns$, and neither did mean self-disclosure, $t(70) = .17, ns$.

Table 3.4

Mean contact at home and university, attitudes, extroversion, self-disclosure and demographic variables (Study 4)

	Student	Parent	Parent
	self-report	self-report	observer-report
Age	18.99 (1.13)	51.55 (4.44)	19.03 (1.11)
Closeness	6.14 (0.80)	6.48 (0.79)	
Extroversion	4.86 (0.81)	4.48 (1.04)	5.13 (0.92)
Attitudes	5.08 (0.89)	5.02 (0.89)	5.22 (0.86)
Quantity of contact at home	3.12 (1.22)	3.34 (1.45)	2.98 (1.33)
Direct cross-group friendship at home	1.97 (0.87)	1.64 (0.70)	1.94 (0.81)
Indirect cross-group friendship at home	3.32 (1.53)	2.23 (0.96)	2.44 (1.17)
Quantity of contact at university	4.02 (1.48)		3.80 (1.26)
Direct cross-group friendship at university	2.52 (0.93)		2.56 (0.82)
Indirect cross-group friendship at university	3.91 (1.57)		2.69 (1.15)
Self-disclosure	4.77 (1.23)	4.85 (1.13)	4.34 (1.25)

Note: Standard deviations in parentheses

However, this analysis compares only the average student and the average parent, not each student with his or her *own* parent. The idiographic analysis described in Study 3 was performed in the same way in Study 4; three discrepancy scores were computed for each dyad on each variable, followed by three pseudo-couple discrepancy scores. The average discrepancies between student and parent self-reports are displayed in Table 3.5, along with the average discrepancy between pseudo-couples - the discrepancies expected by chance. The results of the paired-

samples *t*-test are also presented, testing the differences between real and pseudo discrepancy levels for each dyad.

Table 3.5

Similarity in parent-child dyads: Real discrepancy and discrepancy expected by chance between student and parent self-reports (Study 4)

	<i>Mean</i>	<i>Mean expected</i> by chance	<i>t</i> (70)
Closeness	0.76 (0.80)	0.90 (0.81)	-1.28
Extroversion	1.10 (0.80)	0.98 (0.74)	1.02
Attitudes	0.95 (0.67)	0.97 (0.63)	-0.23
Quantity of contact at home	1.05 (1.06)	1.47 (1.23)	-2.41 *
Direct cross-group friendship at home	0.65 (0.74)	0.96 (0.80)	-3.00 ***
Indirect cross-group friendship at home	1.46 (1.40)	1.49 (1.54)	-0.21
Self-disclosure	1.24 (0.97)	1.56 (1.09)	-2.27 *

Note: Standard deviations in parentheses

**** $p < .001$

*** $p < .005$

** $p < .01$

* $p < .05$

Students' quantity of contact and direct cross-group friendship at home are more similar to those of their parents than to those of an adult chosen at random. Also, a student and his or her parent tend to self-disclose to a more similar degree than would be expected by chance. However, any other similarities between a student and his or her parent are due to similarities between the average student and the average adult. The hypothesis that students would show similar attitudes to those of their parents was not borne out, despite the fact that the nomothetic analysis revealed no

difference between the means. While the average attitudes of students were similar to the average attitudes of parents, students' attitudes were no more similar to those of their parents than would be expected by chance.

The most important discrepancy scores are those which describe the (dis)similarity between self- and observer-ratings. These are displayed in Table 3.6.

Table 3.6

Discrepancy between self-reports and observer-reports: Real discrepancy and discrepancy expected by chance (Study 4)

	<i>Mean</i>	<i>Mean discrepancy</i>	
	discrepancy	expected by chance	<i>t</i> (70)
Age	0.04 (0.20)	1.14 (1.12)	-7.98 ****
Gender	0.00 (0.00)	0.48 (0.50)	-8.02 ****
Extroversion	0.60 (0.60)	1.10 (0.71)	-5.07 ****
Attitudes	0.92 (0.79)	0.93 (0.69)	-0.04
Quantity of contact at home	0.89 (1.01)	1.41 (1.21)	-3.11 ***
Direct cross-group friendship at home	0.56 (0.56)	0.98 (0.85)	-3.83 ****
Indirect cross-group friendship at home	1.19 (1.21)	1.67 (1.47)	-2.94 ***
Quantity of contact at university	1.38 (1.13)	1.47 (1.25)	-0.49
Direct cross-group friendship at university	0.72 (0.61)	1.06 (0.78)	-3.56 ****
Indirect cross-group friendship at university	1.68 (1.35)	1.75 (1.44)	-0.43
Self-disclosure	0.92 (0.63)	1.42 (1.02)	-3.84 ****

Note: Standard deviations in parentheses

**** $p < .001$

*** $p < .005$

** $p < .01$

* $p < .05$

Students' self-reports of extroversion agree with parents' observer-reports, as was expected due to the high visibility of this trait (Borkenau & Ostendorf, 1992; Costa & McCrae, 1988; Emler et al., 2007; Ones et al., 1995; Piedmont et al., 2000). Students' self-reports of contact at home were all validated by parents' observer-reports, whether quantity of contact, direct cross-group friendship, or, more surprisingly, indirect cross-group friendship. All these variables showed less discrepancy between self-and observer-ratings than would be expected by chance. However, the same was not true of attitudes. Parents' ratings of their own child's attitudes were just as similar to their own child's self-reports as to any other student's self-reports. Attitudes are a more subjective construct and arguably more difficult to obtain knowledge about than contact, but it is also likely that attitude self-reports are more susceptible to SDR than are contact self-reports. As hypothesised, self-observer agreement was lower for contact at university. There was no greater self-observer agreement than would be expected by chance for quantity of contact or indirect cross-group friendship, although direct cross-group friendship at university was validated. Parents' ratings of how much their child self-discloses to them seem to agree with students' self-ratings.

The next step was to discover whether level of self-disclosure affected self-observer (dis)agreement. Length of relationship was, of course, not measured in Study 4. A regression analysis was performed for each variable in turn in which discrepancy between self- and observer-reports was regressed onto an average student's self-disclosure score, computed using students' self-report and parents' observer-report of self-disclosure, $M = 4.81$, $SD = 1.04$, $\alpha = .88$.

Students' self-disclosure predicted discrepancy between parents' observer-ratings and students' self-ratings of quantity of contact at home, $\beta = -.23$, $p = .05$, and

of indirect cross-group friendship at home, $\beta = -.25$, $p = .04$. The more a student tends to disclose to their parent, the smaller the discrepancy between the parent's and the student's ratings of how much contact and indirect cross-group friendship the student has at home.

These findings support the hypothesis that self-observer agreement would depend on the student's level of self-disclosure. However, they do not support the secondary hypothesis that this would be particularly true for self-reports at university. It is possible that ratings of self-disclosure were based on self-disclosure which took place face-to-face, at home, rather than at a distance (i.e., by telephone or email), while at university. As all participating students were in their first term at university, these self-disclosure ratings would have been based largely on their relationship with their parents before coming to university. It is also, likely, although rather sad, that self-disclosure decreased somewhat once students went away to university, and that parents were therefore less knowledgeable about the people their child spends time with at university than at home.

These findings support the hypothesis that self-reports of contact at home would be validated, but self-reports of contact at university have not been adequately validated.

Study 5

Study 5 closely replicated Study 4, with one important addition. Study 5 attempted to demonstrate discriminant validity of self-ratings, by exploring whether participants can distinguish between contact and attitudes in two situations: Time 1 - life at home before coming to university, and Time 2 – current life at university. After one term at university, participants were asked to rate their current contact and attitudes, and to estimate the contact and attitudes they had experienced before

coming to university. Participants also recruited one of their parents to act as an observer, and provide ratings of the student's contact and attitudes at home prior to coming to university, and of the student's current contact and attitudes. In this way, discriminant validity could be demonstrated by testing whether self-observer agreement disappeared when cross-matching the Time being rated.

Hypotheses

1. Retrospective self-ratings of Time 1 would agree with observer-reports of Time 1.
2. Self-ratings of Time 2 would agree with observer-reports of Time 2, but to a lesser degree due to parents being less able to make informed judgements about situations at a distance.
3. Demonstrating discriminant validity, this self-observer agreement should *not* appear when cross-matching Time 1 with Time 2. Self-ratings of Time 1 would *not* agree with observer-ratings of Time 2, and self-ratings of Time 2 would *not* agree with observer-ratings of Time 1.
4. Self-observer agreement would be greater for visible constructs such as extroversion and direct cross-group friendship than for attitudes.
5. Self-reports of indirect cross-group friendship would prove impossible to validate due to the difficulty of estimating the contact of all the target's friends.
6. Self-reported contact with Asians would increase from Time 1 to Time 2 due to the more diverse environment in which participants found themselves.
7. There would be a corresponding increase in the positivity of participants' self-reported attitudes towards Asians.

Method

Participants

A total of 62 female and 27 male white British undergraduate students between the ages of 18 and 24 (mean age = 19.25) took part in the study, each in a pair with one of their parents. Parents comprised 58 female and 31 male white British adults between the ages of 38 and 63 (mean age = 50.87). In total, 178 white British individuals (120 females and 58 males) aged between 18 and 63 took part in the research. They comprised 89 dyads, each dyad being composed of one undergraduate student living at home in university holidays, and one of their parents or life-long (as opposed to recently installed) guardians.

Procedure

The same procedure was followed as in Study 4.

Materials

Materials were as in Study 4, with some changes and additions. Unlike in Study 4, two questionnaires were prepared, one for parents and one for students. Following demographic questions, each questionnaire began with a section on preferences for various leisure activities, as in Study 4. Parents were then asked for their own self-reports of extroversion. Extroversion measures were followed with the instruction 'The following questions concern topics which are less comfortable to talk and think about. We understand some of these questions may be difficult but please try to answer them to the best of your ability.' Parents were next asked to report on their attitudes and contact, and were then provided with the following instruction: 'Now we would like to know about your son/daughter who is participating with you. The rest of this booklet contains questions similar to those you just answered, but they will relate to your child (we will refer to him/her as 'X'). We understand that many of

the following questions may be difficult, but please try to answer every question to the best of your ability'. They were then asked to report on their child's current levels of extroversion, attitudes and contact at university, and to rate what their child's levels of extroversion, attitudes and contact had been before they went to university. The order of past and current questions was counterbalanced, as was the order of attitude and contact measures.

Students were simply asked to report on their own current levels of extroversion, attitudes and contact, and to estimate what their levels were before they started at university. The order of past and current questions, as well as of attitude and contact measures, was counterbalanced (no order effects were detected). Reliability statistics are displayed in Table 3.7.

Results and Discussion

The mean levels of extroversion, attitudes and contact of students at Time 1 and Time 2, and of parents at Time 2, are displayed in Table 3.7. Parents' observer-reports of students at Time 1 and Time 2 are also presented. Preliminary nomothetic analyses revealed that, as hypothesised, self-reported contact with Asians increased from Time 1 to Time 2 in terms of quantity of contact, $t(88) = 2.19, p = .03$, and indirect cross-group friendship, $t(88) = 3.77, p < 0.001$, and marginally in terms of direct cross-group friendship, $t(88) = 1.67, p = .09$, due to the cosmopolitan environment in which participants found themselves. However, the hypothesised corresponding increase in the positivity of participants' attitudes towards Asians was only marginal, $t(88) = 1.65, p = .10$. Extroversion was found to increase at university, $t(88) = 3.94, p < .001$, possibly due to an increase in confidence resulting from the added independence.

Table 3.7

Reliability statistics for self- and observer-reported extroversion, attitudes and contact at Time 1 and Time 2 (Study 5)

	Student self- report, Time 1	Student self- report, Time 2	Parent observer- report, Time 1	Parent observer- report, Time 2	Parent self- report, Time 2
<i>N</i> = 89					
Extroversion	$\alpha = .85$, 5 items	$\alpha = .88$, 5 items	$\alpha = .87$, 5 items	$\alpha = .84$, 5 items	$\alpha = .83$, 5 items
Attitudes	$\alpha = .93$, 5 items	$\alpha = .92$, 5 items	$\alpha = .89$, 5 items	$\alpha = .88$, 5 items	$\alpha = .84$, 5 items
Quantity of contact	$r = .87$, $p < .001$	$r = .87$, $p < .001$	$r = .94$, $p < .001$	$r = .93$, $p < .001$	$r = .86$, $p < .001$
Direct cross-group friendship	$r = .65$, $p < .001$	$r = .53$, $p < .001$	$r = .74$, $p < .001$	$r = .63$, $p < .001$	$r = .76$, $p < .001$
Indirect cross-group friendship	$r = .90$, $p < .001$	$r = .87$, $p < .001$	$r = .97$, $p < .001$	$r = .92$, $p < .001$	$r = .75$, $p < .001$
Self-disclosure		$\alpha = .91$, 4 items		$\alpha = .89$, 4 items	$\alpha = .88$, 4 items

Table 3.8

Mean self- and observer-reported levels of extroversion, attitudes and contact at Time 1 and Time 2 (Study 5)

	Student self- report, Time 1	Student self- report, Time 2	Parent observer- report, Time 1	Parent observer- report, Time 2	Parent self- report, Time 2
(N = 89)					
Extroversion	4.54 (1.10)	4.90 (1.01)	4.82 (1.02)	5.06 (0.91)	4.44 (1.01)
Attitudes	5.29 (1.11)	5.40 (1.08)	4.31 (0.80)	4.35 (0.80)	4.16 (0.78)
Quantity of contact	3.73 (1.76)	4.21 (1.59)	3.29 (1.47)	3.89 (1.38)	3.04 (1.37)
Direct cross-group friendship	3.02 (1.53)	3.30 (1.21)	2.42 (1.27)	2.52 (0.85)	1.74 (0.74)
Indirect cross-group friendship	3.64 (1.79)	4.37 (1.70)	2.63 (1.47)	2.84 (1.31)	2.02 (0.77)
Self-disclosure		4.53 (1.46)		4.81 (1.28)	4.73 (1.36)

Note: Standard deviations in parentheses.

In order to examine differences between students and parents, dyadic analyses are again necessary. As in Study 4, discrepancy scores were calculated to discover the degree to which a student is similar to his or her parent, and to examine self-observer agreement. Table 3.8 displays discrepancies between parents' self-reports and students' self-reports of Time 1 and Time 2. It might be expected that a student will show more similarity to his or her parent before attending university than after. This result is found, marginally, in analyses of extroversion⁸. The discrepancy between a student's level of extroversion and his or her parent's level at Time 2 is similar to that expected by chance, but at Time 1 is marginally less than would be expected by chance. This suggests that students tend to have a similar level of extroversion to that of their parents before attending university, but this similarity is attenuated when students spend time at university. Previous research has revealed that, when assessed in adulthood using twin and adoption studies, *non-shared* rather than *shared* environmental factors constitute the major source of environmental influence on most psychological traits (Eysenck, 2008; Plomin & Daniels, 1987). In other words, the non-genetic component of personality tends to be affected very little by parental influence (which is *shared* between siblings) but rather by experiences outside the home (which will *differ* between siblings). Presumably, before adulthood when the student still lives at home full time, the influence of his or her parents is inflated.

The only significant result was that students tended to self-disclose to a similar extent to their parents. However, in all other aspects, students' similarity to their parents was no greater than the average student's similarity to the average adult.

⁸ Using a nomothetic analysis, however, this result is clear: extroversion is similar to parents' before attending university, $t(88) = .72$, ns, but not after, $t(88) = 3.66$, $p < .001$. Partialling out stereotype accuracy decreases the significance of any result which tends to be similar across all dyads.

The discrepancy scores which examine self-observer agreement are the most important in terms of validating self-reports of contact. These are reported in Table 3.9. The hypotheses were fully supported, in that self-observer discrepancy in contact self-reports at Time 1 (when students lived with their parents before attending university) was very significantly less than would be expected by chance. The same was true to a lesser extent at Time 2, as hypothesised, because parents would be less aware of contact that happens in a different location from the one they had shared with their child. Self-reported attitudes were not validated either at Time 1 or Time 2, possibly because they are more likely to be susceptible to SDR than the arguably more objective contact reports. A lower level of self-observer agreement was hypothesised for the less visible constructs, and the lack of self-observer agreement regarding attitudes supports this hypothesis. However, it was hypothesised that self-reports of indirect cross-group friendship would not be validated due to the difficulty of estimating the contact of all the target's friends. This did not prove to be the case; parents' estimates of indirect cross-group friendship agreed with students' self-reports, particularly at Time 1, when students were living with parents.

Table 3.9

Students' similarity to their parents before and after attending university: Discrepancy and discrepancy expected by chance (Study 5)

	Discrepancy	Discrepancy expected		Discrepancy	Discrepancy expected	
	at Time 1	by chance at Time 1	<i>t</i> (88)	at Time 2	by chance at Time 2	<i>t</i> (88)
Extroversion	0.99 (0.74)	1.20 (0.82)	-1.96 (*)	1.04 (0.69)	1.20 (0.84)	-1.44
Quantity of contact	1.65 (1.24)	1.84 (1.43)	-1.06	1.77 (1.33)	1.81 (1.39)	-0.26
Direct cross-group friendship	1.55 (1.25)	1.73 (1.27)	-1.78	1.68 (1.13)	1.80 (1.24)	-1.14
Indirect cross-group friendship	1.83 (1.60)	2.02 (1.64)	-1.84	2.52 (1.55)	2.52 (1.60)	0.00
Attitude to Asians	1.34 (0.89)	1.45 (1.08)	-1.11	1.42 (0.90)	1.53 (1.08)	-1.08
Self-disclosure				1.15 (0.96)	1.56 (1.09)	-2.83 **

Note: Standard deviations in parentheses.

** $p < .01$

(*) $p < .10$

Table 3.10

Self-observer agreement before and after attending university: Discrepancy and discrepancy expected by chance (Study 5)

	Discrepancy	Discrepancy expected		Discrepancy	Discrepancy expected	
	at Time 1	by chance at Time 1	<i>t</i> (88)	at Time 2	by chance at Time 2	<i>t</i> (88)
Extroversion	0.77 (0.66)	1.20 (0.96)	-3.78 ****	0.76 (0.66)	1.04 (0.83)	-2.40 *
Quantity of contact	0.96 (1.07)	1.75 (1.32)	-4.88 ****	1.19 (1.06)	1.65 (1.22)	-3.01 ***
Direct cross-group friendship	1.01 (0.95)	1.48 (1.14)	-3.51 ****	1.06 (0.87)	1.33 (1.04)	-2.68 **
Indirect cross-group friendship	1.40 (1.36)	1.98 (1.62)	-3.59 ****	1.93 (1.44)	2.34 (1.40)	-2.51 *
Attitude to Asians	1.20 (0.76)	1.29 (0.99)	-0.84	1.22 (0.84)	1.34 (1.04)	-1.19
Self-disclosure	0.77 (0.66)	1.20 (0.96)	-3.78 ****	1.04 (0.84)	1.56 (1.13)	-3.77 ****

Note: Standard deviations in parentheses.

**** $p < .001$

*** $p < .005$

** $p < .01$

* $p < .05$

Study 5 now provides an extension of Study 4, in that it examines discriminant validity. Self- and observer-reports were obtained for Time 1 and Time 2, and it has been found that observer-reports of Time 1 agree with self-reports of Time 1, and *vice-versa*. However, these can also be cross-matched, to discover whether observer-reports of Time 2 fail to agree with self-reports of Time 1, and *vice-versa*. By calculating the discrepancy between Time 1 self-reports and Time 2 observer-reports, and between Time 2 self-reports and Time 1 observer-reports, it is possible to test for discriminant validity. It was hypothesised that discriminant validity would be obtained, or in other words that self-observer agreement would *not* be found for the cross-matched discrepancy scores. The results of this analysis are displayed in Table 3.10.

This table clearly shows that there is no more self-observer agreement than would be expected by chance when cross-matching Time 1 and Time 2 self- and observer-reports of contact or attitudes. Observers are not basing their ratings on, for example, some stable aspect of the target's personality which might give them clues about their level of contact across all situations and time periods, but on what they know about how much contact the target actually has at a given time and in a given situation. This strongly demonstrates the discriminant validity of contact self-reports; a student is able to validly rate their current level of contact as well as their level of contact at an earlier date. This suggests, despite the difficulty of obtaining accurate self-insight (Haddock, Rothman, Reber, & Schwarz, 1999), and the inaccessibility of infrequently expressed constructs (Basilli, 1993, 1996a, 1996b), that an individual's level of contact is accessible to that individual, and available to be rated.

The significant agreement shown for extroversion suggests that, in the parent's view, there is little difference in the student's level of extroversion between Time 1 and Time 2. As can be seen in table 3.10, parents' ratings agree more with students' self ratings of extroversion at Time 1 than at Time 2, suggesting that parents may infer extroversion at Time 2 on the basis of extroversion at Time 1.

Table 3.11

Discriminant validity: Cross-matched discrepancy and cross-matched discrepancy expected by chance (Study 5)

	Discrepancy between Time 1 self-reports and Time 2 observer-reports	Discrepancy expected by chance	<i>t</i> (88)	Discrepancy between Time 2 self-reports and Time 1 observer-reports	Discrepancy expected by chance	<i>t</i> (88)
Extroversion	0.85 (0.65)	1.21 (0.96)	-3.24 ***	0.84 (0.72)	1.13 (0.83)	-2.38 *
Quantity of contact	1.72 (1.28)	1.76 (1.25)	-0.20	1.61 (1.30)	1.80 (1.30)	-1.19
Direct cross-group friendship	1.34 (0.98)	1.44 (1.02)	-0.93	1.37 (1.05)	1.63 (1.21)	-1.98
Indirect cross-group friendship	1.66 (1.48)	1.79 (1.43)	-0.74	2.15 (1.55)	2.46 (1.63)	-1.66
Attitude to Asians	1.17 (0.78)	1.29 (0.98)	-1.27	1.31 (0.77)	1.39 (1.01)	-0.70

Note: Standard deviations in parentheses.

**** $p < .001$

*** $p < .005$

** $p < .01$

* $p < .05$

General Discussion

Three studies have contributed to the research literature on intergroup contact by demonstrating for the first time the validity of self-reports of contact through the use of observer-reports, widely seen as the best method developed to date (Howard, 1994; Piedmont et al., 2000). Observers who know the target well are vital when using their reports to validate measures of less easily visible aspects of behaviour (Cronbach et al., 1972; Sudman & Bradman, 1974). These studies used observers who were most likely to know the target better than anyone else – in the case of married adults, their spouse, and in the case of university students, their parents.

Adults' self-reports of quantity of contact and direct and indirect cross-group friendship were validated in Study 3 through observer-reports provided by the spouse. Spouse-ratings tend to be reliable (Costa & McCrae, 1988), and spouses make better observers than peers due to their intimate knowledge of the target (McCrae, 1993). It was found that the level of self-disclosure did not affect self-observer agreement in couples who have been together a longer time, but that, in couples who have been together for a shorter time, those who self-disclose more were more likely to show agreement between self-reports and observer-reports. I argue that this is a result of the different functions fulfilled by self-disclosure at different stages of a relationship, and that during the early stages, self-disclosure is more important for knowledge of the other person than in later stages when intimate knowledge has already been established.

Study 4 validated the self-reports of students, a population commonly used in contact research, concerning quantity of contact, direct cross-group friendship and indirect cross-group friendship in the home town. However, as observers were parents, and were not present at university to witness contact, only self-reports of

direct cross-group friendship were validated at university, presumably because students had shared information with their parents about their university friends. Study 4 also found that self-observer agreement was higher on quantity of contact and indirect cross-group friendship at home for those students who self-disclosed to their parents to a greater degree.

Finally, Study 5 validated students' self-reports of quantity of contact, direct cross-group friendship and indirect cross-group friendship at home and at university, but found again that parents are less able to estimate contact at university due to their being at a distance from the university location. Study 5 demonstrated the discriminant validity of contact self-reports by showing that participants can distinguish between levels of contact in two different situations and at two different times. Observers are not simply rating a stable aspect of the target which gives them a basis for estimating their level of contact overall, but are able to distinguish between two time-periods.

It must be noted here that Study 5 examined concurrent and also retrospective self-reports. Retrospective self-reports have been shown to be unreliable due to their combining the flaws of self-reports with the flaws of autobiographical memory (Henry, Moffitt, Caspi, Langley & Silva, 1994; Hill & Betz, 2005). However, it is sometimes necessary to obtain information about the past through self-reports, such as when pretest data are unavailable (Campbell, Stanley & Gage, 1963). Study 5 can be said to have validated retrospective self-reports of contact, as students' ratings of their contact at a previous time in their lives were validated by observer-ratings made by their parents. It could be argued that parents' observer-reports are also retrospective in that they refer to a previous time period. However, I argue that due to parents' intimate knowledge of their children over the period of many years while they lived at

home, it is unlikely that a significant amount of memory loss could occur during one university term. In any case, Study 5 demonstrates that retrospective self-reports and retrospective observer-reports agree. As it is unlikely that both would err in the same way, this would appear to suggest they capture substantive rather than purely error variance.

The technique used in all three studies was suitable for analysis of dyadic data, and for dealing with non-independence (Kenny et al., 2006). The additional use of pseudo-couple analyses controlled for any agreement between self- and observer-reports which was due simply to norm-based heuristics. These techniques provide conservative tests of the validity of self-reports, and the consistency in which their validity has been demonstrated throughout lends greater weight to the conclusion that self-reports are valid for use in contact research.

A potential weakness of this study was that participants always completed self-reports before observer-reports, in order to get used to the questionnaire before completing the, more difficult, observer-ratings. The issue of self-anchoring is discussed, and steps taken to resolve it, in Study 7.

These findings robustly support the validity of self-reports of contact, particularly contact in the home town. However, they also highlight that choice of observer is important, as parents are less appropriate for validating self-reports of contact at university than of contact at home. Much contact research is carried out at universities, and uses self-reports provided by students of the contact they experience at university. To validate these self-reports more fully, it would be more appropriate to use observers who share a common location and situation with the target.

CHAPTER FOUR:

Validation of Contact Self-Reports using Multiple Observer-ratings

Chapter Three showed that self-reports of contact tend to agree with the observer-reports of close others. However, two important limitations should be addressed. Firstly, Study 3 excluded all non-married participants, which excluded almost all of the most commonly used participant pool; that of university students. Studies 4 and 5 focused on that participant pool, but used as observers individuals who are unlikely to be able to observe much of the participant's contact, as it takes place at university, away from the parents' watchful eyes. I suggest that parents would be an ideal source of observer-ratings for school-age children, but for university students, as shown by studies 4 and 5, parents are more suited to providing observer-ratings of students' contact in their home towns than at university. For this important participant pool, alternative observers must be found for validating self-reports of contact at university.

Secondly, and more importantly, the use of a single observer exposes observer-reports to the error variance created by individual response styles. In addition to response styles, different observers may perceive an intergroup encounter differently and make very different inferences about those involved in the contact (e.g., Correll, Park, Judd & Wittenbrink, 2002; Duncan, 1976; Sagar & Schofield, 1980). Increasing the number of observers would increase validity (Kolar, Funder, & Colvin, 1996) by controlling for response biases in each individual observer, as these biases are unlikely to be shared by all observers. Kenny (1994) suggests that the part of the observers' judgements which is consensual is free from error variance, and provides an ideal criterion by which to validate self-reports. Consensus tends to vary

between $r = .4$ and $r = .6$ (McCrae & Costa, 1989), so such a criterion should not prove too difficult to produce.

Kenny's (1994) 'round-robin' design seems to provide a suitable method for the current research. Robins et al. (2004) and Emler et al. (2007) used this design, in which participants take part in groups of four, with each member rating not only themselves but every other member on each relevant trait or attribute. Diminishing returns are seen when more than two observers are used (McCrae & Costa, 1989) and four observers seems to be the practical maximum (McCrae & Weiss, 2007). The optimal situation would involve groups of four people who are all mutually close friends. However, Robins et al. (2004) used participants who were not acquainted with each other at all (necessarily, as this was ostensibly a set of 'getting acquainted' meetings), and who therefore could not rate others as accurately as if they knew them. Emler et al. (2007) compared self- and observer-reports of political identities which, they argued, and found, tend to be intelligibly communicated to others, particularly in college-educated populations (Emler, 2002). They used participants with a median length of acquaintance of 16 months. However, a weakness of their study was that the lowest length of acquaintance was only one month. It must also be pointed out that data were not gathered in the laboratory under controlled conditions, and were not analysed with a technique suitable for round-robin designs. The current studies made use of the round-robin design and addressed the limitations of prior research by (a) recruiting participants who had been acquainted for at least four months, and requiring that group members be 'close friends' rather than mere 'acquaintances', (b) gathering data in the laboratory under controlled conditions, and (c) analysing the data with a program developed specifically for the analysis of round-robin data.

Non-independence

Chapter Three outlined the problems of analysing non-independent dyadic data as if they were independent individual data. Here the problem exists more prominently, as groups of four mutually-acquainted individuals are rating each other. Data from each member of the group will be affected by their relationships with the three other members of the group, non-independence resulting from each relationship. Variation in this case comes from four separate sources. There is error variance, of course, and there is a rater effect, a target effect, and a relationship effect. The rater effect refers to that variance which is caused by the rater, the person providing the observer-report. Any response strategy he or she uses will contribute to the rater effect. The target effect refers to variance attributable to the target, the person being rated. The attributes of the person being rated will affect the target effect, to the extent that they are visible to others. This is the effect which captures the observer-rating of interest; the part of the rating which is due purely to the observable characteristics of the target. Finally there is a relationship effect, referring to the variance attributable to the effects of the interaction between the rater and target. If a particular target is judged by a particular rater differently than would be predicted based on the rater's rater effect and the target's target effect, this would be attributable to a relationship effect. The Social Relations Model (SOREMO, (Kenny, 1998; 1994; Kenny & La Voie, 1984; Malloy & Kenny, 1986) has the capacity to split the variance in the ratings into these separate components, so that each may be analysed separately.

Study 6

The current studies were able to counteract the problem of non-independence when using a round-robin design (Kenny, 1994). Close acquaintances were recruited in groups of four, and came into the laboratory to make judgments of each other and

of themselves, on judgmental dimensions of outgroup attitudes, degree of intergroup contact, quality of intergroup contact, and a variety of other measures (such as extroversion). The primary goals were to examine whether others (observers or “raters”) made consensual judgments of targets’ degree and quality of intergroup contact, and whether these consensual judgments were correlated with the targets’ own self-ratings. This would provide evidence for the validity of self-reports of intergroup contact. An additional issue of interest is whether observers showed consensus in judgments of targets’ outgroup attitudes that also agreed with self-reports. Finally, by also measuring self-reports and observer-reports of traits such as extroversion these studies hoped to demonstrate discriminant validity, i.e., when observers were judging a target’s intergroup contact they were not simply judging how outgoing the target was, a judgment that past research has suggested can be made consensually and accurately (Borkenau & Liebler, 1995; Borkenau & Ostendorf, 1992; Costa & McCrae, 1988; Emler et al., 2007; Ones et al., 1995; Piedmont et al., 2000).

Hypotheses

1. Observers would show consensus in their ratings of the target’s contact with the outgroup.
2. These consensual ratings would agree with the target’s self-ratings of contact with the outgroup.
3. Observers would show consensus, possibly to a lesser degree, in their ratings of the target’s attitudes towards the outgroup.
4. These ratings would agree, again to a lesser degree, with targets’ self-rated attitudes towards the outgroup.

5. Self-observer agreement would remain when controlling for self- and – observer-ratings of extroversion, demonstrating discriminant validity in that observers are rating what they are asked to rate, rather than basing their judgements on how outgoing the target is.

Method

Participants

One hundred and twenty-eight students at a British university (90 females, 37 males, one undisclosed; mean age = 21 years) took part in the research. They comprised 32 groups of four, group members being well acquainted with each other for at least four months (mean acquaintance time = 19.09 months, $SD = 8.29$). One group had to be removed from the analysis due to missing data and the fact that one of the group members was of Asian origin. The final sample was 31 groups, 124 participants (86 females, 37 males and one undisclosed).

Procedure

Thirty-two students fulfilled a course requirement by each recruiting three close friends and bringing the group of four people to the lab at a prearranged time. Recruits were required to be white English speakers who were not studying psychology. This method has been used in several previous studies to form groups of individuals who knew each other well (Emler, 2002; Emler et al., 2007; Robins et al., 2004). Group members were assigned a letter (A, B, C or D) and asked to remember or take note of each other's letter assignment, so that they could refer to each other by letter. Group members filled in the questionnaire either in separate cubicles or in different parts of a large room, without conferring. Each group member first answered questions about themselves and then about each of the other three group members, in a 'round-robin' design (Kenny, 1994).

Materials

Participants completed a questionnaire consisting of measures of contact with Asians (quantity, quality, and direct and indirect cross-group friendship), attitude towards Asians, political views, and extroversion. Demographic questions were presented first with an introductory paragraph.

The order of the contact measures and the attitude measures was counterbalanced (no order effects were detected), and these were preceded by the same note on ethnicity as in the previous studies.

Quantity of contact was measured by two items: 'How often do you meet Asians in your everyday life at university?' and 'How often do you talk with people who are Asians?'. Answers were given on seven-point scales labelled with 'Never' (1), 'Rarely' (2), 'Sometimes' (3), 'Quite a bit' (4), 'Often' (5), 'Very often' (6) and 'All the time' (7). The two items yielded a reliable index ($r = .81$, $N = 124$, $p < .001$) on which higher scores indicated a greater quantity of contact.

Quality of contact was measured with two seven-point scales anchored by the adjectives '*cooperative – competitive*' and '*superficial – intimate*', presented following the stem question 'Are your interactions with Asians generally: . . ?'. With the '*cooperative – competitive*' item reversed, higher scores on the two-item measure ($r = .31$, $N = 124$, $p < .001$) indicated higher quality of contact.

Direct cross-group friendship was measured with two items ($r = .38$, $N = 124$, $p < .001$): 'About how many of your friends are Asian?', with answers given on a seven-point scale labelled with 'None' (1), 'A few' (2), 'Less than half' (3), 'About half' (4), 'More than half' (5), 'Most' (6) and 'All' (7), and 'How often do you spend time with friends who are Asians?' (answers on the same seven-point scale as that

presented with the contact quantity items), with higher scores indicating more cross-group friendship.

Indirect cross-group friendship was measured with two items ($r = .52$, $N = 124$, $p < .001$): ‘About how many of your closest white friends have friends who are Asian?’ and ‘About how many people in your immediate family have friends who are Asian?’. Answers were given on the previously described ‘*None - All*’ seven-point scales, so that higher scores indicated a greater degree of indirect cross-group friendship.

Outgroup attitudes were measured using a ‘feeling thermometer’ and a set of five bi-polar adjectives on seven-point scales preceded by the instruction: ‘Based on your experience rate the extent to which you have each of the following feelings about Asians in general. (Please circle one number on each scale).’ Bi-polar adjectives were ‘*Negative - Positive*’, ‘*Friendly - Hostile*’, ‘*Suspicious - Trusting*’, ‘*Respect - Contempt*’ and ‘*Admiration - Disgust*’. These five adjectives were taken from Wright et al.’s (1997) six-item scale; the sixth item, ‘*Warm - Cold*’, was not included, as it was made redundant by the feeling thermometer. After transforming the feeling thermometer scores to put them on the same scale as the other attitude scores (by dividing all scores by 100 and multiplying by seven) and reversing ‘*Friendly - Hostile*’, ‘*Respect - Contempt*’ and ‘*Admiration - Disgust*’ items, higher scores on the seven-item measure (Cronbach’s $\alpha = .90$) indicated more positive attitudes towards Asians.

Political views were measured with a single seven-point scale from ‘*Very Liberal / Leftwing*’ (1) to ‘*Very Conservative / Rightwing*’ (7). Extroversion was measured using the four-item extroversion measure (seven-point bi-polar scales anchored at ‘*Restrained- Candid*’, ‘*Active- Passive*’, ‘*Silent- Talkative*’, ‘*Aloof-*

Gregarious') from the NEO-Five Factor Inventory (Costa & McCrae, 1988) as well as an additional single seven-point scale anchored at '*Extroverted*' and '*Introverted*'. After reversing the '*Extroverted - Introverted*' and '*Active - Passive*' measures, the composite scale yielded a reliable index ($\alpha = .83$) on which higher scores reflected higher levels of extroversion.

After these 'self-ratings', participants were presented with the following paragraph: 'Now we would like to know about your friends. The rest of this booklet contains questions similar to those you just answered, but they will relate to the people who accompanied you to the lab today -- one set of questions for each person.' Participants were then given the 'self-rating' questions again, three times, but this time the questions were reworded so that they asked the participant to rate each of the other group members in turn. Participant A received questions about B then C then D, Participant B received questions about A then C then D, and so forth. As the letters were assigned in an arbitrary fashion, this ensured that group members did not rate each other in any meaningful order, for instance rating their closest friend first. The 'other rating' questions included measures of how close a friendship the participant had with each other person being rated (assessed using seven circle-pairs, as in Aron, Aron, & Smollan's 1992 'inclusion of the other in the self' measure, see Appendix 4), how frequently they spent time with that person (a single seven-point scale from 'Very rarely, e.g., once a month' to 'Very frequently, e.g., several times a day', with a midpoint of 'A fair amount, e.g., twice a week') and how many months the participant had known each person being rated.

Results and Discussion

The first step was to check the quality of the friendships within each group, in order to ensure that observer-ratings were based on intimate knowledge of the person

being rated. Each participant was involved in three friendships, so each group provided 12 ratings of closeness, frequency of spending time with the friend in question, and length of acquaintance. These indications of closeness were taken as more than sufficient for the purposes of this experiment. Mean rating of closeness ($N = 372$) was 4.61 ($SD = 1.61$). Mean frequency of spending time with the friend being rated was 5.45 ($SD = 1.49$), indicating a point halfway between 'twice a week' and 'several times a day'. Friends had been acquainted for a mean of 19.09 months ($SD = 8.29$), with a minimum acquaintance length of four months.

Raters rated themselves and the other targets in their group on the following composite variables: contact quantity, contact quality, direct cross-group friendship, indirect cross-group friendship, outgroup attitude, political views, and extroversion. Table 4.1 presents the mean rating of the self and the mean rating given to others in one's group on each of these variables. Using the group as the unit of analysis, tests were performed to determine whether these mean ratings differed between self and the other targets. As indicated in Table 4.1, on average raters reported that they had more contact with Asians than they reported that other group members had, they reported that they had more cross-group friendships than others, they reported significantly higher indirect friendships for themselves than they attributed to others, they indicated that they were more liberal politically than they saw others to be in their group, and they saw others as significantly more extroverted than they judged themselves to be.

Table 4.1

Mean ratings of self and others (Study 6)

Variable	<i>Mean</i> (self)	<i>Mean</i> (others)	<i>t</i> -test
Quantity of contact	4.38	4.18	2.62*
Quality of contact	5.25	5.18	.95
Direct cross-group friendship	3.15	2.95	3.05*
Indirect cross-group friendship	4.31	3.63	7.25*
Attitudes	5.31	5.36	-1.03
Political views	3.21	3.45	-2.41*
Extroversion	4.75	4.91	-2.45*

* Significant mean difference, $p < .05$

Further analyses of these rating data from each group were conducted using SOREMO (Kenny, 1998; 1994; Kenny & La Voie, 1984; Malloy & Kenny, 1986) to analyze the round-robin data structure. The input for this program consists of matrices of ratings on each variable for each group. These matrices are structured with each row indicating the person in the group who gave the ratings (i.e., the rater, or in SOREMO parlance the actor) and the columns indicating the person who is rated (the target, or again in SOREMO parlance the partner). The diagonal elements in each matrix consist of the rater's self-rating on the variable in question and are analyzed separately from the basic SOREMO analysis. In essence, for each such matrix, SOREMO partitions the variance in the rater by target ratings into three components: the rater effect, the target effect, and the relationship (or rater by target interaction) effect, which also includes the error component. The rater effect estimates the degree to which there are stable individual differences in the ratings given by particular raters across targets. For instance, if one rater consistently judged all targets higher than

another rater, there would be a large rater effect. The target effect estimates the degree to which individual targets are rated similarly by the various raters. For instance, if all raters agreed in rating one target higher than another, there would be a large target effect. And the relationship effect estimates the degree to which a particular target is judged by a particular rater differently than what would be predicted based on that rater's rater effect and that target's target effect.

A SOREMO analysis was conducted for contact quantity, contact quality, direct cross-group friendship, indirect cross-group friendship, outgroup attitude, political views, and extroversion. Table 4.2 presents the relative variance in the matrices of ratings attributable to rater, target, and relationship effects, computed across the 31 groups with usable data. As relative variances, these sum to 1.00 for each rating dimension; thus each value indicates the average percentage of variance in each matrix of ratings attributable to each source. Also indicated is whether each component of variance differs significantly from zero, except in the case of the relationship variance component which cannot be tested since it also includes error variance. These are tested across the 31 groups. All of the ratings that concern the outgroup (contact quantity and quality, direct and indirect cross-group friendship, and outgroup attitude) manifest considerable rater variance, thus suggesting large individual differences in ratings attributable to the rater. This demonstrates the necessity of using more than one observer, to control for such individual response styles. Importantly, however, each of these variables also shows evidence of reliable target variance, meaning that there are consensual differences among the targets in how they are rated, with some in each group consistently seen as higher on these variables than others. In other words, raters manifest some degree of agreement with each other on how they rate others in the group on these outgroup judgments.

Interestingly, the two ratings that did not focus on the outgroup, political views and extroversion, show considerable target variance, indeed more than rater variance. This replicates other work (e.g., Borkenau & Liebler, 1995; Borkenau et al., 2004; Funder, 1999) that has shown considerable consensus in how others are judged on the extroversion factor of the Big Five personality inventory. Political views, too, have been shown to be often intelligibly communicated to others, particularly in college-educated populations (Emler, 2002).

Table 4.2

Relative variance partitioning – SOREMO analysis (Study 6)

Variable	Rater	Target	Relationship
Quantity of contact	.68*	.09*	.23
Quality of contact	.50*	.16*	.34
Direct cross-group friendship	.47*	.22*	.30
Indirect cross-group friendship	.79*	.04*	.17
Attitudes	.61*	.18*	.21
Political views	.13*	.38*	.49
Extroversion	.06*	.57*	.37

* Variance component significantly different from zero, $p < .05$

Given that all participants rated themselves on these variables, in addition to rating the others in their group, SOREMO provides (and tests) the correlations between the self-ratings on each variable and both the rater and target effects. These self-observer correlations are given in Table 4.3. The first column of correlations in this table reports the extent to which raters who judge others as high or low on average on a judgmental dimension also judge themselves as high or low on that dimension. Unsurprisingly, in the case of all judgmental dimensions but political

views, there are significant correlations between self –other ratings (e.g., the higher a rater rates him- or herself, the higher he or she rates others on average). This pattern of results is likely to reflect individual differences in response style, with some participants more likely to use more extreme scores when rating, or tending towards using the middle of the scale. The pattern is also consistent with a projection hypothesis, whereby a perceiver uses him/herself as an anchor, and estimates or judges others with reference to the self. Participants may have exaggerated their close friends' similarity to themselves. Typically this is done unconsciously, leads to widespread bias, and contributes to our understanding of a wide range of phenomena, including self-esteem, social consensus, self-enhancement bias, and self-affirmation (for a review, see Krueger, 2000, 2007).

Table 4.3

Self-rater and self-target correlations (Study 6)

Variable	Rater	Target
Quantity of contact	.76*	.55*
Quality of contact	.78*	.31*
Direct cross-group friendship	.40*	.53*
Indirect cross-group friendship	.75*	-.06
Attitudes	.79*	.28*
Political views	.20	.65*
Extroversion	.63*	.72*

* Correlation significantly different from zero, $p < .05$

In terms of the motivation that guided this research, the most interesting results are the correlations in the second numerical column of Table 4.3, between self-ratings on each dimension and how one is judged by others. Having established that

there exists significant, albeit modest, consensus within the groups in how targets are judged on the outgroup-relevant judgments, the question these correlations address is whether self-ratings on those judgmental dimensions are correlated with these target effects. That is, for instance, is it the case that targets who report that they have more outgroup contact are also seen consensually by their friends as having more outgroup contact? If such correlations exist, then they would suggest that self-ratings are valid, in so far as they agree with how the self is consensually rated by others.

As the correlations in Table 4.3 reveal, there are substantial and significant self – target effect correlations in the case of all of the rating variables, with the exception of the indirect cross-group friendship variable. In others words, if a rater says he or she has more outgroup contact, reports higher quality contact, reports more outgroup friends, or reports more positive outgroup attitudes, then other close friends tend to consensually agree with those self-ratings. On the other hand, there is no evidence that friends agree with self-ratings in terms of indirect cross-group friendship. This is not surprising given that others would have to observe, or at least know about, the contact not of the target but of the friends and family of the target to be able to make an informed judgement concerning the target's indirect cross-group friendship. The correlations also reveal considerable self – other agreement in ratings of political views and extroversion; the more one rates oneself as politically liberal, or as extroverted, the more other close friends tend to consensually agree in these assessments.

Given the fact that a target's self-rating of contact with the outgroup is substantially correlated with the consensual judgment of that target's contact by his or her friends, one might wonder whether this correlation might be produced simply by highly correlated judgments of a target's style of interrelating with others more

generally, not specifically focused on outgroup members. In other words, given the substantial self – target effect correlation in extroversion, might it be the case that the substantial self – target effect correlation in contact is entirely due to general appraisals by others (and by oneself) of extroversion generally, and not by appraisals more specifically of intergroup contact? To examine this question, further partial correlation analyses were conducted, examining the partial correlation between self-ratings and target effects for each outgroup variable, partialling out self-ratings on general extroversion as well as consensual target effects on extroversion (i.e., how observers judge a target's extroversion). In the case of all of the outgroup rating variables, with the exception of the indirect contact variable, the self – target effect correlations remained substantially positive and significant even when controlling for general appraisals of self-judged extroversion and friends' consensual appraisals of target extroversion.⁹ In other words, there is substantial agreement between how an individual judges his or her own contact and interactions with the outgroup and how others consensually see that person's outgroup contact and interactions, even when controlling for consensual agreement in generalized extroversion of the individual (both in his or her self-judgments and in how he or she is seen by others). This lends further credence to the claim that the presence of significant target variance in these outgroup judgments and their significant correlations with self-ratings on these dimensions suggest that such self-ratings show considerable validity. Over and above any generalized appraisals of how outgoing someone is and is seen to be, the more

⁹ The values of the partial correlations ($N = 31$) were .72, .29, .44, -.17, .28, and .67 for contact quantity, contact quality, cross group friendship, indirect friendship, outgroup attitude, and political views respectively (all correlations significantly different from zero, $p < .05$, except in the case of indirect cross-group friendship)

contact someone reports with outgroup members, the more others who are close friends consensually agree in that assessment.

Study 7

The results of Study 6 provide clear support for the use of self-reports in research on intergroup contact. A second study was conducted to replicate and extend Study 6. Study 7 closely replicated Study 6, with three main additions and changes.

The first of these changes sought to obtain further evidence of discriminant validity. In the first study, it could be seen that there was discriminant validity between the Asian contact variable and general extroversion. That is, while others agreed with self-judgments of both contact and extroversion, it was not the case that the agreement on contact was due to agreement on extroversion. Study 7 examined an additional and more refined issue of discriminant validity, namely whether participants' ratings discriminate between two discrete target groups. Specifically, are self-reports and target judgments of contact with Asians and contact with gay men the same? Strong support would be provided for the validity of self-reports of contact if they showed discriminant validity between two such different target groups.

Second, whereas Study 6 had included only measures of outgroup contact at university, Study 7 contained additional measures of contact in participants' hometowns. It was predicted that the projection effect might be higher on ratings of contact at university, where student respondents had greater knowledge of situations of intergroup contact and relevant norms, than on contact in targets' home town.

Third, the order of self- and observer-reports was counterbalanced. In Study 6 participants always completed self-reports first, because it was thought that this would be easier for them. The limitation of this fixed order, however, is that it may have inflated the extent to which participants projected their ratings of others from the self.

The extent to which a person perceives others as similar to themselves increases when a person focuses on themselves during contact with another (Cartwright, 1957), and completing a self-report would arguably cause participants to focus on themselves while completing ratings of others. If Study 7 detects no order effect, and if correlations are still high between self-ratings and own ratings of others, this would be more likely to indicate similarity between rater and target, rather than self-projection.

Finally, some unnecessary demographic questions were also removed, and the 'politics' and 'extroversion' measures, which were no longer necessary as Study 6 had compared self-other agreement on contact measures with these baselines and controls. The observer-rating of indirect cross-group friendship was also removed, as it had proved impossible for participants to do more than guess at. The question concerning how often a pair saw or spent time with each other was replaced with a set of questions regarding the amount of self-disclosure between each pair. It was reasoned that this would constitute a better measure of the quality of the friendship, as previous studies have shown that self-disclosure is central in the development of relational intimacy and depth (Laurenceau, Barrett, & Pietromonaco, 1998).

Hypotheses

1. Replicating Study 6, observers would show consensus in their ratings of the target's contact with the outgroup and these consensual ratings would agree with the target's self-ratings of contact with the outgroup.

2. Also replicating Study 6, observers would show consensus in their ratings of the target's attitudes towards the outgroup and these consensual ratings would agree with the target's self-ratings of attitudes towards the outgroup.

3. Self-ratings of contact with Asians would agree with observer-ratings of contact with Asians and *not* with gay men, and vice versa – that discriminant validity would be obtained between these two groups.

4. The projection effect would be higher on ratings of contact at university than on contact in targets' home town.

5. No effect would be obtained for the order of self- and other-ratings – this would indicate that 'projection effects' are more likely to be due to actual similarity between rater and target.

Method

Participants

One hundred and thirty-six students at a British university (88 females and 48 males; mean age 20 years) took part in the research. They comprised 34 groups of four, group members being well acquainted with each other for at least six months, with a mean acquaintance length of 20 months, $SD = 9.54$).

Procedure

The same procedure as in Study 6 was followed.

Materials

The same questionnaire was used as in Study 6, with some changes and additions. The contact items were included twice, once with reference to contact at university, and once for contact in the participant's hometown. To keep the questionnaire of manageable length, questions about attitudes and contact with Asians included only the feeling thermometer and the measure of direct cross-group friendship, at home and at university. A feeling thermometer was also included for measuring attitudes towards homosexual men, and direct cross-group friendship was measured using the same two questions as for Asians but with the words 'homosexual

men' substituted (at home, $r = .70$, $N = 136$, $p < .001$, at university, $r = .66$, $N = 136$, $p < .001$).

Self-disclosure between each pair was measured with 4 items; 'How often do you talk about how you are feeling to [A/B/C/D]?', 'How often do you discuss intimate or personal issues with [A/B/C/D]?', 'How often does [A/B/C/D] talk to you about how they are feeling?' and 'How often does [A/B/C/D] discuss intimate or personal issues with you?'. Seven-point scales were anchored at '*Never*' and '*All the time*', with '*Sometimes*' at the midpoint. Items were collapsed to form a reliable scale ($\alpha = .94$) on which higher scores indicated greater self-disclosure.

Results and Discussion

No order effects were detected as a result of the counterbalancing of the contact and attitude measures, nor of the self- and other-ratings. The next step was to check the quality of the friendships within each group. Ratings indicated a similar level of closeness to that found in Study 6, and certainly acceptable for the purposes of the study. Mean rating of closeness ($N = 408$) was 4.28, ($SD = 1.53$), and friends had been acquainted for a mean of 20.05 months ($SD = 11.11$). Mean self-disclosure with the friend being rated was 3.40 ($SD = 1.50$), indicating a level of self-disclosure near the midpoint of the scale. Participants felt that they disclosed more to their friends ($M = 3.45$, $SD = 1.58$) than their friends disclosed to them ($M = 3.35$, $SD = 1.56$), $t(407) = 2.16$, $p = .03$, but both scores indicate self-disclosure only occurred 'sometimes'.

The main analysis was started by comparing mean self-ratings with how others are perceived on average. These means and tests of the differences between them are reported in Table 4.4. Unlike Study 6, there were no differences in how respondents judged the amount of contact they themselves had with Asians, either at home or at

university, compared with the amount of contact others were judged to have had with Asians. On the other hand, respondents' self-reported contact with gays, both at home and at university, was substantially lower than the amount of contact that was attributed to others. That is, respondents generally said that others had substantially more contact with gays than they indicated for themselves. At the same time, interestingly, raters rated their own warmth towards gays as significantly higher than the warmth that they attributed to others.

Table 4.4

Mean ratings of self and others (Study 7)

Variable	<i>Mean</i> (self)	<i>Mean</i> (others)	<i>t</i> -test
Home contact with Asians	1.32	1.41	1.18
University contact with Asians	1.60	1.52	1.37
Thermometer towards Asians	78.13	77.96	.20
Home contact with gays	1.15	2.15	-15.06*
University contact with gays	1.53	2.47	-20.77*
Thermometer towards gays	81.74	78.34	3.38*

* Significant mean difference, $p < .05$

The relative variances attributable to raters, targets, and relationship are shown in Table 4.5. As in Study 6, there was considerable rater variance, or individual differences in the way in which people make ratings, particularly for the thermometer ratings. In other words, some raters consistently rated targets higher than other raters and this was more true on the thermometer ratings than on the contact ones. Importantly, however, on all variables there was also significant variance attributable to the target, meaning that there was consensus across raters within groups as to who was higher on a particular variable and who was lower. The target variances for

judgments of Asian contact and liking were roughly similar in magnitude to those found in Study 6, although interestingly there was substantially more target variance in judged contact with Asians at home than in judged contact with Asians at university. Target variance in the case of contact and liking for gays was clearly more substantial than in the case of target variance for Asians. Thus, there was considerable consensus within these groups of four close friends about who has more contact with gays and who feels warmer towards them.

Table 4.5

Relative variance partitioning – SOREMO analysis (Study 7)

Variable	Rater	Target	Relationship
Home contact with Asians	.36*	.19*	.45
University contact with Asians	.59*	.07*	.33
Thermometer towards Asians	.69*	.06*	.25
Home contact with gays	.16*	.50*	.34
University contact with gays	.37*	.24*	.39
Thermometer towards gays	.54*	.17*	.29

* Variance component significantly different from zero, $p < .05$

Given that there was reliable target variance for all of the contact and liking variables, both for Asian and gay targets, then the important question was whether one's own reports of contact and liking were correlated with how others perceived oneself. That is, if an individual sees herself as having high contact with a target group, do others in her group of friends see her as having relatively high contact? The self-correlations, both with the rater effect and the target effect, are reported in Table 4.6.

Table 4.6

Self-rater and self-target correlations (Study 7)

Variable	Rater	Target
Home contact with Asians	.57*	.54*
University contact with Asians	.71*	.28
Thermometer towards Asians	.91*	.46*
Home contact with gays	.35*	.60*
University contact with gays	.67*	.33*
Thermometer towards gays	.81*	.23

* Correlation significantly different from zero, $p < .05$

All of the self-rater correlations in the second column of the table are large and significant. As in Study 6, this is consistent with a large projection and/or scale usage effect: If one judges oneself as high on liking or contact for one of the target groups, then one judges others in one's group of friends as high on liking or contact for that target group. Interestingly, while these self-rater correlations are all substantial, they are clearly higher for the liking and university contact variables than for the home contact variables. Attitudes, being more subjective, are most likely more difficult to rate, leading to increased reliance on such response strategies as self-projection (Hodges, Johnsen & Scott, 2002). The higher self-rater correlation for university contact variables supports my hypothesis, indicating participants may be more inclined to infer from their own lives when rating contact in a situation of which they have personal experience. Of course, it is possible that many friendships are shared, and that the 'projection effect' seen here is simply due to rater and target sharing a number of Asian friends, and therefore having similar levels of cross-group friendship.

The self-target correlations are also positive and most are significant. These indicate that others largely agree with the individual in self-judgments. Interestingly, the most substantial of these correlations are for the home contact variables, both with Asians and with gays. In essence, if an individual reports more contact than others with these groups in his home environment, then others agree with him in this assessment of his degree of contact. The self-target correlations are substantially smaller for the two university contact variables, where it has been seen there is greater projection, i.e., higher self-rater correlations. Finally, the self-target correlation is substantial and significant for the Asian thermometer variable, whereas it is smaller and nonsignificant for the gay thermometer variable; the more we say we feel warmly towards Asians, the more others concur in this judgment, but this is less true in the case of our expressed warmth of feeling towards gays.

Table 4.7

Correlations among target effects and self-ratings: Asian and gay contact and liking (Study 7)

		Target effects				Self-ratings		
		Asian contact	Asian warmth	Gay contact	Gay warmth	Asian contact	Asian warmth	Gay contact
Target effects	Asian contact							
	Asian warmth	.88*						
	Gay contact	-.09	-.27					
	Gay warmth	-.12	.17	.58*				
Self-ratings	Asian contact	.54*	.34*	-.16	.08			
	Asian warmth	.38*	.46*	-.34*	.00	.29*		
	Gay contact	.18	.12	.60*	.24	.17	-.10	
	Gay warmth	.20	.09	-.05	.23	.18	.67*	.14

* Correlation significantly different from zero, $p < .05$

Turning next to the additional and more refined investigation of discriminant validity, it must be considered whether self-reports and target judgments of contact with one target group (e.g., Asians) are the same as those for another, quite different outgroup (e.g., gays). In Table 4.7 are presented the correlations among the target effects and self-ratings for four different variables: Home contact with Asians, Asian thermometer, home contact with gays, and gay thermometer.¹⁰ The target effects represent how one is judged consensually by others in one's group. The matrix of correlations is divided into three sections. Correlations that lie above the boxed set are correlations of target effects with each other. Correlations to the right of the boxed set are correlations of self-ratings with each other. Correlations inside the box are correlations of target effects with self-ratings.

Looking first at the correlations between the target effects, it is clear that if one is consensually judged as having high contact with a group, then one is consensually judged to have positive attitudes towards that group ($r = .88$ for Asians; $r = .58$ for gays). On the other hand, all other correlations among these target effects are nonsignificant, and three of the four of them are negative. In other words, individuals who are consensually seen as having high contact and positive attitudes towards one of the two groups are *not* seen as having high contact and positive attitudes towards the other group.

Turning to the correlations among the self-judgments, the correlation between reported contact towards the two groups is positive, but not significant. On the other

¹⁰ Only the "at home" contact variables were included in this table because the "at university" contact variables showed less target variance and lower self-target correlations. Additionally, including both self and target variables for each of these, would have more than doubled the size of the resulting matrix (66 total correlations versus the current 28 reported in the matrix). The full matrix is reproduced in Appendix 1.

hand, the correlation between the two thermometer measures is very large and highly significant ($r = .67$). Thus, liking towards the two target groups is redundant in the self-ratings, in a way that they are not in the target effects. This difference is surely in part due to the large rater effects on these thermometer measures – if one gives high thermometer ratings for oneself towards one target group, one gives high thermometer ratings for oneself for other target groups and also for others in general.

Taking the correlations between the target effects, on the one hand, and the self-ratings on the other (i.e., those in the box), the correlations on the diagonal are those reported earlier – are self-judgments on a variable congruent with how one is consensually judged by others on that variable? Importantly, these self-target correlations are all larger than the off-diagonal correlations. For instance, others consensually agree with one in one's self-rating of contact with Asians ($r = .54$) and they consensually agree with one in one's self-rating of contact with gays ($r = .60$). But when the target groups are mismatched between self-ratings and target effects, the correlations are not significant (and are negative in one of the two cases). That is, one's self-rating of contact with Asians is correlated $-.16$ with one's target effect on contact with gays. And one's self-rating of contact with gays is correlated $.18$ with one's target effect on contact with Asians. So the self-ratings agree with the target effects only when the same target group is referenced.

General Discussion

These two studies constitute an original contribution to the research literature on intergroup contact by demonstrating for the first time the validity of self-report measures of intergroup contact based on consensually supported observer-reports. Using a round-robin design, it was possible to show significant agreement between participants' self-reports and observers' ratings, and to demonstrate the discriminant

validity of the findings. In providing evidence for the validity of self-reports of contact, these studies remove a source of criticism of many of the studies comprising the large data base of research on intergroup contact. I discuss the conclusions and implications of this research in terms of the utility and limitations of self-reports of contact, and discriminant validity.

Utility and limitations of self-reports of contact

As has been done previously in personality research (e.g., for measures of extroversion; see Borkenau & Liebler, 1995; Borkenau et al., 2004; Funder, 1999) and for reputational measurement of political affiliation (Emler, 2002; Emler et al., 2007), it was found in both studies that observers (or “partners”) made consensual judgments of targets’ (actors’) degree and quality of intergroup contact, and that these consensual judgments were correlated with the targets’ own self-ratings. Study 6 also replicated the results of previous research using measures of extroversion and political affiliation; as in previous work, extroversion and political views appeared to be the most visible, easily-rated attributes, as they manifested the greatest target variance, the smallest self-projection, and the highest self-target correlations.

Compelling evidence was also reported of reliable target variance, meaning that there were consensual differences among the targets in how they were rated by observers. Thus raters manifest some degree of agreement with each other on how they rate others in the group. As expected, both political views and extroversion showed considerable target group variance, but target effects were significant, albeit smaller, for all measures of contact, especially cross-group friendship and contact quality, and for outgroup attitude.

Substantial self – target effect correlations substantiated these findings, evident on all rating variables except indirect friendship. Thus, when a rater reported

that he or she had more outgroup contact, higher quality contact, more outgroup friends, or more positive outgroup attitudes, then other close friends tended to consensually agree with those self-ratings. But, not surprisingly, the self-target effect correlation was not significant for indirect contact (Study 6); others would have to observe, or at least have knowledge of, the contact not of the target but of the friends and family of the target to be in receipt of this information.

Notwithstanding the general support provided here for self-reports of intergroup contact, there remain two potential limitations with this method of assessing contact, namely socially desirable responding, and self-projection. However, only partial evidence of socially desirable response bias was found. In Study 6, raters reported having significantly more contact with Asians, and more direct and indirect cross-group friendship than the others in their group. In Study 7 respondents reported more positive attitudes towards gay men, but actually reported lower contact with gays than they thought their friends had. These results do not suggest a strong influence of social desirability; nonetheless given that people have a tendency to rate themselves more favourably than they rate others (Dunning, 2005), this remains a threat to the validity of self-reports of intergroup contact. Evidence indicates moreover that accuracy tends to be greater when rating others than when rating oneself (Dunning, 2005; Epley & Dunning, 2000), and this may be due to an over-reliance on case-based information (to the detriment of distributional information) when rating the self, which is much less pronounced when rating others (Buehler, Griffin & Ross, 1994; Dunning, 2005). People also tend to look at others more objectively, and at themselves more optimistically (Epley & Dunning, 2004). Future research might try to demonstrate whether participants actually under-estimate others' contact rather than (or as well as) over-estimating their own.

Turning to projection, it was found that participants did indeed rate others as similar to themselves to a large extent. In Study 6, for all judgmental dimensions but political views, there were significant correlations between self–other ratings. It could be argued that this self-projection weakens observer-reports as a method for evaluating self-reports. However, as SOREMO calculates the variance caused solely by the target and unaffected by the rater, and as this variance was found to be significantly different from zero, I argue that self-projection effects did not weaken this research. In any case, self-projection may have been overestimated, as people tend to be friends with similar others (Daniels & Plomin, 1985; Kiesner, Maass, Cadinu & Vallese, 2003; Rowe & Osgood, 1984; Rushton, 1989), and ratings of others are therefore necessarily correlated with self-ratings. To reduce any impact of self-anchoring the order of self- and other-ratings was counterbalanced in Study 7, but no effect of this order variable was found. Self-projection occurs most frequently when a person focuses on themselves during contact with another (Cartwright, 1957), and answering questions about oneself would be likely to increase this effect (Schwarz & Sudman, 1992). I therefore suggest that this similarity between ratings was due more to actual similarity between the friends, or to scale usage styles, than to self-projection. This theory is supported by the fact that, in Study 7, it was predicted, and found, that the projection effect would be higher on ratings of contact at university. Close friends are likely to share many of the same friends, thereby increasing the similarity between their cross-group friendship.

Discriminant validity

Two types of evidence are also presented for discriminant validity. First, in Study 6, self-reports and observer-reports of extroversion were measured, demonstrating that when observers judged a target's intergroup contact they were not

simply judging how outgoing the target was, a judgment that past research has suggested can be made consensually and accurately. Of course, it is plausible that more extroverted people (people who like being around others in general) would be judged higher on ratings of contact with outgroup members (in particular). It proved possible to exclude this alternative explanation for the agreement between targets' self-ratings of contact and consensual judgment by their friends in terms of targets' levels of extroversion. Correlations between self-ratings and target effects remained reliable when partialling out self-ratings on general extroversion as well as consensual target effects on extroversion.

Secondly, in Study 7, an additional and more refined issue of discriminant validity was examined, namely whether participants' ratings discriminate between two discrete target groups, Asians and gays. Specifically, are self-reports and target judgments of contact with Asians and contact with gays the same? A comparison was performed of the correlations among the target effects and self-ratings for four different variables, based on contact and attitude measures for the Asian target group, and contact and attitude measures for the gay target group. As should be the case, if measures have discriminant validity, contact and attitude measures were correlated within, but not between, target groups. Thus, if an individual is seen consensually as having high contact with and positive attitudes towards one of the two groups, this does *not* mean that he or she is seen as having high contact with and positive attitudes towards the other group. Further evidence of discriminant validity comes from the correlations between target effects and self-ratings. These correlations assess, for example, the congruence between self-judgments on a variable and how one is consensually judged by others on that variable (e.g., if I report high contact with Asians, do others also report that I have high contact with Asians?). These

correlations are significant only when there is a match between the target group for the two ratings, but not when there is a mismatch. For example, self-rated contact with Asians [gays] is *not* significantly correlated with own target effect on contact with gays [Asians].

To conclude, the results of these two studies confirm this innovative attempt to validate self-reports of intergroup contact by observer-reports. Consistent results have been reported across two studies, and two target groups, showing that observer-reports validate self-reports. This research also yielded two types of evidence for discriminant validity, showing that consensus across observers about targets is not simply based on judged extroversion, and that measures of contact and attitude are only meaningfully related within, but not between, target groups. Overall, these findings increase the confidence with which conclusions can be drawn from the large literature on intergroup contact that relies on self-report measures.

CHAPTER FIVE:

Validation of contact self-reports using online communities

While observer-ratings are widely recognised as an appropriate criterion against which to judge the validity of self-ratings, they are by no means without flaw. It is already known that self-reports are prone to bias, but observer-reports have shown bias also (Hoyt & Kerns, 1999). It is likely, for example, that observer-reports are based on limited knowledge, or knowledge that has been gained simply from listening to what the target has told the observer about themselves - a source which is just as prone to bias as a self-report on a questionnaire. Consensus between observers should support the value of observer-reports as a means of validating self-reports (McCrae & Weiss, 2007), but consensus might also result from targets presenting a similar 'front' to all three observers. Consensus may also reflect similarities existing between friends (Daniels & Plomin, 1985; Kiesner et al., 2003; Rowe & Osgood, 1984; Rushton, 1989), as friends are likely to share group norms concerning what constitutes a high or low level of contact. Consensus may even be to some degree artifactual, the correlations inflated due to a restricted range imposed by friends rating themselves and each other based on the relatively small variation about this group norm.

There is also a chance that both friends and family members, feeling loyalty for the target, will over-report socially desirable characteristics and under-report undesirable ones in the same way that the targets themselves might, although to a lesser degree (Krahé, 2008). Alternatively, observers may act like 'cognitive misers' (Taylor & Fiske, 1978), considering only the most recent or salient occasion which

provides information relevant to the task of rating a target's level of contact, and ignoring all other aspects and occasions.

Triangulation is the use of multiple cross-checked sources and methodologies. Common method variance (see Chapter Two) tends to exaggerate the relationship between two constructs measured with the same method (Wiggins, 1973). Including multiple types of data in the research design reduces this exaggeration and allows weaknesses associated with each type of data to be controlled for (Block & Block, 1980). McCrae and Weiss (2007) advocate multimethod approaches to personality assessment because no method, including observer ratings, is perfect. Even assuming that observer-reports are an ideal criterion for validation of self-reports of contact, it remains the case that in order to test this validity thoroughly, self-reports must be validated by a variety of methods, not simply this one external criterion.

In order to demonstrate *criterion-related validity*, an instrument must predict some criterion external to the instrument itself but which can be expected to arise from having a certain score on the test (Nunnally, 1978). In other words, self-reports of contact should predict actual contact. As comprehensively observing even one individual's actual contact over the course of a representative period of time would be unfeasible, an adequate alternative seems to be the analysis of mutually-nominated cross-group friendships (e.g., Lee, Howes & Chamberlain, 2007). This method requires every friend nominated by every participant to be asked for their own nominations, in order to obtain the knowledge that the friendship is indeed mutual rather than one-sided, or simply a desire for friendship. Of course, this would be very impractical if boundaries were not placed upon the potential nominations. This has traditionally been achieved by examining only friendships existing in a closed environment such as a workplace or school. Lee et al. (2007), for example, explored

cross-group friendship in school children by asking 350 children to nominate their friends, yielding 956 mutually-nominated friendships. Among other findings, they discovered that girls were more inclined towards cross-race friendships than were boys. While it may be reasonable to assume that the friendships within a classroom may comprise most of the total friendships of a child, this is unlikely to be true of adults. The large amount of contact and friendship potential existing outside of the relatively small closed environment under study limits the capacity of the research to form a representative picture of outgroup contact in a group of individuals. Fortunately, Oxford University constitutes a very large closed environment, within which there is an unusually high tendency for real, mutually-nominated, friendships to be reflected in an online community; 'Facebook'.

The online community

Facebook is an online network used by the majority of Oxford University students (around 40,000 students, staff and alumni; Facebook Network Information, 2009). Every Oxford student has access to the internet, most commonly in their room in student accommodation, or at the very least through libraries and computer rooms, and as such is able to use and interact on the Facebook network whenever they wish. Facebook gives every member a page where they describe themselves: their name, their photograph, their educational background, their hometown, and any interests, courses, groups or characteristics they wish to share with others on the network. Most importantly, Facebook allows a person to 'add' other members as 'Friends'¹¹, thereby adding a link to their Friend's Facebook page on their own page. In this way, all a person's Friends are named and pictured for any member of the Oxford Facebook

¹¹ Note that I refer to a person added on a Facebook Friends page via mutual nomination as a 'Friend', while 'friend' refers to the conventional use of the word.

network to view. Importantly, a 'Facebook friendship' is, by its nature, mutual, as both members must agree that they are friends before Facebook will record the friendship. One friend nominates the other, who is then automatically sent a message with two links, one to confirm and one to deny the friendship. Mutually-nominated direct cross-group friendship can therefore be measured by counting the number of people on a participant's Facebook Friends Page whose picture or name shows them to be Asian (see Appendix 2 for example/screenprint). Another important feature of Facebook is that members have a 'Profile' page which includes a place (called the 'Wall') where Friends can leave messages or comments about the member's news, photographs, items (such as web journals or general writings) they have posted on their pages, or simply to chat. These messages are visible to anyone in the Oxford network, and as such can be used for collecting data about online contact. The Friend's name and photograph are included with each message or comment, allowing information to be collected concerning their ethnicity, gender, and approximate age.

It could be argued that online friends are not the same as real friends, and that friendship with a person of another race online does not fulfil the essential component for contact as described by Allport (1954), as the contact is not face to face. This would only be the case for very few friends in this instance, however, as Facebook is *not* an online network where it is common to meet people online and converse with them but never meet in person. Rather, Facebook is a method of sending messages to all one's 'real world' friends, or communicating with them in an informal fashion, reading their news, or exchanging digital photographs. Most Facebook members' 'Friends Pages' are composed almost exclusively of people whom they know from their day-to-day lives, rather than people they have met online. Indeed, when a Friend is being 'added', Facebook requests an explanation of how the user met the person

concerned. Options include having lived, worked or travelled together, taken an academic course together, 'hooked up' at a party, dated, met through a mutual friend, taken part in extra-curricular activities together and so forth. Only one of the 14 options states 'met through Facebook'.

Online contact

Even so, as it is possible some 'Facebook Friendships' may be purely online, the issue of online contact must be addressed. Given the direction of the society in which we live, it is arguable that to dismiss online friendships would be unwise (Amichai-Hamburger & McKenna, 2006). The Internet has become a significant field for social interactions (Kraut, Patterson, Lundmark, Kiesler, Mukopadhyay, & Scherlis, 1998). Indeed, intimate friendships can form more easily and quickly online than offline (McKenna, Green, & Gleason, 2002; Walther, 1996) due to factors such as increased control over how one presents oneself, and the absence of inhibitions caused by anxiety over one's personal appearance or habits (Ben Ze'ev, 2004; McKenna & Bargh, 2000; McKenna & Green, 2002; McKenna et al., 2002). Online communication is associated with increased self-disclosure, particularly among purely 'online' friends who have never met physically, but also among 'real-world' friends during their interactions online (McKenna, Buffardi, & Seidman, 2005), due to increased feelings of anonymity and decreased vulnerability (Cooper & Sportolari, 1997). This increased self-disclosure and decreased vulnerability can create a feeling of greater closeness, or 'hyper-intimacy' (Bonebrake, 2002) and a hyperpersonal effect (Walther, 1996) meaning that the contact feels very personal in nature. Meeting a person online has, surprisingly, been found to lead to greater liking and a greater feeling of having 'got to know' the person than meeting that person in a face-to-face encounter (McKenna et al., 2002, Study 3). Rather than being superficial, therefore, it

can be seen that online friendships can, in fact, be closer than relationships devoid of any online communication, and thus cannot be dismissed.

Internet-mediated contact in general may, in fact, be more valuable than face-to-face contact (Amichai-Hamburger & McKenna, 2006). Anxiety during a face-to-face intergroup encounter can cause the individuals involved to ignore stereotype-disconfirming information about the outgroup (Wilder, 1993) and apply heuristics such as stereotypes in their evaluation of outgroup members (Bodenhausen, 1990; Bodenhausen & Wyer, 1985). Participants tend to feel greatly decreased anxiety during electronic encounters (McKenna et al., 2005) as they have the ability to withdraw completely and immediately from the contact whenever they choose; participating in internet-mediated contact is seen as less 'risky' than a face-to-face encounter (Bargh & McKenna, 2004; McKenna & Bargh, 1999). Optimal conditions for contact such as equal status are far easier to effect over the internet (Spears, Postmes, Lea, & Wolbert, 2002), and practical issues surrounding bringing together groups separated geographically are removed.

By allowing participants to engage in contact from the privacy of their own homes, internet-mediated contact reduces the negative effects of public self-consciousness which can otherwise lead to a narrowing of cognitive and perceptual focus, and increased reliance on stereotypes for person evaluation (Lambert, Payne, Jacoby, Shaffer, Chasteen & Khan, 2003). An individual is more likely to be open to altering their dominant responses (Zajonc, 1965) in a private setting than in a public one, and although a 'chat room' is a public electronic venue, participants tend to feel that the setting is private because they can take part in their own home (McKenna & Bargh, 2000; McKenna et al, 2002). Beneficial effects of contact are also more likely to generalize to the outgroup as a whole when that contact is anonymous, as category

salience is increased (Spears et al., 2002). It can be argued therefore that measuring online contact with real-world friends (and potentially some purely online friends) would not only be a valuable secondary criterion for assessing the validity of self-reports of contact, but may well provide useful information in its own right, as a measure of online contact, a variety of contact which may well prove to be immensely valuable in intergroup conflict resolution in the near future.

Two studies examined instances of online contact with Asians (Study 8) and number of Asian Friends in the online community (Study 9), and compared these data with self-reports of contact with Asians. Study 8 also included observer-reports, to ascertain whether online contact was itself valid, when compared to consensual observer-reports, arguably the best standard against which to measure validity (Howard, 1994).

Study 8

This study was designed to ascertain whether actual online contact bore any relation to self-reported contact, and more importantly to consensually-rated observer-reported contact (as the latter might be seen as the more reliable measure). Participants who had taken part in Study 7 were therefore asked for permission to view their Facebook Profiles¹². The self-reports of contact that these participants had already provided, and the observer-reports provided by their three friends, were compared with actual contact taking place online, six weeks after the round-robin

¹² This was to ensure ethical guidelines were followed, although I argue that viewing Facebook Profiles does not constitute an invasion of privacy, as any message left on the 'Wall' (public comment space) is a public message, intended to be viewed by anyone. There is also a private messaging function which members can use if they do not want their comments to be visible to the public. Members also have the option to make their Facebook Profile private so that only their friends can see it. If they choose not to do so, they are opting to allow anyone within the Oxford network to view their Profile. None of the participants of Study 8 had opted to make their Profiles private.

study had taken place. In this way, self- and observer-ratings were examined for evidence of predictive criterion-related validity, the criterion in question being actual experiences of online contact six weeks after self- and observer-report data were gathered.

Instances of actual online contact included having an Asian person leave a comment on a participant's Facebook Profile, or comment on a participant's photographs, or interact with the participant in any other way through Facebook. Private messages could not, of course, be taken into consideration. Only recent contact was accounted for, as the 'Wall' only has space for a limited number of messages; more recent messages push older ones off the 'Wall'.

Hypotheses

1. Those participants reporting higher levels of contact would be more likely to have experienced recent online contact with Asians six weeks later, than those reporting lower levels of contact.

2. Participants whose observer-reports of contact were higher would be more likely to display evidence of recent online contact with Asians six weeks later, than those whose observer-reports of contact were low. Observers are less likely to know about contact that happens online, but as online friends in this case are also likely to be real-world friends, observer-ratings of contact were expected to predict instances of online contact with Asian Friends.

3. Those participants who had reported more positive attitudes towards Asians would be more likely, six weeks later, to have experienced recent online contact with Asians.

Method

Participants

Fifty individuals who had taken part in Study 7 consented to take part in the current study and to have their Facebook Profiles examined. Forty-one of them were female and nine male, and they were aged between 19 and 35 (mean age = 20.48). All were students at Oxford University, and members of Facebook. Importantly, non-members of Facebook were not excluded from the study; the fact that all participants were members of Facebook highlights the overwhelming tendency of Oxford undergraduate students to choose to belong to this online community.

Procedure

Six weeks after Study 7 took place, Facebook Profiles of those participants who had given their consent were examined and participants were divided into two groups based on whether their Facebook Profile showed them to have any recent contact with Asians on Facebook. Those whose pages provided evidence of online communication with someone whose name or photograph showed them to be Asian were grouped as 'with recent contact' and those with no comments written by Asians as 'without recent contact'. Self-report measures of quantity and quality of contact, as well as direct and indirect cross-group friendship, and attitudes to Asians, were taken from the results of Study 7, and were matched to individuals using a Research Participation Scheme code given to each student for the duration of their studies at Oxford, and which they had reported in the round-robin study in which they had taken part. The only observer-report measures which had been assessed in Study 7 were direct cross-group friendship, and the feeling thermometer. Only these were therefore included in the current study.

Materials

The only new measure was the categorical variable ‘Recent instances of online contact’, with two levels: ‘*contact*’ or ‘*no contact*’. The number of instances of online contact was calculated, but this variable was extremely skewed, due to the majority of participants having no instances of contact [$N(0) = 30$, $N(1) = 16$, $N(2) = 3$, $N(3) = 0$, $N(4) = 1$]. No transformation was sufficient to normalize this distribution, so a dichotomous variable was created. Although crude, this measure was expected to suffice to determine whether actual contact online bears any relation to reported contact. Measures taken from the two round-robin studies were self-reports of quantity and quality of contact with Asians, direct cross-group friendship at university, indirect cross-group friendship at university, extroversion, attitude, and the feeling thermometer¹³ (see Chapter Four). Observer-reports were also used of direct cross-group friendship at university and the feeling thermometer. To create the observer-reports, each participant’s three friends’ observer-ratings were averaged. Emler et al. (2007) used this method, and reported the consensus between the observers also. Following their procedure, consensus was calculated for direct cross-group friendship and the feeling thermometer. For each measure, a correlation was produced for each of the three observers, and this was averaged to produce a measure of the consensus existing between observers regarding the construct of interest. Observers agreed with each other to a significant degree concerning the participant’s level of direct cross-group friendship, $r = .27$, $N = 128$, $p = .02$, but did not agree with each other significantly regarding the participant’s attitudes as measured by the feeling thermometer, $r = .07$, $N = 128$, ns.

¹³ This was kept separate from the attitude measure for this study, in order to facilitate comparison with the observer-report.

Results and Discussion

A one-way between-groups analysis of variance revealed that self-reported quality and quantity of contact, as well as direct and indirect cross-group friendship, were all significantly higher for those who had recent online contact with Asians than for those who did not. These results are displayed in Table 5.1. Self-reported attitudes, however, did not differ between those with and those without recent online contact. This is not unexpected, as the measure did not take into account offline contact. Observer-reports, assumed to be a valid measure of a participant's true level of contact, are also displayed in Table 5.1. Direct cross-group friendship, as rated consensually by three of the participants' good friends (see Chapter Four), was significantly higher for those with recent online contact than for those without. Observer-reported attitudes as measured by the feeling thermometer were also significantly higher for those who had had recent online contact with Asians than for those who had not, although it must be noted that these observer-ratings were not consensual, and may not validly reflect the targets attitudes.

The first hypothesis was supported in that all self-report contact measures differed significantly depending on whether or not the participant had engaged in recent online contact. The second hypothesis was also supported, as both observer-measures also differed significantly depending on whether online contact existed or not. However, the third hypothesis, that those with a recent instance of online contact would show higher self-reported attitudes, was not supported. As observer-reports had failed to consensually validate self-reported attitudes, this finding may be explained by a lack of validity in attitude self-reports. An implicit measure of attitudes would be a useful addition to the study (see Study 9).

Table 5.1

F tests of self- and observer-reported outgroup contact and attitudes between those with recent online contact and those without (Study 8)

	<i>Mean</i> (no recent online Contact) <i>N</i> = 30	<i>Mean</i> (recent online contact) <i>N</i> = 20	<i>F</i> (1,48)
Self-report of quantity of contact	2.69 (1.40)	4.66 (1.70)	20.05****
Self-report of quality of contact	2.41 (0.68)	3.10 (0.61)	13.42****
Self-report of direct cross-group friendship	1.85 (1.17)	3.52 (1.18)	24.13****
Self-report of indirect cross-group friendship	2.83 (1.67)	4.30 (1.36)	10.69***
Self-report of feeling thermometer	79.00 (16.47)	73.00 (15.93)	1.63
Self-report of attitudes	5.47 (0.95)	5.42 (0.90)	0.03
Observer-report of direct cross-group friendship	1.88 (0.68)	3.28 (0.69)	50.08****
Observer-report of feeling thermometer	76.67 (6.55)	83.83 (9.07)	10.54***

Note: Standard deviations provided in parentheses.

**** $p < .001$

*** $p < .005$

As observer-reports are well recognised as an appropriate criterion for validating other measures, and as all observer-measures also differed significantly depending on whether online contact existed or not, the use of Facebook was concluded to be a valid criterion also for validation of self-reports. Facebook contact succeeded in validating self-reports of contact in this study. However, it is likely that most contact examined by this study took place between close friends, as close friends can be expected to communicate more frequently on Facebook than do mere acquaintances. It is therefore possible that the contact measured in this study reflects *close* cross-group friendship, and does not adequately reflect participants' cross-group friendship in general. Study 9 therefore examined all of each participant's mutually-nominated friendships, and determined how many of those friendships were cross-group.

Another potential weakness associated with the dependent variable used in this study emerged during data collection. Those participants who have more instances of online contact with Friends in general (perhaps due to higher levels of extroversion, or simply by being more frequent users of Facebook) may be incorrectly classed as having no recent online Asian contact. As only the most recent instances are available for viewing on the 'wall', an instance of contact with an Asian Friend would have to be more recent for those who have more instances of contact with Friends in general, than for those who have fewer instances of contact in general. To attempt to assess whether this issue affected the data, a one-way analysis of covariance was conducted for all the self- and observer-report variables, controlling for extroversion ($M = 4.90$, $SD = .82$). The results were unchanged by the step of controlling for extroversion, suggesting that extroversion did not influence the independent variable by increasing the number of *non*-Asian instances of contact. Interestingly, however, extroversion

has been found to *decrease* use of the internet for social interaction (Hamburger & Ben-Artzi, 2000), although this pattern was only found for women. Clearly, controlling for extroversion cannot determine whether this variable fairly assesses whether a person has experienced a recent instance of online intergroup contact. The frequency of participants' Facebook use is more likely than extroversion to have affected these results, and as this cannot be assessed post-hoc, this dependent variable remains potentially problematic. The alternative dependent variable used in Study 9 did not share this flaw, and therefore corrected for this limitation of Study 8.

Study 9

Predictive criterion-related validity of self-and observer-reported contact having been demonstrated, Study 9 sought to demonstrate *concurrent* criterion-related validity. Data concerning self-reports of direct cross-group friendship were gathered concurrently with data concerning actual cross-group friendship as displayed on 'Facebook Friends Pages'. The ethnicity of each of a participant's Facebook Friends can usually be determined by examining the Friend's name and photograph. It was expected that the *total* number of Friends on a person's Facebook page would increase the number of *Asian* Friends on the page simply by virtue of the proportion remaining constant, assuming that Asian friendship is not being specifically avoided. Having more Friends in general cannot be expected to increase a person's tendency to *prefer* Asian Friends, but simply to increase the opportunity for outgroup friendships. For this reason, it is necessary to control for total number of Facebook Friends when testing relationships between Asian Friends and other variables.

It was expected that, as people vary in their levels of extroversion (Costa & McCrae, 1988), they would also vary in their number of Facebook Friends. As I have reported, the way in which extroversion influences the use of the internet for social

interaction has been shown to be different for males and females. While extroverted women are *less* likely to use the internet for social interaction, extroverted men do not show the same pattern (Hamburger & Ben-Artzi, 2000). However, as Facebook Friendships are primarily also real-world friendships, extroversion was expected to be positively related to Facebook Friends.

Observer-reports of contact were not included in this study, as the purpose of their inclusion in Study 8 had already been achieved, namely that online contact as measured using Facebook had been shown to agree with consensually-rated observer-reports of direct cross-group friendship. Instead, an implicit measure of attitudes was included, to ascertain whether the proposed unobtrusive and explicit measures of contact could predict scores on the implicit and explicit measures of attitudes. Explicit measures of contact have consistently been shown to predict explicit measures of attitudes (for a review, see Pettigrew & Tropp, 2006), and, to a lesser extent, to predict implicit measures of attitudes (Brown & Hewstone, 2005). However, no work has yet sought to discover whether unobtrusive contact predicts implicit attitudes, as no suitable unobtrusive measure of contact has been found. I argue that mutually-nominated online friendships, as reported on the Facebook, provide such an unobtrusive measure of contact, as the examination of pre-existing information does not involve participants consciously reporting on their cross-group friendship. Petty, Fazio, and Briñol (2009) define an implicit measure as one which measures without making individuals aware of what the measure is assessing. Although Facebook Friendships cannot constitute a truly ‘implicit’ measure, this definition supports their use as an unobtrusive measure of contact.

Implicit Association Test

As already described (see Chapter One), the Implicit Association Test (IAT, Greenwald et al., 1998) assesses the strength of an association between an attitude object and an evaluation. Participants are presented with attitude objects of interest, such as Asian and white names or faces, and asked to press a certain key as fast as possible when they see one category, and another key for the other category. They must then repeat this categorising task with positive and negative evaluations (or stereotypical and non-stereotypical, if that is the issue being explored). Finally, participants are instructed to categorise both groups of items at once, in two blocks of trials. One block defines one response key as the correct category response for positive OR white (for example) and the other as negative OR Asian (i.e., a block of compatible trials), and one block defines one response key as the correct category response for positive OR Asian (for example) and the other as negative OR white (i.e., a block of incompatible trials). The higher the implicit prejudice of the participant, the more difficult they will find it to associate the target group with positive and the ingroup with negative evaluations, or in other words to make one response for white and negative, and the other response for Asian and positive. The difference in response times for compatible versus incompatible trials is therefore taken to represent the level of implicit prejudice of the participant (Greenwald & Banaji, 1995).

Despite the transparency of this technique, it reliably shows a tendency among the majority ethnic group towards higher response times for incompatible trials (McConnell & Leibold, 2001). In other words, even if a participant is aware that his or her attitudes are being assessed, he or she cannot 'fake good'.

Hypotheses

1. Self-reports of direct cross-group friendship would predict number of Asian Friends on Facebook while controlling for total number of Facebook Friends.
2. Self-reports of attitudes would not agree with scores on the IAT, as socially sensitive issues tend to produce greater disparities between implicit and explicit attitudes (Greenwald et al., 1998).
3. Self-reports of direct cross-group friendship would predict self-reports of attitudes.
4. Number of Asian Facebook Friends (while controlling for total number of Facebook Friends) would predict implicit attitudes, as measured using the IAT.

Method

Participants

Nineteen males and 44 females aged between 18 and 32 (mean age = 19.46) consented to take part in the study and to have their Facebook Profiles examined. Again, all 63 participants were students at Oxford University, and were members of Facebook (although, again, not chosen specifically because of this).

Procedure

Facebook Profiles of participants were examined and data gathered concerning two variables: 'total Friends on Facebook' (the total number of people on the participant's 'Friends Page') and 'Asian Friends on Facebook' (the total number of Asian people on the participant's 'Friends Page'). Participants also completed the IAT, and a questionnaire consisting of measures of contact with Asians (quantity of contact, and direct and indirect cross-group friendship), attitude towards Asians, and extroversion. Demographic questions were presented first with an introductory

paragraph. The order in which participants completed the questionnaire and the IAT was counterbalanced (no order effects were detected). Facebook data were gathered while participants completed the IAT, so that concurrent measures could be taken. This eliminated any effects of measurement delay, such as the possibility of changes in Facebook Friends Pages in the time intervening between participants completing the questionnaire and IAT, and their Friends page being examined.

Materials

Extroversion was then measured using the same items as in Study 3. After reversing the '*Extroverted - Introverted*' and '*Active - Passive*' measures, the composite scale yielded a reliable index ($\alpha = .77$), on which higher scores reflected higher levels of extroversion.

The order of the contact measures and the attitude measures was counterbalanced (no order effects were detected), and these were preceded by the usual note regarding ethnicity.

Quantity of contact was measured by two items ($r = .94$, $N = 63$, $p < .001$): 'How often do you meet Asians in your everyday life at university?' and 'How often do you *talk with* people who are Asians?' Answers were given on seven-point scales labelled with '*Never*' (1), '*Rarely*' (2), '*Sometimes*' (3), '*Quite a bit*' (4), '*Often*' (5), '*Very often*' (6) and '*All the time*' (7).

Direct cross-group friendship was measured with two items ($r = .63$, $N = 63$, $p < .001$): 'About how many of your *friends at university* are Asian?', with answers given on a seven-point scale labelled with '*None*' (1), '*A few*' (2), '*Less than half*' (3), '*About half*' (4), '*More than half*' (5), '*Most*' (6) and '*All*' (7), and 'How often do you *spend time* with friends at university who are Asians?', with answers on the same seven-point scale as that presented with the 'Quantity' items.

Indirect cross-group friendship was measured with two items ($r = .48$, $N = 63$, $p < .001$): ‘About how many of your closest white friends have friends who are Asian?’ and ‘About how many people in your immediate family have friends who are Asian?’. Answers were given on the previously described ‘None - All’ seven-point scales, so that higher scores indicated a greater degree of indirect cross-group friendship.

Attitudes were assessed using the same measures as in Study 3. After transforming the feeling thermometer scores to put them on the same scale as the other attitude scores (by dividing all scores by 100 and multiplying by seven) and reversing ‘*Friendly - Hostile*’, ‘*Respect - Contempt*’ and ‘*Admiration - Disgust*’ items, the six items were collapsed to form a reliable index ($\alpha = .85$), on which higher scores indicated more positive attitudes towards Asians.

IAT

Participants were presented with Muslim and Christian names (see Figure 5.1) and asked to categorise them as ‘Asian’ or ‘white’, pressing a separate response on the keyboard for each category. They were then presented with positively or

Figure 5.1

Schematic description and illustration of the IAT as used in Study 9

Sequence	1	2	3	4	5	6
Task description	Initial target-concept discrimination	Associated attribute discrimination	Reversed target-concept discrimination	Reversed attribute discrimination	Compatible combined task	Incompatible combined task
Categories & correct response key	White names - S Asian names - K	Positive words - S Negative words - K	White names - K Asian names - S	Positive words - K Negative words - S	White names - S Asian names - K Positive words - S Negative words - K	White names - K Asian names - S Negative words - K Positive words - S
Sample Stimuli	Christopher Charles David James Victoria Elizabeth Catherine Sarah Basimah Ikraam Aziz Sadiq Mohammed Fatima Jamilah Hanif	Lucky Honour Gift Happy Pleasure Miracle Rainbow Peace Poison Disaster Grief Bomb Accident Evil Filth Hatred	Christopher Charles David James Victoria Elizabeth Catherine Sarah Basimah Ikraam Aziz Sadiq Mohammed Fatima Jamilah Hanif	Lucky Honour Gift Happy Pleasure Miracle Rainbow Peace Poison Disaster Grief Bomb Accident Evil Filth Hatred	Lucky Christopher Honour Charles Rainbow David Happy James Pleasure Victoria Miracle Elizabeth Gift Catherine Peace Sarah Poison Basimah Disaster Ikraam Grief Aziz Bomb Sadiq Evil Mohammed Filth Fatima Hatred Jamilah Accident Hanif	Lucky Basimah Honour Ikraam Rainbow Aziz Happy Mohammed Pleasure Sadiq Miracle Fatima Gift Jamilah Peace Hanif Poison Christopher Disaster Charles Grief David Bomb James Evil Victoria Filth Elizabeth Hatred Catherine Accident Sarah

negatively valenced words¹⁴ (also see Figure 5.1) and asked to categorise these as ‘positive’ or ‘negative’. These two blocks were then presented again, with the required response switched, so that where a key had previously meant ‘positive’ it now meant ‘negative’, and where a key had previously meant ‘Asian’ it now meant ‘white’. Participants were then presented with a ‘compatible’ block of trials in which they would see either a name or a word, and be required to press one response key for ‘Asian’ or ‘negative’ and the other response key for ‘white’ or ‘positive’. The final block, of ‘incompatible’ trials, presented participants with either a name or a word, requiring them to press one response key for ‘Asian’ or ‘positive’ and the other response key for ‘white’ or ‘negative’. The difference in response latencies, in milliseconds, between compatible and incompatible trials, is the *IAT effect* (Greenwald et al., 1998). Greenwald, Nosek and Banaji’s (2003) improved scoring algorithm adjusts for weaknesses they have identified in the conventional scoring algorithm. They suggest dividing the difference between the means by the standard deviation of all the latencies from the two ‘test’ blocks (both compatible and incompatible). The resulting ‘D’ score provides a measure on which higher scores can be taken to reflect greater prejudice. As suggested by Greenwald et al., (2003) the latencies of any trials in which the participant had made an error response were replaced with the block mean + 600ms, and the D score was then computed.

Results and Discussion

Table 5.2 shows descriptive statistics for all variables. Extroversion was found, as expected, to be associated with an increased number of Facebook Friends in general, $r = .44, p < .001$. As extroversion has been found to decrease use of the

¹⁴ Taken from Greenwald et al. (1998).

internet for social interaction (Hamburger & Ben-Artzi, 2000), if Facebook Friends represented mainly online relationships, one would expect extroversion to be negatively correlated with number of Facebook Friends. However, the idea that Facebook Friends represent *real-world* friends is supported by the finding that higher extroversion is associated with a larger number of Facebook Friends. In turn, as expected, the number of Facebook Friends in general was associated with a corresponding increase in Asian Facebook Friends, by virtue of the proportion remaining constant, $r = .47, p < .001$. Total Facebook Friends was therefore controlled for in all further analyses of Asian Friends.

As most participants were found to have a very low number of Asian Facebook Friends, modal values being five ($N = 7$) and seven ($N = 7$), the positive skew observed in this variable was not surprising. A logarithmic transformation was applied to successfully normalize the distribution. A slight positive skew was also found in the distribution of total Facebook Friends, and this was corrected with a square root transformation.

Partial correlations were performed to assess the relationship between the log of Asian Facebook Friends and each of the explicit measures of contact, while controlling for the square root of total Facebook Friends. The log of Asian Facebook Friends correlated significantly with quantity of contact, $r = .33, p = .009$, and with direct cross-group friendship, $r = .35, p = .007$, but not with indirect cross-group friendship, $r = .11, p = .42$. This supports the most relevant of the hypotheses; that explicit measures of contact would be validated by unobtrusive measures. Indirect cross-group friendship could not be expected to agree with Asian Facebook Friends, as it does not measure the participant's own number of Asian friends but those of the participant's white friends.

Table 5.2

Means & standard deviations (Study 9)

<i>N</i> = 63	<i>Mean</i>	<i>SD</i>
Total Friends on Facebook	337.38	181.06
Square root of total Friends on Facebook	17.70	4.95
Asian Friends on Facebook	16.32	21.09
Log of Asian Friends on Facebook	.99	.43
Implicit attitudes (D score)	.37	.25
Extroversion	4.48	.88
Quantity of contact	4.01	1.56
Direct cross-group friendship	3.49	1.29
Indirect cross-group friendship	3.52	1.35
Attitudes to Asians	5.07	0.97

As hypothesised, scores on the IAT failed to significantly correlate with self-reported attitudes, $r = -.02$, $p = .91$, presumably due to the increased disparity usually found between implicit and explicit measures of attitudes concerning socially sensitive issues (Greenwald et al., 1998).

The contact hypothesis was supported by the explicit measures, in that explicit attitudes correlated with quantity of contact, $r = .31$, $p = .02$, direct cross-group friendship, $r = .29$, $p = .02$, and indirect cross-group friendship, $r = .25$, $p = .047$. However, unobtrusive measures failed to provide support for the contact hypothesis.

Controlling for the square root of total Facebook Friends, the log of Asian Facebook Friends did not predict scores on the IAT, $r = .14$, ns¹⁵.

General Discussion

Two studies have demonstrated the criterion-related validity of self-reports of contact. Study 8 demonstrated predictive criterion-related validity by showing that self-reported contact (in terms of quantity and quality of contact as well as direct and even indirect cross-group friendship) predicted whether or not participants were found, at a later date, to have experienced a recent instance of online contact with a person of Asian origin, while Study 9 demonstrated concurrent criterion-related validity by showing that self-reported contact (in terms of quantity of contact and direct cross-group friendship) was related to the number of mutually-nominated friendships a participant had with Asians, as measured by the online network, Facebook.

Additionally, Study 8 found that observer-reports of quantity of contact as well as direct cross-group friendship predicted whether the target was likely to have an instance of online contact with an Asian person. This demonstrates the value of the Facebook network for unobtrusively measuring contact, as it can be said to agree with observer-reports, a recognised criterion for demonstrating validity. Interestingly, observer-reports of attitude also predicted online contact, while self-reported attitudes did not. Observers may have inferred attitude from contact, as it would be fairly unusual for individuals to discuss their attitudes towards a certain outgroup. The failure of self-reported attitudes to predict online contact may be due to the increased susceptibility of explicit attitude measures to SDR and other response biases, as they

¹⁵ A series of moderated regression analyses were performed to assess whether scores on the IAT moderated the relationship between explicit and unobtrusive measures of contact. No significant results were found.

might be argued to be relatively subjective when compared to measures of contact. Observer-reports of attitudes did not validate self-reports, suggesting a lack of validity in these explicit attitude measures. As a result of this, an implicit measure of attitudes was included in Study 9, but this also was not successfully predicted by the number of Asians on a participant's Friends page.

The failure of the unobtrusive measures to support the contact hypothesis may be linked to the fact that Study 9 failed to find a significant relationship between explicit and implicit attitudes. This could be due to the socially sensitive nature of the explicit attitude measure (Greenwald et al., 1998). However, there is some controversy in the literature concerning whether the results of the IAT reflect participants' own implicit evaluations of the target group (Greenwald & Banaji, 1995) or simply automatic associations due to a lifetime of cultural learning of stereotypes (Karpinski & Hilton, 2001). The important difference is whether the participants endorse the associations or not. If the IAT measures associations which are endorsed, it can be said to measure implicit prejudice, which will correlate with explicit prejudice to the extent that participants are willing to report honestly and accurately on their level of prejudice (i.e., to the extent that explicit attitude measures are valid). If the IAT measures associations which are culturally learned but *not* endorsed, the lack of relationship between explicit and implicit attitudes found here may simply be due to the fact that the two types of measure are not measuring similar constructs. If this is the case, it is not surprising that the unobtrusive measure of contact did not predict the implicit measure of attitudes.

Study 9 introduced a realistic, practical alternative to analysing mutually-nominated friendships in a relatively small closed environment (e.g., Lee et al., 2007). A far larger proportion of a participant's total number of friendships could be studied,

as it was not necessary to ask every nominated friend to confirm whether the friendship was mutual. Friendships were also not limited to those existing within one sphere of the participant's life; Facebook Friendships can exist between people who know each other from primary school, people who live in the participant's home town, people who the participant met while travelling in another country, people who the participant used to work with, and so forth. Every sphere of the participant's life can be represented, the only limitation being the choice of each individual concerning whether or not to belong to Facebook. Of course, it is impossible to calculate the proportion of people known to the participant who belong to Facebook and the proportion who do not, but I argue that a very large majority of people similar in age to the participants would tend to belong to Facebook. Facebook membership worldwide is currently at over 175 million (Facebook Statistics, 2009) and still climbing fast - 25 million members joined in the last month (Facebook Timeline, 2009). More than 3 billion minutes are spent on Facebook each day (Facebook Statistics, 2009). In addition, this method of measuring contact need not be limited to the student population, as Facebook is not purely a 'student' phenomenon – more than half the members are non-students (Facebook Statistics, 2009). I conclude that Facebook constitutes a practical and effective means of measuring cross-group friendship.

CHAPTER SIX:

Summary and Conclusions

Over nine studies, this thesis examined the hypothesis that self-reports are a valid method of measuring intergroup contact. Self-reports have been used in the vast majority of intergroup contact research (Pettigrew & Tropp, 2006), and the question of their validity for measuring contact has not been addressed until this point. All the studies reported in this thesis lend support to the hypothesis examined by the thesis, leading to the conclusion that self-reports are indeed valid for the measurement of intergroup contact. This chapter summarises the findings of this thesis and outlines the contribution that this thesis has made to the existing research literature concerning the social-psychological study of intergroup contact. The limitations of the research are discussed, and suggestions are made as to what might constitute the next steps to pursue in this line of research.

Thesis Aims

This thesis aimed to explore the validity of self-reports as a measure of contact in terms of their resistance to context effects, their agreement with the ratings of intimate observers, and their concurrent criterion-related validity. Discriminant validity was also demonstrated. Table 6.1 provides an overview of all studies in this thesis.

Table 6.1

Summary of studies in thesis

Study	Participant population	Sample Size	Outgroup	Contact measures included	Variety of validity examined	Method of analysis	Attitude self-reports validated?	Notes
1	White British Students	165	Asians	Direct cross-group friendship at home Direct cross-group friendship at university	Resistance to context effects.	ANCOVA	No	Recent events -> less cont
2	White British Students	118	Rival university	Quantity of contact at university Direct cross-group friendship at university Indirect cross-group friendship at university	Resistance to context effects	One-way ANOVA	No	
3	White British Adults	62 pairs	Asians	Quantity of contact at home Direct cross-group friendship at home Indirect cross-group friendship at home	Agreement with observer-ratings made by spouse	Comparison of dyadic index and pseudo-couple dyadic index	Yes	Skewed data; paired t violated - conservative tes Male self-reports of quant not validated.

4	White British Students	71 pairs	Asians	Quantity of contact at home Direct cross-group friendship at home Indirect cross-group friendship at home Quantity of contact at university Direct cross-group friendship at university Indirect cross-group friendship at university	Agreement with observer-ratings made by parent	Comparison of dyadic index and pseudo-couple dyadic index	No	Quantity of contact and indirect cross-group friendship at university not validated
5	White British first-year Students	89 pairs	Asians	Quantity of contact before university Direct cross-group friendship before university Indirect cross-group friendship before university Quantity of contact after a term at university Direct cross-group friendship after a term at university Indirect cross-group friendship after a term at university	Agreement with observer-ratings made by parent	Comparison of dyadic index and pseudo-couple dyadic index	No	Discriminant validity between time-periods.
6	White British Students	31 groups of 4	Asians	Quantity of contact at university Quality of contact at university Direct cross-group friendship at university Indirect cross-group friendship at university	Agreement with consensual observer-ratings made by 3 close friends	Social Relations Model analysis	Yes	Discriminant validity between extroversion and contact. Indirect cross-group friendship not validated

7	White British Students	34 groups of 4	Asians & gays	Direct cross-group friendship at home Direct cross-group friendship at university	Agreement with consensual observer-ratings made by 3 close friends	Social Relations Model analysis	Asians – Yes Gays – No	Discriminant validity between outgroups. University cross-group friendship with Asians not validated.
8	White British Students	50	Asians	Quantity of contact at university Quality of contact at university Direct cross-group friendship at university Indirect cross-group friendship at university	Criterion-related validity – criterion = online contact with Asians	One-way ANOVA	No	Observer-reports of direct cross-group friendship & attitudes also validated
9	White British Students	63	Asians	Quantity of contact at university Direct cross-group friendship at university Indirect cross-group friendship at university	Criterion-related validity – criterion = number of Asian friends in a very large online community	Linear Regression	No	IAT failed to agree with attitude self-reports or Asian Facebook Friends

Resistance to context effects

Two studies demonstrated that self-reports of contact are resistant to context effects. Previous research has shown that attitude measurement is extremely context-dependent (Schwarz, 1999), with preceding questions influencing the answers to subsequent ones (e.g., Schwarz & Sudman, 1992). Studies 1 and 2 sought to demonstrate that self-reports do not share this weakness. Study 1 validated self-reports of direct cross-group friendship with Asians at home and at university by demonstrating their resistance to the effects of being asked to recall a positive or negative encounter with an Asian person prior to completing the contact measures. In comparison with a control group who were asked to recall an encounter *after* completing contact and attitude measures, the act of recalling an encounter affected participants' self-reported attitudes in the direction of the valence of the encounter, but did not affect contact self-reports at all. This study was complicated by the fact that participants who recalled a recent encounter tended to report less cross-group friendship than did those who chose a less recent encounter to recall. I argue that the encounters of those participants who had relatively low contact with Asians would tend to be more superficial and more difficult to remember (Craig & Lockhart, 1972), leading them to be forced to draw on more recent events. The improved design of Study 2 overcame this limitation by providing an event for participants to read, rather than asking them to recall an event freely. Study 2 validated quantity of contact, direct cross-group friendship, and even indirect cross-group friendship with students from a rival university, through demonstrating that, while attitudes were affected in accordance with the valence of the prime, self-reported contact was unaffected.

Agreement with observer-reports

Observer-reports are arguably the best method developed to date for validating self-reports (Howard, 1994) and have been used extensively for that purpose in

personality research (e.g., Piedmont et al., 2000). Five studies assessed this important criterion for validity. The analytic technique was of importance in these studies, as they involve non-independent data, which is often erroneously treated as if it were independent (Kenny et al., 2006). An analytic technique appropriate to both the design and the non-independent nature of the data was chosen for each study.

Study 3 demonstrated the validity of self-reports of quantity of contact, direct and even indirect cross-group friendship with Asians by comparing them with the reports of the observer who is, arguably, best equipped to judge - the spouse (McCrae, 1993). Dyadic indices (Kenny et al., 2006) were computed for married couples and for pseudo-couples (Corsini, 1956, i.e., a male and a female paired together at random) and these were compared so that any self-observer agreement due purely to stereotype accuracy (Cronbach, 1955, self-observer agreement resulting from the similarity of the target to the norm) could be separated from self-observer agreement due to unique knowledge between spouses.

An additional finding was that length of relationship moderated the relationship between self-disclosure and self-observer agreement. The degree to which a couple disclosed to one another predicted the discrepancy between their self- and observer-reports of direct cross-group friendship, but only if they had been together a relatively short time. Those couples who had been together a longer time showed very high levels of self-observer agreement independent of self-disclosure. I argue that couples who have been together a longer time already know the most intimate details about each other, and self-disclosure can have no further effect. It may also be the case that the issues about which couples disclose may change over the course of their relationship. Self-disclosure may be relevant to self-observer agreement on direct cross-group friendship in the early stages of the relationship because couples may discuss the ethnicities of their friends,

while older couples may self-disclose on very different topics. Further research into the content of self-disclosure at various stages of a relationship would be necessary to illuminate this question.

Study 4 addressed a limitation of Study 3 – the population most commonly studied in contact research is the student population, due to the proximity, ease and practicality of participant recruitment within this population. Studies 4 and 5 focused on validating the self-reported contact of students, through the observer-reports of the people who, arguably, are likely to know them best, their parents. Using the same dyadic technique as in Study 3, Study 4 validated students' self-reports of quantity of contact, direct and indirect cross-group friendship at home, and direct cross-group friendship at university.

This study failed to validate quantity of contact or indirect cross-group friendship at university. There are several possible explanations for this finding. First, discrepancy in this study was very low, with scores tending towards zero, and this caused a positive skew in the data, to the extent that no transformation had the capacity to normalise the distribution. This violated the paired *t*-test's assumption of normality, and decreased the likelihood of finding a significant result. Secondly, with regard to indirect cross-group friendship specifically, it is a very difficult construct about which to make an informed estimate, as it requires knowing not only about the participant's friends, but the friends of their friends. While parents might be able to interact with participants' friends at home, and find out a little, at least, about the friends of their child's friends, this would not be possible at university. It is also likely that parents know less about their child's quantity of contact at university than at home, where they can directly observe a larger proportion of the child's contact, and where they are in constant contact with the child, with the opportunity for increased self-disclosure. Indeed, Study 4 found that the more a student

tends to disclose to their parent, the smaller the discrepancy between the parent's and the student's ratings of how much contact and indirect cross-group friendship the student has at home.

Study 5 replicated the design of Study 4, with the addition of a time-period factor. In this way, Study 5 was able to demonstrate discriminant validity. Study 5 validated current quantity of contact and direct and indirect cross-group friendship with Asians, as well as quantity of contact and direct and indirect cross-group friendship with Asians taking place at home before the participant had attended university. In addition, it was found that students and parent observers can distinguish between current contact and contact occurring before coming to university. Parents' observer-reports of retrospective contact did not agree with students' self-reported current contact, and neither did parents' observer-reports of current contact agree with students' self-reported retrospective contact. This indicates that observers are not basing their ratings on, for example, some stable aspect of the target's personality which might give them clues about their child's level of contact across all situations and time periods, but on what they know about how much contact the target actually has at a given time and in a given situation. Despite the difficulty of obtaining accurate self-insight (Haddock, Rothman, Reber, & Schwarz, 1999), and the inaccessibility of infrequently-expressed constructs (Basilli, 1993, 1996a, 1996b), Study 5 demonstrated that an individual's level of contact is a construct which *is* accessible to that individual, and available to be validly rated.

Studies 6 and 7 addressed two limitations of the previous studies. Firstly, contact at university had not been consistently validated by observer-reports, possibly due to the difficulty for parents of estimating levels of contact in a situation they could not directly observe. Close friends at university were deemed to be useful observers for validating contact at university, while parents remained the most useful observers for validating

contact at home. Secondly, the previous studies used only a single observer, exposing observer-reports to the error variance created by individual response biases. Validity would be increased by increasing the number of observers (Kolar, Funder, & Colvin, 1996) as the error variance associated with each individual observer would be unlikely to be shared, and that part of the variance which is consensual would therefore provide an ideal criterion by which to validate self-reports (Kenny, 1994).

Study 6 validated self-reports of quantity and quality of contact and direct cross-group friendship with Asians at university through examining agreement with the consensual ratings of three of the participants' closest friends. The Social Relations Model was used to analyse the data, as this is the best way to deal with non-independence (Kenny, 1994). Indirect cross-group friendship was not validated, suggesting that it is difficult for even very close friends to know about the contact of their friend's other friends. Discriminant validity was again demonstrated, showing that friends were not basing their observer-ratings of contact on a general appraisal of how outgoing the target is. Self-observer agreement was unchanged by partialling out the effects of extroversion as rated by the target and by the observers.

Study 7 replicated this design, with a reduction in measures to allow for the addition of a second outgroup – that of homosexual men. Clear indications of discriminant validity were provided, as observer-reports of contact with one outgroup did *not* agree with targets' self-reports of contact with the *other* outgroup. In other words, people do not rate themselves or someone else as higher in contact with outgroups in general; they are able to distinguish between outgroups. Study 7 succeeded in validating self-reports of direct cross-group friendship with gay men at home and at university, as well as self-reports of direct cross-group friendship with Asians at home. Self-reports of

direct cross-group friendship with Asians at university were not validated, and further research would be useful to discover whether this failure is meaningful or artifactual. These two studies demonstrate the validity of self-report measures of intergroup contact based on consensually supported observer-reports, analysed using a technique appropriate to non-independent data.

Criterion-related validity

Two studies examined criterion-related validity. Taking as a criterion Asian friendship as represented by the online network Facebook, Study 8 sought to discover whether such a criterion was itself valid, through comparison with observer-reports of direct cross-group friendship. It was concluded that Facebook contact did indeed represent real-world contact in that observer-reports corroborated contact on Facebook. Study 8 also successfully validated self-reports of quantity and quality of contact as well as direct and indirect cross-group friendship with Asians at university, by demonstrating that they show predictive criterion-related validity. Explicit contact measures are able to predict which group a participant will be found to belong to at a later date: those with, or those without a recent instance of online contact with Asians. The dependent variable used in Study 8 was found to be less than ideal, and in response to this, Study 9 determined the total number of Asian friendships mutually-nominated to appear on each participant's Facebook Friends page. Study 9 also successfully validated quantity of contact and direct and indirect cross-group friendship, by demonstrating that they show concurrent criterion-related validity.

Studies 8 and 9 introduce a potentially very useful method of assessing mutually-nominated cross-race friendships. Facebook is a very large online community which reflects real-world friends rather than merely 'e-relationships'. As mutually-nominated friendships from all spheres of life can be examined, there is no need to recruit all a

participant's friends to take part themselves in order to corroborate a friendship, nor to limit the nominated friendships to a relatively small closed environment in order to recruit everyone within that environment (e.g., Lee et al., 2007). It may also be possible to recruit participants directly from Facebook, as any member can contact any other member (unless that member has specifically requested not to be contacted). Participants can be recruited from all over the world for online participation. I suggest that Facebook constitutes a practical and effective means of measuring cross-group friendship, whose potential has not yet begun to be explored.

Theoretical implications of the research

Validity of self-reports for measuring contact

This thesis goes some way toward answering a question which has troubled intergroup contact researchers. The weaknesses of self-reports have been well documented, and yet the majority (81%, Pettigrew & Tropp, 2006) of research on the contact hypothesis has relied on self-report measures of contact (necessarily so, with regards to the quality of that contact, and to direct and indirect cross-group friendships). The question of whether self-reports are a valid method of assessing intergroup contact is an important one, as if it transpires that they are not, the derived implications of the whole research endeavour on the contact hypothesis are potentially flawed. The studies presented in this thesis support the validity of self-reports as a contact measure, providing evidence that research on intergroup contact has relied on a valid technique, and promoting the continued use of self-reports to measure contact.

While this thesis has supported the validity of contact self-reports, it has also introduced a useful means of objectively measuring contact, in as far as a person's friendships are for the most part represented on the online network, Facebook. I have described how this network can be used to conduct an entire experiment, with

participants recruited through Facebook, online questionnaires presented to them, and their mutually-nominated cross-race friendships counted.

Attitude self-reports

While this thesis does not focus on the validity of self-reported outgroup attitudes, they have consistently been included in every study presented here. Two studies found, in accordance with the research literature (e.g., Schwarz, 1999; Schwarz & Sudman, 1992) that attitude self-reports are not resistant to context effects. One study demonstrated that a married couple agree on each other's outgroup attitudes, but two studies demonstrate that a parent does not agree with their child's self-ratings of outgroup attitudes. Two studies demonstrated that three close friends consensually agree with a participant concerning his or her attitudes towards Asians, but not towards gay men. Finally, an Implicit Association Test failed to corroborate self-reported attitudes towards Asians, replicating previous work showing that socially sensitive issues increase disparity between implicit and explicit measures (Greenwald et al., 1998). It seems likely that self-reports are, at best, inconsistent in their validity for the measurement of attitudes, and that non-self-report measures ought to be adopted for measuring the dependent variable in future research into the effects of intergroup contact.

Low contact

Both Studies 8 and 9 found a surprisingly low level of contact as measured unobtrusively using the online network 'Facebook'. Study 8 found that the overwhelming majority (30 out of 50 participants) had not had a recent instance of online contact with an Asian, and Study 9 found, despite the range of Asian Friends extending up to 111, that the majority (52.4%) of participants had 10 or fewer Asian Facebook Friends. Approximately 4.6% of students accepted to study at Oxford each year are Asian while 86.6% are white (Undergraduate Admissions Office, 2007). The most recent information

concerning the ethnic make-up of students at Oxford University reports that, out of a total of 19,993 students, Oxford has 404 Asian undergraduates and 421 Asian postgraduates, with 253 undergraduates and 80 postgraduates describing themselves as mixed-race Asian and white (Oxford University Gazette, 2008). These figures suggest that Asians constitute as little as .04% of the total student population of Oxford, and go some way to explaining the low contact found in Studies 8 and 9.

Potential limitations of the research

One weakness of the research presented in this thesis is that every study reported here assessed intergroup contact from the perspective of the majority status group; whites. This was necessary, as it is the attitudes of this group with whom the majority of contact research has been concerned, as it is their prejudice which constitutes a social problem. It is therefore the self-reports of the majority group which need to be assessed for validity. However, it must be pointed out that the effects of contact have been found to be weaker among members of minority status groups than among majority status groups (Tropp & Pettigrew, 2005). Tropp and Pettigrew (2005) reported that the concerns of majority and minority groups during an interaction are very different, with minority groups being concerned about being the target of prejudice, while majority groups are more concerned that the minority group should not see them as prejudiced (e.g., Plant, 2004). Members of minority status groups may not interpret an intergroup interaction as positively as majority status groups do, and may not share the same criteria for evaluating whether Allport's (1954/1979) optimum conditions have been met (Robinson & Preston, 1976). An intergroup encounter which might be ideal for fostering improved attitudes in a majority group member may have greatly reduced benefits for the attitudes of a minority group member (Tropp & Pettigrew, 2005), partly due to an ongoing awareness of their own group's devalued status. While this thesis has validated the self-reported contact of

majority status whites, no attempt has been made to validate the self-reports of minority group members. Further research could examine this important issue, as there is a great deal of work to be done in ensuring that an intergroup encounter is beneficial not only for a majority status participant but also for the potential prejudice target.

It was found in Chapter Three that the paired *t*-test's assumption of normality was consistently violated in the distribution of the discrepancy variables and the pseudo-couple discrepancy variables. I have outlined that this is inevitable when dealing with data which tends towards zero to such an extreme degree. However, it remains the case that the test on which the three studies reported in that chapter base their conclusions was conducted under conditions which violated one of its assumptions. This will have led to a decrease in power, and the risk of Type II errors. However, I argue that when a test is used to support the validity of a research methodology it is better to be too conservative than to risk erroneously supporting the use of that research methodology. The non-parametric alternative, the Wilcoxon signed-rank test, assumes that the data be symmetrically distributed, which they are not. The use of Structural Equation Modelling (SEM) would control for substantive variance to the extent that each member of a dyad is similar to the other member (Cline, 1964), and this reduces the effectiveness of SEM as a technique for analysing dyadic data where a high degree of similarity is expected within the pairs. Controlling for self-projection and stereotype accuracy is only beneficial to the extent that the target is unlike the stereotype, and target and judge are dissimilar. I suggest that there seems to be no preferable alternative to the paired *t*-test.

While direct cross-group friendship and quantity of contact have been consistently validated throughout each study presented in this thesis, the same is not the case for indirect cross-group friendship. Study 4 failed to find self-observer agreement between students and their parents for indirect cross-group friendship at university, and Study 6

found that the consensual observer-ratings of three close friends did not agree with the target's self-ratings of indirect cross-group friendship. Study 7 did not attempt to validate indirect cross-group friendship as a result of the finding in Study 6. I argue that it is more surprising that indirect cross-group friendship was validated in so many studies than that it was not validated in Studies 4 and 6. An observer-report of indirect cross-group friendship requires observers to estimate the contact not of the target but of the target's friends, which is understandably very difficult. The fact that indirect cross-group friendship was validated in three out of four cases by the observer-reports of a parent, and was consistently validated by the observer-reports of a spouse, seems to promote optimism for the research endeavour into extended contact. Nevertheless, an alternative, unobtrusive measure of indirect cross-group friendship would be a boon (see below).

Potential directions for future research.

Study 1 found, when analysing primed and unprimed data together, that positive events proved to be slightly easier to recall than negative events. It would appear that, in accordance with previous research (Haddock et al., 1999; Schwarz et al., 1991) ease of recall may have amplified the effects of valence of prime on attitudes. This highlights a potential prejudice intervention paradigm, which could be explored by future research. I have already described (see Chapter Two) the work by Schwarz et al. (1991) and Haddock et al. (1999) which required participants to think of either many (difficult) or few (easy) examples of the variable of interest. In accordance with this method, participants could be asked to think of either many or few examples of positive or negative outgroup encounters they have experienced. If, as might be predicted by the aforementioned research, the 'difficult – negative' and 'easy - positive' conditions result in improved attitudes in comparison with controls, this would offer positive implications for the prejudice intervention research effort. Such a technique would attempt to alter a

participant's standard of comparison in order to change the way in which they view their existing intergroup encounters, rather than providing new intergroup encounters.

Although this method has not yet been attempted as a prejudice intervention paradigm, it might prove interesting to explore it further.

Moving on to the next empirical chapter, Study 5 examined retrospective self- and observer-ratings of contact before going to university, in comparison with concurrent self- and observer-ratings of contact whilst at university. The next logical step to take in pursuing this line of research is to gather data from students immediately on arrival at university, for example at a departmental orientation event, concerning their levels of contact and attitudes at home. These same students can then be tested again after they have been at university for a term, or perhaps at the end of each term, to examine how their contact changes. Most importantly, it would be interesting to examine whether the presumed increase in contact is associated with an increase in positive attitudes towards the outgroup.

In the third empirical chapter, Study 7 showed that targets and observers can distinguish between an Asian outgroup and a homosexual outgroup. This distinction may be relatively easy, given that contact with one of these outgroups can be readily observed while the other requires knowledge of the sexuality of the target's acquaintances. I suggest that it would be useful to conduct further research to determine whether observers can distinguish between two ethnic outgroups, such as blacks and Asians. The difficulty here is that, in the United Kingdom, the size of the black minority group (25% of the minority ethnic population, Office of National Statistics, 2001) is very different from the size of the Asian minority group (50% of the minority ethnic population, Office of National Statistics, 2001). Observers are likely to be able to estimate that any particular target will have less contact with blacks than with Asians. It may be difficult to find two

minority ethnic groups of similar size in the United Kingdom, so I suggest that such research should take place elsewhere.

Turning to the final empirical chapter, I have already described the benefits of using Facebook to measure direct cross-group friendship, and suggested that future research would benefit from using this method. However, I now suggest that future research might also use Facebook to measure indirect cross-group friendship. The Friends pages of each white person on a participant's Friends page could be examined, with a view to recording the ethnic origins of all the Friends of the participant's Friends. Although such an endeavour would be time-consuming, researchers could thereby obtain a measure of indirect cross-group friendship which would arguably be based on a far more concrete footing than the participant's self-reported indirect cross-group friendship. Another avenue for future research in this area might be to explore the effects of *online* indirect cross-group friendship. Facebook contact is visible to others. People can see how many Asian Friends a person has, although they would have to make the effort to find out. However, people can easily see when an Asian person comments on someone's page, or photos, without specifically looking for this information. This constitutes indirect online contact, and its effects might be worth exploring.

An acknowledged limitation of this thesis is that it examines only the self-reports of majority group members. A great deal of further work is necessary to evaluate the validity of self-reports for examining the contact experienced by minority groups. Such a task was beyond the scope of this thesis, due to the lack of availability of members of multiple minority groups in large enough numbers. Cultural differences may also be a factor in obtaining the observer-reports of family members in order to validate participants' self-reports. However, online research following the guidelines of Vazire (2006) may facilitate the acquisition of observer-corroborated self-reports, and future

research could examine the validity of at least the largest minority group in the United Kingdom; that of Asians.

Summary

This thesis constitutes an original contribution to the research literature on intergroup contact by demonstrating for the first time the validity of self-report measures of intergroup contact based on a variety of validation methods, and, where necessary, analysed using techniques appropriate to non-independent data. Taken together, these nine studies provide substantial support for the continued use of self-reports in contact research. Whereas attitude self-reports have revealed by-now-well-established priming or context effects, self-reports of contact appear to be robust to context effects. While attitude self-reports are inconsistently corroborated by observer-reports, contact self-reports show consistent agreement with the reports provided by spouses, parents, and close friends. Finally, while an implicit measure of attitudes failed to agree with the explicit measure, unobtrusive and objective measures of contact agreed without exception with self-reports of contact. While explicit attitude measures assess a fairly subjective construct, the same does not appear to be true of explicit contact measures. These appear to be accessible to the individual and available to be reported upon. This is heartening news for advocates of intergroup contact, and those who apply it in social interventions, given the very high reliance on self-report measures in the voluminous literature on intergroup contact (see Pettigrew & Tropp, 2006).

REFERENCES

- Aboud, F. E., Mendelson, M. J., & Purdy, P. (2003). Cross race peer relationships and friendship quality. *International Journal of Behavioural Development*, 27, 165-173.
- Acitelli, L. K., Kenny, D. A., & Weiner, D. (2001). The importance of similarity and understanding of partners' marital ideals to relationship satisfaction. *Personal Relationships*, 8, 167-185.
- Allport, G. (1954/1979). *The nature of prejudice*. Reading, UK: Addison-Wesley.
- Amichai-Hamburger, Y., & McKenna, K. Y. A. (2006). The contact hypothesis reconsidered: Interacting via the Internet. *Journal of Computer-Mediated Communication*, 11(3), article 7.
- Amir, Y. (1969). Contact hypothesis and ethnic relations. *Psychological Bulletin*, 71, 319-342.
- Arbisi, P. A., & Ben-Porath, Y. S. (1995). An MMPI-2 infrequent response scale for use with psychopathological populations: The infrequency psychopathology scale, F(p). *Psychological Assessment*, 7, 424-431.
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63, 596-612.
- Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including other in self. *Journal of Personality and Social Psychology*, 60, 241-253.
- Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin*, 23, 363-377.

- Aronson, E., & Bridgeman, D. (1979). Jigsaw groups and the desegregated classroom: in pursuit of common goals. *Personality and Social Psychology Bulletin*, 5, 438-446.
- Babington, C. (2008, September 22). Poll: Views still differ sharply by race [Electronic version]. *Associated Press*.
- Baker, P. E. (1934). *Negro-White adjustment*. New York, NY: Association Press.
- Bargh, J. A., & McKenna, K. Y. A. (2004). The Internet and social life. *Annual Review of Psychology*, 55, 573-590.
- Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, 81, 261-272.
- Bass, B. M., & Yammarino, F. J. (1991). Congruence of self and others' leadership ratings of Naval officers for understanding successful performance. *Applied Psychology*, 40, 437-454.
- Bassili, J. N. (1993). Response latency versus certainty as indexes of the strength of voting intentions in a CATI survey. *Public Opinion Quarterly*, 57, 54-61.
- Bassili, J. N. (1996a). Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, 71, 637-653.
- Bassili, J. N. (1996b). The how and why of response latency measurement in telephone surveys. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative process in survey research* (pp. 319-346). San Francisco, CA: Jossey-Bass.

- Bechtel, R. B., Achelpohl, C., & Akers, R. (1972). Correlates between observed behavior and questionnaire responses on television viewing. In E. A. Rubinstein, G. A. Comstock, & J. P. Murray (Eds.), *Television and human behavior: A technical report to the surgeon general's advisory committee on television and social behavior* (Vol. 4, pp. 274-344). Washington, DC: Government Printing Office.
- Becker, G., & Bakal, D. (1970). Subject anonymity and motivational distortion in self-report data. *Journal of Clinical Psychology*, 26, 207–209.
- Ben Ari, R., & Amir, Y. (1986). Contact between Arab and Jewish youth in Israel: Reality and potential. In R. Brown & M. Hewstone (Eds.), *Contact and conflict in intergroup encounters* (pp. 45–58). Cambridge, MA: Basil Blackwell
- Ben-Ze'ev, A. (2004). *Love online: Emotions on the Internet*. Cambridge University Press: Cambridge.
- Bergholz, R. (1982, November 7). Anti-Bradley vote seen as key election factor [Electronic version]. *Los Angeles Times*.
- Bertrand, M. & Mullainathan, S. (2003). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market distribution. *MIT Working Paper Series*, 03–22.
- Bettencourt, B. A., Brewer, M. B., Croak, M. R., Miller, N. (1992). Cooperation and the reduction of intergroup bias: The roles of reward structure and social orientation. *Journal of Experimental Social Psychology*, 28, 301-319.
- Blair, I. (2001). Implicit stereotypes and prejudice. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 359–374). Mahwah, NJ: Erlbaum.

- Blanchard, F. A., Adelman, L., & Cook, S. W. (1975). Effect of group success and failure upon interpersonal attraction in cooperating interracial groups. *Journal of Personality and Social Psychology*, 31, 1020-1030.
- Block, J. H., & Block, J. (1980). The role of ego-control and egoresiliency in the organization of behavior. In A. W. Collins (Ed.), *The Minnesota Symposia on Child Psychology: Vol. 13. Development of cognition, affect, and social relations* (pp. 39-101). Hillsdale, N J: Erlbaum.
- Bodenhausen, G. V (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science*, 1, 319-322.
- Bodenhausen, G. V., & Wyer, R. S. (1985). Effects of stereotypes in decision making and information-processing strategies. *Journal of Personality and Social Psychology*, 48, 267-282.
- Bond, C. F. J. & Kenny, D. A. (2002). The triangle of interpersonal models. *Journal of Personality and Social Psychology*, 83, 355-366.
- Bonebrake, K. (2002). College students' Internet use, relationship formation, and personality correlates. *Cyber Psychology & Behavior*, 5, 551-557.
- Borkenau, P., & Liebler, A. (1993). Consensus and self-other agreement for trait inferences from minimal information. *Journal of Personality*, 61, 477-496.
- Borkenau, P., & Liebler, A. (1995). Observable attributes as manifestations and cues of personality and intelligence. *Journal of Personality*, 63, 1-25.
- Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behaviour as cues of personality and intelligence. *Journal of Personality and Social Psychology*, 86, 599-614.
- Borkenau, P., & Ostendorf, F. (1992). Social desirability as moderator and suppressor variable. *European Journal of Personality*, 6, 199-214.

- Boruch, R. F., & Cecil, J. S. (1979). *Assuring the confidentiality of social research data*. Philadelphia, PA : University of Pennsylvania Press.
- Boulton, M. J., & Smith, P. K. (1996). Liking and peer perceptions among Asian and white British children. *Journal of Social and Personal Relationship*, 13, 163-177.
- Bradburn, N. M. (1983). Response effects. In P. Rossi, J. Wright, & A. Anderson (Eds.), *Handbook of Survey Research* (pp. 289-328). New York, NY: Academic Press.
- Bradley, J. V. (1968). *Distribution-free statistical tests*. Englewood Cliffs, NJ: Prentice-Hall.
- Braha, V., & Rutter, D. R. (1980). Friendship choice in a mixed-race primary school. *Educational Studies*, 6, 217–223.
- Brewer, M. B. and Miller, N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In N. Miller and M. B. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 281-302). New York, NY: Academic Press.
- Brooks, D. (1975). *Race and labour in London transport*. London, UK: Oxford University Press.
- Brown, J. D., & Dutton, K. A. (1995). Truth and consequences: The costs and benefits of accurate self-knowledge. *Personality and Social Psychology Bulletin*, 21, 1288–1296.
- Brown, R. & Hewstone, M. (2005). An integrative theory of intergroup contact. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 37, pp. 255-343). San Diego, CA: Elsevier Academic Press.

- Brown, R., & Wade, G. (1987). Superordinate goals and intergroup behaviour: The effect of role ambiguity and status on intergroup attitudes and task performance. *European Journal of Social Psychology, 17*, 131-142
- Brownstein, A. L. (2003). Biased predecision processing, *Psychological Bulletin, 129*, 545–568.
- Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology, 67*, 366– 381.
- Burisch, M. (1984b). You don't always get what you pay for: Measuring depression with short and simple versus long and sophisticated scales. *Journal of Research in Personality, 18*, 81-98.
- Campbell, D. T. (1965). Ethnocentric and other altruistic motives. In D. Levine (Ed.), *Nebraska symposium on motivation*, (pp. 283-311). Lincoln, NE: University of Nebraska Press.
- Campbell, D. T., Kruskal, W., & Wallace, W. (1966). Seating aggregation as an index of attitude. *Sociometry, 29*, 1-15.
- Campbell, D. T., Stanley, J. C., & Gage, N. L. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and Validity Assessment*. Beverly Hills: SAGE.
- Cartwright, D. (1957). Social psychology and group processes. *Annual Review of Psychology, 8*, 211-236.
- Casselden, P. A., Hampson, S. E. (1990). Forming impressions from incongruent traits. *Journal of Personality and Social Psychology, 59*, 353-62

- Center on Congress. (2008). *Learn about congress*. Retrieved 20 February, 2009, from http://www.centeroncongress.org/learn_about/feature/qa_members.html#makeup
- Chen, M., & Bargh, J. (1997). Nonconscious behavioural confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, 33, 541-560.
- Christensen-Szalanski, J. J., & Bushyhead, J. B. (1981). Physicians' use of probabilistic information in a real clinical setting. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 928-935.
- Christensen-Szalanski, J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior & Human Decision Processes* 48, 147-168.
- Citrin, J., Green, D. P. & Sears, D. O. (1990). White reactions to black candidates: When does race matter? *The Public Opinion Quarterly*, 54, 74-96.
- Clack, B., Dixon, J. A. & Tredoux, C. (2005). Eating together apart: Patterns of segregation in a multiethnic cafeteria. *Journal of Community and Applied Social Psychology*, 15, 1-16.
- Clark, M. L., & Ayers, M. (1992). Friendship similarity and proximity during early adolescence: Gender and racial patterns. *The Journal of Psychology*, 126, 393-405.
- Clayton, D. (2007). The audacity of hope. *Journal of Black Studies*, 38, 51-63.
- Cline, V. A. Interpersonal perception. In B. A. Maher (Ed.), *Progress in experimental personality research* (Vol. 1, pp. 221-284). New York, NY: Academic Press.
- Clymer, A. (1989, November 12). Election day shows what the opinion polls can't do [Electronic version]. *New York Times*.
- College Board. (1976-1977). *Student descriptive questionnaire*. Princeton, NJ: Educational Testing Service.

- Colvin, C. R., & Block, J. (1994). Do positive illusions foster mental health? An examination of the Taylor and Brown formulation. *Psychological Bulletin*, 116, 3–20.
- Commission for Racial Equality. (2007, September 19). *A lot done, a lot to do: Our vision for an integrated Britain*. Final report of the CRE. Retrieved 9 February, 2009, from <http://www.equalityhumanrights.com/Documents/Race/General%20advice%20and%20information/A%20lot%20done,%20a%20lot%20to%20do%20-%20Our%20vision%20for%20an%20integrated%20Britain.pdf>
- Cook, S. W. (1962). The systematic analysis of socially significant events: A strategy for social research. *Journal of Social Issues*, 18, 66-84.
- Cooper, A. C., & Sportolari, L. (1997). Romance in Cyberspace: Understanding online attraction. *Journal of Sex Education and Therapy*, 22, 7-14.
- Cooper, A. C., Woo, C. Y., & Dunkelberg, W. C. (1988). Entrepreneurs' perceived chances for success. *Journal of Business Venturing*, 3, 97–108.
- Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, 83, 1314–1329.
- Corsini, R. J. (1956). Understanding and similarity in marriage. *Journal of Abnormal and Social Psychology*, 52, 327-332.
- Cose, E. (2006, October 30). The “Bradley effect.” [Electronic version]. *Newsweek*.
- Costa, P. T., Jr., & McCrae, R. R. (1988). Personality in adulthood: A 6-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology*, 54, 853-863.

- Cracknell, R. (2005). *Social background of MPs*. Standard Note number 1528, Social & General Statistics Section, House of Commons Library. Retrieved 9 February, 2009, from <http://www.parliament.uk/commons/lib/research/notes/snsg-01528.pdf>
- Craik, F. I. M. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour*, 11, 671-684.
- Cronbach, L. J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity." *Psychological Bulletin*, 52, 177- 193.
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 443-507). Washington, DC: American Council on Education.
- Cronbach, L. J., Glaser, G. C., Nanda, H., & Raiaratnam, N. (1972). *The Dependability of Behavioral Measurements*. New York, NY: John Wiley & Sons.
- Crosby, R. A., & Yarber, W. L. (2001). Perceived versus actual knowledge about correct condom use among U. S. adolescents. *Journal of Adolescent Health*, 28, 415–420.
- Cross, P. (1977). Not can but will college teaching be improved? *New Directions for Higher Education*, 17, 1–15.
- Crowne, D. P. (1979). *The experimental study of personality*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Cummings, S. & Lambert, T. (1997). Anti-Hispanic and anti-Asian sentiments among African Americans. *Social Science Quarterly*, 78, 338-353.
- Dambrun, M. & Guimon, S. (2004). Implicit and explicit measures of prejudice and stereotyping: Do they assess the same underlying knowledge structure? *European Journal of Social Psychology*, 34, 663-676.

- Daniels, D., & Plomin, R. (1985). Differential experiences of siblings in the same family. *Developmental Psychology*, 21, 747-760.
- Davey, A. (1983). *Learning to be prejudiced. Growing up in multi-ethnic Britain*. London, UK: Arnold.
- Denscombe, M., Szulc, H., Patrick, C., & Wood, A. (1986). Ethnicity and friendship: The contrast between sociometric research and fieldwork observation in primary school classrooms. *British Educational Research Journal*, 12, 221–236.
- DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review*, 1, 346–357.
- Deutsch, M. & Collins, M. (1951). *Interracial housing: A psychological evaluation of a social experiment*. Minneapolis: University of Minnesota Press.
- Dicken, C. (1963). Good impression, social desirability, and acquiescence as suppressor variables. *Educational and Psychological Measurement*, 23, 699-720.
- Diener, E., Sandvik, E., Pavot, W., & Gallagher, D. (1991). Response artifacts in the measurement of subjective well-being. *Social Indicators Research*, 24, 35-56.
- Dixon, J., & Durrheim (2003). Contact and the ecology of racial division: Some varieties of informal segregation. *British Journal of Social Psychology*, 42, 1-23
- Dixon, J. A., Tredoux, C., Durrheim, K., Finchilescu, G., & Clack, B. (2008). Mapping the micro-ecology of racial segregation ‘the inner citadels of the color line’: Mapping the micro-ecology of racial segregation in everyday life spaces. *Social and Personality Psychology Compass*, 2, 1547–1569.
- Dodd, V. (2005, July 23). Religious hate crimes rise fivefold [Electronic version]. *Guardian*.

- Dovidio, J. F., & Fazio, R. H. (1992). New technologies for the direct and indirect assessment of attitudes. In J. Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 204-237). New York, NY: Russell Sage Foundation.
- Dovidio, J. F., Kawakami, K., & Beach, K. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In R. Brown & S. L. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup relations* (pp. 175-197).
- Dowd, M. (2008, May 14). Raspberry for Barry [Electronic version]. *New York Times*.
- DuBois, D., & Hirsch, B. (1990). School and neighbourhood friendship patterns of Black and Whites in early adolescence. *Child Development*, 61, 524-536.
- Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of blacks. *Journal of Personality and Social Psychology*, 34, 590-598.
- Duncan, S., Kanki, B., Mokros, H., & Fiske, D. W. (1984). Pseudounilaterality and other ills to which interaction research is heir. *Journal of Personality and Social Psychology*, 46, 1335- 1348.
- Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York, NY: Psychology Press.
- Dunning, D., & Parpal, M. (1989). Mental addition and subtraction in counterfactual reasoning: On assessing the impact of actions and life events. *Journal of Personality and Social Psychology*, 57, 5-15.

- Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, 23, 316-326.
- Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, 37, 90-93.
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2004). *Why the unskilled are unaware: Further explorations of (absent) self insight among the incompetent*. Unpublished manuscript, Cornell University, Ithaca, NY.
- Emler, N. (2002). Morality and political orientations: an analysis of their relationship. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology* (Vol. 13, pp. 259-291). Chichester, UK: Wiley.
- Emler, N., Tarry, H., & St. James, A. (2007). Postconventional moral reasoning and reputation. *Journal of Research in Personality*, 41, 76-89.
- Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self or social prediction? *Journal of Personality and Social Psychology*, 79, 861– 875.
- Epley, N., & Dunning, D. (2004). *The mixed blessings of self-knowledge in behavioral prediction*. Unpublished manuscript, Harvard University, Cambridge, MA.
- Espenshade, T. J. (1997). Taking the pulse of public opinion toward immigrants. In T. J. Espenshade, (Ed.), *Keys to successful immigration: Implications of the New Jersey experience* (pp. 89-116). Washington, DC: The Urban Institute Press.
- Eysenck, H. J. (2008). Genetic and environmental contributions to individual differences: The three major dimensions of personality. *Journal of Personality*, 58, 245-261.

- Facebook Network Information. (2009). Retrieved February 9, 2009, from
<http://oxford.facebook.com/networks/?nk=16777585>
- Facebook Statistics. (2009). Retrieved February 23, 2009, from
<http://www.facebook.com/press/info.php?statistics>
- Facebook Timeline. (2009). Retrieved February 23, 2009, from
<http://www.facebook.com/press/info.php?timeline>
- Felson, R. B. (1981). Ambiguity and bias in the self-concept. *Social Psychology Quarterly*, 44, 64-69.
- Festinger, L., & Kelley, H. (1951). *Changing attitudes through social contact*. Ann Arbor, MI: Institute for Social Research, University of Michigan.
- Finkel, S. K., Guterbock, T. M. & Borg, M. J. (1991). Race of interviewer effects in a preelection poll. *Public Opinion Quarterly*, 55, 313-330.
- Fischhoff, B. (1975). Hindsight $\neq$ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 104, 288-299.
- Fischhoff, B., P. Slovic and S. Lichtenstein (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception & Performance*, 3, 552-564.
- Fiske, D. W. (1982). Convergent-discriminant validation in measurements and research strategies. In D. Brinbirg & L. H. Kidder (Eds.), *Forms of validity in research* (pp. 77-92). San Francisco, CA: Jossey-Bass.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 357-411). Boston, MA: McGraw-Hill.

- Fossett, M. A., & Kiecolt, K. J. (1989). The relative size of minority populations and white racial attitudes. *Social Science Quarterly*, 70, 820-835.
- Friedrich, J. (1996). On seeing oneself as less self-serving than others: The ultimate self-serving bias? *Teaching of Psychology*, 23, 107-109.
- Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception*. San Diego, CA: Academic Press.
- Funder, D. C. (2001). Personality. *Annual Review of Psychology*, 52, 197-223.
- Fussell, S. R., & Krauss, R. M. (1991). Accuracy and bias in estimates of others' knowledge. *European Journal of Social Psychology*, 21, 445-454.
- Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology*, 62, 378-391.
- Gaertner, S. L., Dovidio, J. F., Banker, B. S., Houlette, M., Johnson, K. M., & McGlynn, E. A. (2000). Reducing intergroup conflict: From superordinate goals to decategorization, recategorization, and mutual differentiation. *Group Dynamics: Theory, Research and Practice*, 4, 98-114.
- Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, 78, 708-724.
- Giles, M. W. (1977). Percent black and racial hostility: an old assumption reexamined, *Social Science Quarterly*, 58, 412-417.
- Graham, J. A., & Cohen, R. (1997). Race and sex as factors in children's sociometric ratings and friendship choices. *Social Development*, 6, 355-372.

- Green, J. D., & Sedikides, C. (2004). Retrieval selectivity in the processing of self-referent information: Testing the boundaries of self-protection. *Self and Identity*, 3, 69-80.
- Green, J. D., Sedikides, C., & Gregg, A. P. (2008). Forgotten but not gone: The recall and recognition of self-threatening memories. *Journal of Experimental Social Psychology*, 44, 547-561.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4-27.
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, 109, 3-25.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464 – 1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85, 197 – 216.
- Haddock, G., Rothman, A. J., & Reber, R., & Schwarz, N. (1999). Forming judgments of attitude certainty, intensity, and importance: The role of subjective experience. *Personality and Social Psychology Bulletin*, 25, 771-782.
- Hallinan, M. T., & Teixeira, R. A. (1987). Students' interracial friendships: Individual characteristics, structural effects and racial differences. *American Journal of Education*, 95, 563-583.

- Hamburger, Y. A., & Ben-Artzi, E. (2000). The relationship between extroversion and neuroticism and the different uses of the Internet. *Computers in Human Behaviour, 16*, 441-449.
- Hartup, W. W., & Stevens, N. (1997). Friendships and adaptation in the life course. *Psychological Bulletin, 121*, 355–370.
- Henry, B., Moffitt, T. E., Caspi, A., Langley, J., & Silva, P. A. (1994). On the “remembrance of things past”: A longitudinal evaluation of the retrospective method. *Psychological Assessment, 6*, 92-101.
- Hewstone, M. (1996). Contact and categorization: Social psychological interventions to change intergroup relations. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and Stereotyping* (pp. 323–368). New York, NY: The Guilford Press.
- Hewstone, M. & Brown, R. (1986). Contact is not enough: An intergroup perspective on the 'contact hypothesis'. In M. Hewstone & R. Brown (Ed.), *Contact and conflict in intergroup encounters* (pp. 1-44). Oxford, UK: Blackwell.
- Hill, L. G., & Betz, D. L. (2005). Revisiting the retrospective pretest. *American Journal of Evaluation, 26*, 501-517.
- Hodges, S. D., Johnsen, A. T., & Scott, N. S., (2002). You’re like me, no matter what you say: Self projection in self-other comparisons. *Psychologica Belgica, 42*, 107-112.
- Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self stereotyping and the salience of social categories. *British Journal of Social Psychology, 26*, 325-340.
- Hood, M. V., & Morris, I. L. (1997). Amigo o enemigo?: Context, attitudes, and Anglo public opinion toward immigration. *Social Science Quarterly, 78*, 309-323.

- Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology, 75*, 581-595.
- Howard G. (1994). Why people say nasty things about self-reports. *Journal of Organizational Behaviour 15*, 399-404.
- Howes, C., & Wu, F. (1990). Peer interactions and friendships in an ethnically diverse school setting. *Child Development, 61*, 537-541.
- Hoyt, W. T., & Kerns, M. D. (1999). Magnitude and moderators of bias in observer ratings: A meta-analysis. *Psychological Methods, 4*, 403-424.
- Ickes, W. J., & Simpson, J. A. (1997). Managing empathic accuracy in close relationship. In W. Ickes (Ed.), *Empathic accuracy* (pp. 218-250). New York: Guilford Press.
- Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 774-788.
- Keeter, S. & Samaranayake, N. (2007, March 9). Can you trust what polls say about Obama's electoral prospects? [Electronic version] *Pew Research Center for the People and the Press*. Retrieved February 9, 2009, from <http://pewresearch.org/pubs/408/can-you-trustwhat-polls-say-about-obamas-electoral-prospects>
- Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: ease of generation as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32*, 1-24.
- Kenny, D. A. (1988). The analysis of data from two-person relationships. In S. Duck (Ed.), *Handbook of interpersonal relationships* (pp. 57-77). London, UK: Wiley.
- Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York, NY: Guilford.

- Kenny, D. A. (1998), *SOREMO Version V.2*. Unpublished manuscript, University of Connecticut, Storrs, CT.
- Kenny, D. A. (2000). *PERSON: A general model for understanding interpersonal perception*. Unpublished manuscript, University of Connecticut.
- Kenny, D. A. & Acitelli, L. K. (2001). Accuracy and bias in the perception of the partner in a close relationship. *Journal of Personality and Social Psychology*, 80, 439-448
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford Press.
- Kenny, D. A., & La Voie, L. (1984). The social relations model. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 18, pp. 142-182). Orlando, FL: Academic Press.
- Khmelkov, V. T., & Hallinan, M. T. (1999). Organizational effects of race relations in schools. *Journal of Social Issues*, 55, 627-645.
- Kiesner, J., Maass, A., Cadinu, M., & Vallese, I. (2003). Risk factors for ethnic prejudice during early adolescence. *Social Development*, 12, 288-308.
- Knudson, D. O., Pope, H., & Irish, D. P. (1967). Response differences to questions on sexual standards: An interview questionnaire comparison. *Public Opinion Quarterly*, 3, 290-297.
- Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality*, 64, 311-337.
- Konstabel, K., (2006). *The structure and validity of self- and peer-reported personality traits*. Doctoral dissertation, University of Tartu, Tartu, Estonia: Tartu University Press.

- Krahé, B. (2008). Aggression. In M. Hewstone, W. Stroebe, & K. Jonas (Eds.), *Introduction to social psychology* (4th ed., pp. 156-175). Oxford, UK: Blackwell.
- Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukopadhyay, T., & Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being? *American Psychologist*, 53, 1017-1031.
- Krueger, J. (2000). The projective perception of the social world: A building block of social comparison processes. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 323-351). New York, NY: Plenum/Kluwer.
- Krueger, J. (2007). From social projection to social behaviour. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology* (Vol. 18, pp. 1-35). Hove, UK: Psychology Press.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77, 1121-1134.
- Kunda, Z., Fong, G. T., Sanitioso, R., & Reber, E. (1993). Directional questions direct self-conceptions. *Journal of Experimental Social Psychology*, 29, 63-86.
- Lambert, A. J., Payne, B. K., Jacoby, L. L., Shaffer, L. M., Chasteen, A. L., & Khan, S. R. (2003). Stereotypes as dominant responses: On the "social facilitation" of prejudice in anticipated public contexts. *Journal of Personality and Social Psychology*, 84, 277-295.
- Larwood, L. (1978). Swine flu: A field study of self-serving biases. *Journal of Applied Social Psychology*, 8, 283-289.
- Larwood, L., & Whittaker, W. (1977). Managerial myopia: Self-serving biases in organizational planning. *Journal of Applied Psychology*, 62, 194-198.

- Laurenceau, J. P., Barrett, L. F., & Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: The importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *Journal of Personality and Social Psychology*, 74, 1238–1251.
- Lee, L., Howes, C., & Chamberlain, B. (2007). Ethnic heterogeneity of social networks and cross-ethnic friendships of elementary school boys and girls. *Merrill-Palmer Quarterly*, 53, 325-346.
- Lefkowitz, M. M., Eron, L. D., Walder, L. O., & Huesmann, L. R. (1977). *Growing up to be violent*. New York, NY: Pergamon.
- Lenton, A. P., Bryan, A., Hastie, R., & Fischer, O. (2007). We want the same thing: Projection in judgements of sexual intent. *Personality and Social Psychology Bulletin*, 33, 975-988.
- Lentz, T. F. (1938). Acquiescence as a factor in the measurement of personality. *Psychological Bulletin*, 35, 659.
- Levinger, G., & Breedlove, J. (1966). Interpersonal attraction and agreement: A study of marriage partners. *Journal of Personality and Social Psychology*, 3, 367–372.
- Liebkind, K. & McAlister, A. (1999). Extended contact through peer modelling to promote tolerance in Finland. *European Journal of Social Psychology*, 29, 765-800.
- Linton, M. 1982. Transformations of memory in everyday life. In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 77-91). San Francisco, CA: W. H. Freeman.
- Loftus, E. F., & Wagenaar, W. A. (1988). Lawyers' predictions of success. *Jurimetrics Journal*, 29, 437–453.

- Mabe, P. A. III, & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67, 280–286.
- Mackenzie, B. K. (1948). The importance of contact in determining attitudes towards Negroes. *Journal of Abnormal and Social Psychology*, 43, 417-441.
- Malloy, T. E., & Kenny, D. A. (1986). The social relations model: An integrative method for the study of personality. *Journal of Personality*, 54, 101-124.
- Marks, G., & Miller, N. 1987. Ten years of research on the false-consensus effect: An empirical and theoretical review. *Psychological Bulletin* 102:72–90.
- Marottoli, R. A., & Richardson, E. D. (1998). Confidence in, and self-rating of, driving ability among older drivers. *Accident Analysis and Prevention*, 30, 331-336.
- Marteau, T. M., Johnston, M., Wynne, G., & Evans, T. R. (1989). Cognitive factors in the explanation of the mismatch between confidence and competence in performing basic life support. *Psychology and Health*, 3, 173–182.
- McCauley, C., Plummer, M., Moskalenko, S., & Mordkoff, J. T. (2001). The exposure index: A measure of intergroup contact. *Peace and Conflict: Journal of Peace Psychology*, 7, 321-336.
- McConnell, A. R., & Leibold, J. M. (2001). Relations among the implicit association test, discriminatory behaviour, and explicit measures of racial attitude. *Journal of Experimental Social Psychology*, 37, 435-442.
- McCrae, R. R. (1993). Agreement of personality profiles across observers. *Multivariate Behavioural Research*, 28, 25-40.
- McCrae, R. R., & Costa, P T., Jr. (1989). Different points of view: Self-reports and ratings in the assessment of personality. In J. P. Forgas & J. M. Innes (Eds.), *Recent advances in social psychology: An international perspective* (pp. 429-439). North-Holland, The Netherlands: Elsevier Science.

- McCrae, R. R., & Weiss, A. (2007). Observer ratings of personality. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 259-272). London, UK: The Guildford Press.
- McKenna, K. Y. A., & Bargh, J. A. (1999). Causes and consequences of social interaction on the Internet: A conceptual framework. *Media Psychology, 1*, 249-269.
- McKenna, K. Y. A., & Bargh, J. A. (2000). Plan 9 from cyberspace: The implications of the Internet for personality and social psychology. *Personality and Social Psychology Review, 4*, 57-75.
- McKenna, K. Y. A., Buffardi, L. E. & Seidman, G. (2005). Self presentation to friends and strangers online. In K. H. Renner, A. Schutz, & F. Machilek (Eds.), *Internet and personality* (pp. 176-189). Göttingen, Germany: Hogrefe & Huber Publishers.
- McKenna, K. Y. A., & Green, A. S. (2002). Virtual group dynamics. *Group Dynamics: Theory, Research and Practice, 6*, 116-127.
- McKenna, K. Y. A., Green, A. S., & Gleason, M. (2002). Relationship formation on the Internet: What's the big attraction? *Journal of Social Issues, 58*, 9-31.
- McSmith, A. (2007, September 20). The search for racial equality: 31 years of fighting racism [Electronic version]. *The Independent*. Retrieved February 9, 2009, from <http://www.independent.co.uk/news/uk/this-britain/the-search-for-racial-equality-31-years-of-fighting-racisim-402938.html>
- Meacham, J. (2008, November 17). The age of Obama [Electronic version]. *Newsweek*, p. 18-20.
- Meer, B., & Freedman, E. (1966). The impact of Negro neighbors on white home owners. *Social Forces, 45*, 11-19.

- Merida, K. (2008, May 13). Racist incidents give some Obama campaigners pause [Electronic version]. *Washington Post*.
- Middleton, W., Harris, P., & Surman, M. (1996). Give 'em enough rope: Perception of health and safety risks in bungee jumpers. *Journal of Social and Clinical Psychology, 15*, 68–79.
- Minard, R. D. (1952). Race relations in the Pocahontas coal field. *Journal of Social Issues, 8*, 29-44.
- Moaz, I. (2000). Power relations in intergroup encounters. *International Journal of Intercultural Relations, 24*, 259-277.
- Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996). The benefits of positive illusions: Idealization and construction of satisfaction in close relationships. *Journal of Personality and Social Psychology, 70*, 79-98.
- Murstein, B. I., & Beck, G. D. (1972). Person perception, marriage adjustment, and social desirability. *Journal of Consulting and Clinical Psychology, 39*, 396–403.
- Nagourney, A. (2006, December 10). The pattern may change, if ... [Electronic version]. *New York Times*.
- Neale, M. A., & Bazerman, M. H. (1983). The role of perspective taking ability in negotiating under different forms of arbitration. *Industrial and Labor Relations Review, 36*, 378-388.
- Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology, 15*, 263-280.
- Nickerson, R. S., Baddeley, A., & Freeman, B. (1987). Are people's estimates of what other people know influenced by what they themselves know? *Acta Psychologica, 64*, 245-259.

- Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology* 4, 681–691.
- Nunnally, J. C. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.
- Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *Journal of Finance*, 8, 1887–1934.
- Office for National Statistics. (2001). *Population size of minority ethnic groups*. Retrieved 20 February, 2009, from <http://www.statistics.gov.uk/cci/nugget.asp?id=273>
- Office for National Statistics. (2001, April). *Census*.
- Oksam, J., Kingma, J., & Klasen, H. J. (2000). Clinicians' recognition of 10 different types of distal radial fractures. *Perceptual and Motor Skills*, 91, 917–924.
- Ones, D. S., Viswesvaran, C., & Reiss A. D. (1995). The role of social desirability in personality testing for personnel selection: The Red Herring. *Journal of Applied Psychology*, 81, 660-679.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Clinical Psychology*, 29, 261–265.
- Oskamp, S. & Jones, J. M. (2000). Promising practices in reducing prejudice: A report from the President's initiative on race. In S. Oskamp (Ed.), *Reducing Prejudice and Discrimination* (pp. 319-334). Hillsdale, NJ: Lawrence Erlbaum.
- Paolini, S., Hewstone, M., Cairns, E., & Voci, A. (2004). Effects of direct and indirect cross-group friendships on judgments of Catholics and Protestants in Northern Ireland: The mediating role of an anxiety-reduction mechanism. *Personality and Social Psychology Bulletin*, 30, 770-786.
- Paulhus, D. L (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.

- Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaire* (pp. 143-165). New York, NY: Springer-Verlag.
- Paulhus, D. L. (1988). *Assessing self-deception and impression management in self-reports: The Balanced Inventory of Desirable Responding*. Unpublished manual, University of British Columbia, Vancouver Canada.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman, (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, 66, 1025-1060.
- Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, 49, 65–85.
- Pettigrew, T. F. & Alston, D. A. (1988). *Tom Bradley's Campaign for Governor*. Washington, DC: Joint Center for Political and Economic Studies.
- Pettigrew, T. F., & Tropp, L. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90, 751-783.
- Petty, R. E., Fazio, R. H., & Briñol, P. (2009). The new implicit measures: An overview. In R. E. Petty, R. H. Fazio & P. Briñol (Eds.). *Attitudes: Insights from the new implicit measures*. New York: Psychology Press.
- Phillips, D. L. and Clancy, K. J. (1972). Some Effects of 'Social Desirability' in Survey Studies. *American Journal of Sociology*, 77, 921-938.

- Piedmont, R. L., McCrae, R. R., Riemann, R., & Angleitner, A. (2000). On the invalidity of validity scales: Evidence from self-reports and observer ratings in volunteer samples. *Journal of Personality and Social Psychology*, 78, 582-593.
- Plant, E. A. (2004). Responses to interracial interactions over time. *Personality and Social Psychology Bulletin*, 30, 1458-1471.
- Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from each other? *Behavioural Brain Science*, 10, 1-60.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 879-903.
- Potok, M. (2008, September 12). Neo-Nazi magazine to feature Obama assassination cover [Electronic version]. *Hatewatch* (Southern Poverty Law Center).
- Price, V., & Zaller, J. (1993). Who gets the news? Alternative measures of news reception and their implications for public opinion research. *Public Opinion Quarterly*, 57, 133-164.
- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28, 369-381.
- Pyszczynski, T. & Greenberg, J. (1987). Self-regulatory perseverance and the depressive self-focusing style: A self-awareness theory of the development and maintenance of reactive depression. *Psychological Bulletin*, 102, 122-138.
- Reis, H. T. & Judd, C. M. (2000). *Handbook of research methods in social and personality psychology*. Cambridge University Press.
- Riggio, R. E., Widaman, K. F., & Friedman, H. S. (1985). Actual and perceived emotional sending and personality correlates. *Journal of Nonverbal Behavior*, 9, 69-83.

- Risucci, D. A., Torolani, A. J., & Ward, R. J. (1989). Ratings of surgical residents by self, supervisors and peers. *Surgical Gynecology and Obstetrics*, 169, 519–526.
- Robins, R. W., Hendin, H. M. & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the rosenberg self-esteem scale. *Personality and Social Psychology Bulletin*, 27, 151-161.
- Robins, R. W., Mendelsohn, G. A., Connell, J. B., & Kwan, V. (2004). Do people agree about the causes of behavior? A social relations analysis of behavior ratings and causal attributions. *Journal of Personality and Social Psychology*, 86, 334-344.
- Robinson, J. W., & Preston, J. D. (1976). Equal status contact and modification of racial prejudice: A re-examination of the contact hypothesis. *Social Forces*, 54, 911-924.
- Rosenthal, A. (1989, November 9). The 1989 elections: Predicting the outcome; broad disparities in votes and polls raising questions [Electronic version]. *New York Times*.
- Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13, 279-301.
- Rothbart, M., & John, O. P. (1985). Social categorization and behavioral episodes: A cognitive analysis of the effect of intergroup contact. *Journal of Social Issues*, 41, 81-104.
- Rowe, D. C., & Osgood, D. W. (1984). Heredity and sociological theories of delinquency: A reconsideration. *American Sociological Review*, 49, 526-540.

- Rubin, K. H., Bukowski, W. M., & Parker, J. G. (1998). Peer interactions, relationships, and groups. In W. Damon (Series Ed.) & N. Eisenberg (Ed.), *Handbook of child psychology: Vol 3. Social, emotional, and personality development* (5th ed., pp. 619–700). New York, NY: Wiley.
- Rushton, J. P. (1989). Genetic similarity, human altruism, and group selection. *Behavioral and Brain Sciences*, 12, 503-559.
- Rutter, D. R., Quine, L., & Albery, I. P. (1998). Perceptions of risk in motorcyclists: Unrealistic optimism, relative realism and predictions of behaviour. *British Journal of Psychology*, 89, 681–696.
- Sagar, H. A., & Schofield, J. W. (1980). Racial and behavioral cues in black and white children's perceptions of ambiguously aggressive acts. *Journal of Personality and Social Psychology*, 19, 590-598.
- Salancik, G. R., & Conway, M. (1975). Attitude inferences from salient and relevant cognitive content about behavior. *Journal of Personality & Social Psychology*, 32, 829-840.
- Scheier, M. E, & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*, 4, 219-247.
- Schofield, J. W. (1979). The impact of positively structured contact on intergroup behavior. *Social Psychology Quarterly*, 42, 280-284.
- Schofield, J. W. (1995). Review of research on school desegregation's impact on elementary and secondary school students. In J. A. Banks & C. A. McGee Banks (Eds.), *Handbook of research on multicultural education* (pp. 597–616). New York, NY: Macmillan.

- Schofield, J. W., & Sagar, H. A. (1977). Peer interaction patterns in an integrated middle school. *Sociometry*, 40, 130-138.
- Schuman, H., Stech, C., & Bobo, L. (1985). *Racial attitudes in America: Trends and interpretations*. Cambridge, MA: Harvard University Press.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54, 93-105.
- Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217-245). Hillsdale, NJ: Erlbaum.
- Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, 61, 195-202.
- Schwarz, N., & Strack, F. (1999). Evaluating one's life: A judgement model of subjective well-being. In N. Schwarz (Ed.), *Subjective well-being* (pp. 27-47). Oxford, UK: Pergamon Press.
- Schwarz, N., & Sudman, S. (Eds.). (1992). *Context effects in social and psychological research*. New York, NY: Springer Verlag.
- Sedikides, C., & Green, J. D. (2000). On the self-protective nature of inconsistency/negativity management: Using the person memory paradigm to examine self-referent memory. *Journal of Personality and Social Psychology*, 79, 906-922.
- Sedikides, C., & Green, J. D. (2004). What I don't recall can't hurt me: Information negativity versus information inconsistency as determinants of memorial self-defense. *Social Cognition*, 22, 4-29.

- Sedikides, C., & Strube, M. J. (1997). Self-evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 209–269). New York, NY: Academic Press.
- Sherif, M. (1966). *Group conflict and cooperation: Their social psychology*. London, UK: Routledge & Kegan Paul.
- Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *Intergroup conflict and cooperation. The Robbers Cave experiment*. Norman, OK: University of Oklahoma Book Exchange.
- Sherif, M., & Sherif, C. W. (1953). *Groups in harmony and tension: An integration of studies on inter-group relations*. New York, NY: Harper & Row.
- Shrum, W., Cheek, N., Jr., & Hunter, S. (1988). Friendship in school: Gender and racial homophily. *Sociology of Education*, 61, 227–239.
- Sigall, F. L., & Page, R. (1971). Current stereotypes: A little fading, a little faking. *Journal of Personality and Social Psychology*, 18, 247–255.
- Sillars, A. L., & Scott, M. D. (1983). Interpersonal perception between intimates: An integrative review. *Human Communication Research*, 10, 153–176.
- Sims, V. F., & Patrick, J. R. (1936). Attitude towards the negro of northern and southern college students. *Journal of Social Psychology*, 7, 192–204.
- Sinclair, S., Dunn, E., & Lowery, B. (2005). The relationship between parental racial attitudes and children's implicit prejudice. *Journal of Experimental Social Psychology*, 41, 283–289.
- Singer, H. A. (1948). The veteran and race relations. *Journal of Educational Sociology*, 21, 397–408.

- Skowronski, J. J., Betz, A. L., Thompson, C. P., & Shannon, L. (1991). Social memory in everyday life: The recall of self-events and other-events. *Journal of Personality and Social Psychology*, 60, 831-843.
- Slavin, R. E. (1978). Student teams and achievement divisions. *Journal of Research and Development in Education*, 12, 381-387.
- Smith, P. B., & Bond, M. H. (1994). *Social psychology across cultures: Analysis and perspectives*. Boston, MA: Allyn & Bacon.
- Snyder, M. & Cantor, N. (1979). Testing hypotheses about other people: The use of historical knowledge. *Journal of Experimental Social Psychology*, 15, 330-342.
- Spears, R., Postmes, T., Lea, M., & Wolbert, A. (2002). When are net effects gross products? The power of influence and the influence of power in computer-mediated communication. *Journal of Social Issues*, 58, 91-107.
- Stephan, W. G. (1999). *Reducing prejudice and stereotyping in schools*. New York, NY: Teachers College Press.
- Stephan, W. G., & Rosenfield, D. (1978). Effects of desegregation on racial attitudes. *Journal of Personality and Social Psychology*, 36, 795-804.
- Stouffer, S. A., Lumsdaine, A. A., Lumsdaine, M. H., Williams, R. M., Jr., Smith, M. B., Janis, I. L., Star, S. S., & Cottrell, L. S., Jr. (1949b). *The American soldier: Combat and its aftermath, Volume II*. Princeton, NJ: Princeton University Press.
- Strack, E, Schwarz, N., & Gschneidinger, E. (1985). Happiness and reminiscing: The role of time perspective, affect, and mode of thinking. *Journal of Personality and Social Psychology*, 47, 1460-1469.
- Sudman, S. & Bradburn, N. (1974). *Response effects on surveys: A review and synthesis*. Chicago, IL: Aldine.

- Sussman, B. (1985, November 28). Hidden racial attitudes distorted Virginia polls [Electronic Version]. *Washington Post*.
- Tajfel, H. (1970). Experiments in intergroup, discrimination. *Scientific American*, 223, 96-102.
- Tajfel, H. (1978). *Differentiation between social groups: studies in the social psychology of intergroup relations*. London, UK: Academic Press.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.
- Tan, A., Tan, G., Avdeyeva, T., Crandall, H., Fukushi, Y., Nyandwi, A., Chin, H. Y., & Wu, C. G. (2001). Changing negative racial stereotypes: The influence of normative peer information. *Howard Journal of Communications*, 12, 171-180.
- Taylor, D. M., Dubé, L., & Bellerose, J. (1986). Intergroup contact in Quebec: Myth or reality? In M. Hewstone, & R. Brown (Eds.), *Contact and conflict in intergroup encounters* (pp. 107–118). Oxford, UK: Basil Blackwell.
- Taylor, S. E., & Brown, J. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193–210.
- Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, 36, 778-793.
- Tellegen, A. (1988). The analysis of consistency in personality assessment. *Journal of Personality*, 56, 621-663.
- Terkildsen, N. (1993). When white voters evaluate black candidates: The processing implications of candidate skin colour, prejudice, and self-monitoring. *American Journal of Political Science*, 37, 1032-1053.

- Tracey, J. M., Arroll, B., Richmond, D. E., & Barham, P. M. (1997). The validity of general practitioners' self assessment of knowledge: Cross sectional study. *British Journal of Medicine*, 315, 1426– 1428.
- Tropp, L. R., & Pettigrew, T. F. (2005). Relationship between intergroup contact and prejudice among minority and majority status groups. *Psychological Science*, 16, 951-957.
- Turner, R. (2005). *Reducing explicit and implicit prejudice among Asians and whites in the UK: Processes and consequences of direct and extended intergroup contact*. Unpublished doctoral thesis, Oxford University.
- Turner, R. N., Hewstone, M., & Voci, A. (2007). Reducing explicit and implicit outgroup prejudice via direct and extended contact: The mediating role of self-disclosure and intergroup anxiety. *Journal of Personality and Social Psychology*, 93, 369-388
- Turner, R. N., Hewstone, M., Voci, A., Paolini, S., & Christ, O. (2008). Reducing prejudice via direct and extended cross-group friendship. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology* (Vol. 18, pp. 212-255). Hove, UK: Psychology Press.
- Turner, R. N., Hewstone, M., Voci, A., & Vonofakou, C. (2008). A test of the extended intergroup contact hypothesis: The mediating role of intergroup anxiety, perceived ingroup and outgroup norms, and inclusion of the outgroup in the self. *Journal of Personality and Social Psychology*, 95, 843-860.
- Undergraduate Admissions Office. (2007, December). Oxford University undergraduate admissions statistics for entry 2007 [electronic version]. *Oxford University Gazette, Supplement 2, Vol. 138*. Retrieved March 5th, 2009, from <http://studyatoxford.files.wordpress.com/2008/02/adstats.pdf>

- U.S. Census Bureau. (2006). *American community survey*. Retrieved 20 February, 2009, from http://factfinder.census.gov/servlet/DTTable?_bm=y&-context=dt&-ds_name=ACS_2006_EST_G00_&-CONTEXT=dt&-mt_name=ACS_2006_EST_G2000_B02001&-tree_id=306&-redoLog=false&-currentselections=ACS_2006_EST_G2
- U.S. Census Bureau. (2007). *Population Estimates*. Retrieved 20 February, 2009, from http://factfinder.census.gov/servlet/QTTable?_bm=y&-geo_id=01000US&-qr_name=PEP_2007_EST_DP1&-qr_name=PEP_2007_EST_DP1PR&-parsed=true&-ds_name=PEP_2007_EST&-_lang=en&-format=&-CONTEXT=qt
- U.S. Department of Commerce (2004). *A nation online: Entering the broadband age*. Retrieved 9 February, 2009, from <http://www.ntia.doc.gov/reports/anol/nationonlinebroadband04.pdf>
- Vazire, S. (2006). Informant reports: a cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, 40, 472-481.
- Vogt, D. S., Colvin, C. R. (1999). *The good judge of personality: Gender differences, personality correlates, and Chronbackian "artifacts."* Northeastern University, unpublished.
- Vonofakou, C. (2008). *The effects of direct and extended cross-group friendships on meta-attitudinal strength and attitude accessibility*. Unpublished doctoral thesis, University of Oxford, UK.
- Wagenaar, W. A. (1986). My memory: A study of autobiographical memory over six years. *Cognitive Psychology*, 18, 225-252.
- Wagner, U., Hewstone, M., & Machleit, U. (1989). Contact and prejudice between Germans and Turks: A correlational study. *Human Relations*, 42, 561-574.

- Wagner, U., & Machleit, U. (1986). 'Gastarbeiter' in the Federal Republic of Germany: Contact between Germans and migrant populations. In R. Brown & M. Hewstone (Eds.), *Contact and conflict in intergroup encounters* (pp.59-78). Cambridge, MA: Basil Blackwell.
- Walker, I., & Crogan, M. (1998). Academic performance, prejudice, and the Jigsaw classroom: New pieces to the puzzle. *Journal of Community and Applied Social Psychology*, 8, 381-393.
- Walker, W. R., Skowronski, J. J., & Thompson, C. P. (2003). Life is pleasant—and memory helps to keep it that way! *Review of General Psychology*, 7, 203-210.
- Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23, 3-43.
- Weber, R. & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, 45, 961-977.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 58, 806–820.
- West, R., French, D., Kemp, R., & Elander, J. (1993). Direct observation of driving, self reports of driver behaviour, and accident involvement. *Ergonomics*, 36, 557-67.
- Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.
- Wilder, D. A. (1984). Intergroup contact: The typical member and the exception to the rule. *Journal of Experimental Social Psychology*, 20, 177-194.

- Wilder, D. A. (1993). The role of anxiety in facilitating stereotypic judgments of out-group behavior. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in group perception* (pp. 87-109). San Diego, CA: Academic Press.
- Williams, R. M. (1964). *Strangers next door: Ethnic relations in American communities*. Englewood Cliffs, NJ: Prentice-Hall.
- Wilson, T. D., & Dunn, E. W. (2004). Self-knowledge: Its limits, value, and potential for improvement. *Annual Review of Psychology*, 55, 493-518.
- Wilson, T. D., & LaFleur, S. J. (1995). Knowing what you'll do: Effects of analyzing reasons on self-prediction. *Journal of Personality and Social Psychology*, 68, 21-35.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101-126.
- Winters, K. C., & Neale, J. (1985). Mania and low self-esteem. *Journal of Abnormal Psychology*, 94, 282-290.
- Wiseman, F. (1972), Methodological considerations in segmentation studies. In F. C. Allvine (Ed.), *Combined proceedings* (pp. 306-311). Chicago, IL: American Marketing Association.
- Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, 72, 262-274.
- Wojciszke, B. (1994). Multiple meanings of behaviour: Construing actions in terms of competence or morality. *Journal of Personality and Social Psychology*, 67, 222-232.

- Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, 73, 73-90.
- Yourish, K. (2008, August 22). For Black politicians, a rocky road but a steady climb [Electronic version]. *The Washington Post*.
- Zajonc, R. B. (1965). Social facilitation. *Science*, 149, 269-274.

APPENDICES

Appendix 1:

Full matrix of correlations among target effects and self-ratings – Asian and gay contact and liking (Study 7)

		Target						Self					
		Asian Contact (H)	Asian Contact (U)	Asian Therm.	Gay Contact (H)	Gay Contact (U)	Gay Therm.	Asian Contact (H)	Asian Contact (U)	Asian Therm.	Gay Contact (H)	Gay Contact (U)	
Target	Asian Contact (U)	.42* .88* .38* -.09 .00 -.27 -.25 .06 -.14 -.12 -.17 .17 .58* .56*											
	Asian Therm.												
	Gay Contact (H)												
	Gay Contact (U)												
	Gay Therm.												
Self	Asian Contact (H)	.54* .10 .34* -.16 -.24						.08 .34 .29* -.01 .17 .12 -.10 .18 .09 .01 .15 .18 -.07 .67* .14 .08					
	Asian Contact (U)	-.09 .28* -.09 .17 .05											
	Asian Therm.	.38* .20 .46* -.34* -.22											
	Gay Contact (H)	.18 .09 .12 .60* .15											
	Gary Contact (U)	-.04 -.10 .12 .17 .32*											
	Gay Therm.	.20 -.07 .09 -.05 -.08											

Appendix 2:

Facebook Friends Page screenprint

facebook

Profile edit Friends ▾ Networks ▾ Inbox

home account privacy logou

Search ▾



 Photos

 Notes

 Groups

 Events

 Posted Items

 Marketplace

Friend List Find Friends Status Updates Social Timeline

Show: Recently Updated ▾

 Search within friends

You have 49 friends with recently updated profiles.



Name: **Bella Lat**
Network: London
Details: You worked at LU Cinema from 2002 to 2003.
[edit details]
Status: Bella is cable laying.
Updates: Educational Info
Profile updated on Wednesday

Send Message
Poke Bella!
View Friends
View Photos
Remove Friend



Name: **Catherine Oakley**
Network: Oxford '07
Details: You have lived in OB1 since 2006.
You have been members of Hertford College Chapel Choir since 2006.
[edit details]
Status: Catherine is DONE WITH FINALS!
Updates: Profile Picture, Interests and About Me
Profile updated on Wednesday

Send Message
Poke Catherine!
View Friends
View Photos
Remove Friend



Name: **Peter Hampson**
Network: Manchester
Details: You went to college with Peter.
[edit details]
Updates: Profile Picture
Profile updated on Wednesday

Send Message
Poke Peter!
View Friends
View Photos
Remove Friend



Name: **Kate Louw**
Networks: Cape Town Grad Student '07
South Africa
Details: You traveled to scotland in winter 2006.
You know Kate through Abby Loebenberg.
[edit details]

Send Message
Poke Kate!
View Friends
View Photos
Remove Friend

Appendix 3:

Outgroup contact and attitude questions

Quantity of contact with outgroup – Included in Studies 2-6, 8, and 9

How often do you meet [outgroup] in your everyday life at [university/home]?

1	2	3	4	5	6	7
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Quite a bit</i>	<i>Often</i>	<i>Very often</i>	<i>All the time</i>

How often do you *talk with* people who are [outgroup]?

1	2	3	4	5	6	7
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Quite a bit</i>	<i>Often</i>	<i>Very often</i>	<i>All the time</i>

Quality of contact with outgroup – Included in Studies 6 and 8

Are your interactions with [outgroup] generally:

1	2	3	4	5	6	7
<i>Cooperative</i>						<i>Competitive</i>
1	2	3	4	5	6	7
<i>Superficial</i>						<i>Intimate</i>

Direct cross-group friendship with outgroup - Included in all studies

About how many of your *friends* are [outgroup]?

1	2	3	4	5	6	7
<i>None</i>	<i>A few</i>	<i>Less than half</i>	<i>About half</i>	<i>More than half</i>	<i>Most</i>	<i>All</i>

How often do you spend time with friends who are [outgroup]?

1	2	3	4	5	6	7
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Quite a bit</i>	<i>Often</i>	<i>Very often</i>	<i>All the time</i>

Indirect cross-group friendship with outgroup - Included in Studies 2-6, 8, and 9

About how many of your [ingroup]friends do you think have friends who are [outgroup]?

1	2	3	4	5	6	7
<i>None</i>	<i>A few</i>	<i>Less than half</i>	<i>About half</i>	<i>More than half</i>	<i>Most</i>	<i>All</i>

About how many people in your immediate family have friends who are [outgroup]?

1	2	3	4	5	6	7
<i>None</i>	<i>A few</i>	<i>Less than half</i>	<i>About half</i>	<i>More than half</i>	<i>Most</i>	<i>All</i>

Attitude towards outgroup – Subsection of these included in all studies

Please use this ‘feeling thermometer’ to rate how warm or cold you feel toward [outgroup] in general:

10°	20°	30°	40°	50°	60°	70°	80°	90°	100°
<i>Cold</i>									<i>Warm</i>

Based on your experience rate the extent to which you have each of the following feelings about [outgroup] in general. (Please circle one number on each scale).

Do you feel.....

1	2	3	4	5	6	7
<i>Negative</i>						<i>Positive</i>

1	2	3	4	5	6	7
<i>Friendly</i>						<i>Hostile</i>

1	2	3	4	5	6	7
<i>Suspicious</i>						<i>Trusting</i>

1	2	3	4	5	6	7
<i>Respect</i>						<i>Contempt</i>

1	2	3	4	5	6	7
<i>Admiration</i>						<i>Disgust</i>

Appendix 4:

Closeness measures

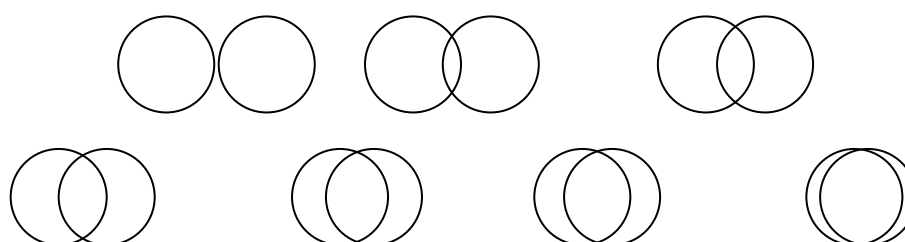
Studies 3, 4 and 5

How close do you feel to X?

1	2	3	4	5	6	7
<i>Not at all close</i>						<i>Very Close</i>

Study 6¹⁶

In the pictures of the circle-pairs below, please imagine that one circle represents *you* and one represents *your friend, [A/B/C/D]*. Please circle the picture that best describes your relationship with each other; how close you are, how well you know each other.



Studies 6 and 7

How long have you known [A/B/C/D]? _____ months (approx)

How often do you see/spend time with [A/B/C/D]?

1	2	3	4	5	6	7
<i>Very rarely</i> <i>(eg once a month)</i>			<i>A fair amount</i> <i>(eg twice a week)</i>			<i>Very frequently</i> <i>(eg several times a day)</i>

¹⁶ Taken from the inclusion of the other in the self scale (Aron et al., 1992).

Appendix 5:

Self-disclosure measures used in Studies 3, 4, 5 and 7

Imagine you had a problem that was worrying you. How likely is it that you would tell X?

1	2	3	4	5	6	7
Not at all likely						Very likely

Imagine you had an important secret. How likely is it that you would choose X to tell?

1	2	3	4	5	6	7
Not at all likely						Very likely

If X was worried about something, how likely is it that they would tell you about the problem?

1	2	3	4	5	6	7
Not at all likely						Very likely

If X had an exciting secret, how likely is it that they would tell you about the secret?

1	2	3	4	5	6	7
Not at all likely						Very likely

How often do you talk about how you are feeling to X?

1	2	3	4	5	6	7
<i>Never</i>		<i>Sometimes</i>			<i>All the time</i>	

How often do you discuss intimate or personal issues with X?

1	2	3	4	5	6	7
<i>Never</i>		<i>Sometimes</i>			<i>All the time</i>	

How often does X talk to you about how they are feeling?

1	2	3	4	5	6	7
<i>Never</i>		<i>Sometimes</i>			<i>All the time</i>	

How often does X discuss intimate or personal issues with you?

1	2	3	4	5	6	7
<i>Never</i>		<i>Sometimes</i>			<i>All the time</i>	

Appendix 6:

Measures of ‘visible’ traits, for providing a baseline of comparison for self-observer agreement

Extroversion - used in Studies 3, 4, 5, & 6

Please rate what you think your own personality is on the following dimensions:

1	2	3	4	5	6	7
<i>Restrained</i>						<i>Candid</i>
1	2	3	4	5	6	7
<i>Active</i>						<i>Passive</i>
1	2	3	4	5	6	7
<i>Silent</i>						<i>Talkative</i>
1	2	3	4	5	6	7
<i>Aloof</i>						<i>Gregarious</i>
1	2	3	4	5	6	7
<i>Extroverted</i>						<i>Introverted</i>

Political views – used in Study 6

Please mark where your political views lie on a scale from left (liberal) to right (conservative):

1	2	3	4	5	6	7
<i>Very Liberal / Leftwing</i>						<i>Very Conservative / Rightwing</i>