

THE IMPACT OF INTRA-GROUP INTERACTION ON IDENTITY AND ACTION

Submitted by Laura Grace Elizabeth Smith to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Psychology, October 2008.

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Laura Grace Elizabeth Smith

.....

ABSTRACT

The unifying theme of the chapters presented in this thesis is that intra-group interaction impacts on in-group identity content, and this content provides a foundation for social action and social behaviour. The primary goals of this thesis are first, to demonstrate that social realities can be established and transformed through interaction; and second, to investigate why the process of intra-group interaction can spark and exacerbate social conflict. In Chapter 1, I review and attempt to theoretically integrate the disparate literatures on group discussion, identity and action.

In Chapter 2, I investigate the effect of interaction on the positive-negative asymmetry effect (PNAE). In Study 2.1, participants were more likely to discriminate on rewards than fines, and find allocating rewards to be a more legitimate and pleasant act than allocating fines. Conversely, participants thought allocating fines would have a more negative effect on recipients and felt more negative about allocating fines than rewards. In Study 2.2, when in-group advancement was obstructed, no PNAE was found: obstruction was sufficient justification for out-group punishment in its own right. When in-group advancement was *not obstructed*, the PNAE reversed after group discussion, such that more hostility occurred when participants administered fines than when they awarded rewards. This reversal was mediated by processes of norm formation.

In Chapter 3, I describe three studies which show that consensual intra-group discussions about a negatively regarded out-group increased inter-group hostility. Study 3.1 compared group discussion about immigrants with individual reflection. Results showed that group discussion informed the content of stereotypes, which led to support for anti-immigrant policies. In Study 3.2, participants discussed either an irrelevant topic, the out-group stereotype, or the out-group stereotype plus what

concrete actions should be taken towards that group. Only discussion of the stereotype significantly increased hostility, suggesting that the psychological products of discussion *per se* (cohesion, identification, etc.) are not solely responsible for hostility. Rather, social validation of the stereotype explained why its discussion increased hostility. Study 3.3 replicated these results with a behavioural measure.

In Chapter 4, I present two studies which controlled for the content of interaction by showing participants short films of similar others having a group discussion. Study 4.1 investigated the paradoxical finding that when groups discuss potential courses of action against an out-group, they are less likely to act than when they discuss simply the out-group stereotype (Chapter 3). Results suggested that when group discussions imply that there is social consensus about a course of action, even the advocacy of extreme actions can increase support for (more moderate) social action. Study 4.2 manipulated whether or not the discussants consensualised on the out-group stereotype, whilst controlling for discussion content. Only when the discussion ended in consensus did participants identify with the discussants and perceive norms for social action.

In Chapter 5, I address how social identities and their associated (self-) stereotypes can disadvantage members of low status groups, but how they can also promote social change. The data demonstrates that consensualisation in small groups can transform (or reconfirm) such stereotypes, thereby eliminating (or bolstering) stereotype threat effects. In Study 5.1, female participants were asked why men are (or are not) better at maths. They generated their answers individually or through group discussion. Stereotype threat was undermined only when they collectively challenged the stereotype. Content analyses suggest that discussions redefined in-group and out-group stereotypes, providing the basis for stigma reversal or confirmation. In Study 5.2, male and female participants confirmed or challenged the stereotype in same-

gender discussion groups or no discussion, baseline conditions. After a discussion that confirmed the stereotype, women displayed signs of stereotype threat and men's performance was "lifted". When they challenged the stereotype, the difference between men and women on the maths test was eliminated.

Overall, the results reported in this thesis suggest that intra-group interaction enables group members to develop an understanding of their common ideology, which may establish the consensual basis of their identity content. If such consensualisation occurs, this provides them with a sense that their perceptions of reality are socially valid, and gives rise to (implicit or explicit) in-group norms. This provides individuals with a solid foundation upon which they may act. The implications of these conclusions are discussed in Chapter 6.

TABLE OF CONTENTS

<i>Abstract</i>	Page	ii
<i>List of tables and figures</i>		x
<i>Acknowledgements</i>		xii
1 Introduction		1
Aims and objectives of this thesis		2
Definition of terms		4
Overview		6
Intra-group interaction		6
Early research using group discussion		6
<i>Group decision</i>		6
The group polarisation phenomenon		7
<i>Normative influence</i>		8
<i>Social comparison theory</i>		9
<i>Pluralistic ignorance</i>		11
<i>Informational influence</i>		12
<i>Persuasive arguments theory</i>		12
<i>Persuasion</i>		14
<i>Critique</i>		15
Identity		17
The social identity approach		18
<i>Social identity theory</i>		19
<i>Self-categorisation theory</i>		21
<i>SCT and social influence</i>		22
<i>The SCT account of group polarisation</i>		23
Theories of interdependence		27
The inductive model of identity formation		28
The influence of identity content on inter-group relations		32
<i>Stereotype consensus</i>		32
<i>Recent advances</i>		34
Action		35
Attitudes in action		35
Interaction in action		37
<i>Theories of oppression</i>		39
<i>Theories of collective action</i>		41
Summary		44
The current research		45
2 Intra-group interaction and the development of norms which promote inter-group hostility		48
The PNAE and its causes		49
Consensualisation and norm formation		50

<i>In-group obstruction as an aggravating condition</i>	52
Study 2.1.....	53
Method	54
<i>Dependent measures</i>	54
Results and discussion	55
Study 2.2.....	56
Method	57
<i>Participants and design</i>	57
<i>Independent variables</i>	57
<i>Dependent measures</i>	58
<i>Procedure</i>	59
<i>Pre-consensus phase</i>	60
<i>Consensus phase</i>	61
Analytic strategy	62
<i>Condition effects</i>	62
<i>Tests of specific hypotheses</i>	63
Results	64
<i>Discrimination</i>	64
<i>Condition effects</i>	64
<i>Specific hypotheses</i>	65
<i>Pre-consensus phase</i>	65
<i>Consensus phase</i>	65
<i>Hostile norm</i>	67
<i>Condition effects</i>	67
<i>Specific hypotheses</i>	68
<i>Pre-consensus phase</i>	68
<i>Consensus phase</i>	68
<i>Mediation</i>	69
<i>Social identification</i>	70
Proportion of respondents behaving fairly	71
General discussion.....	72
<i>Conclusion</i>	76
3 The power of talk: legitimising hostile social action.....	78
<i>The power of talk</i>	79
Normative and informational explanations	80
<i>Stereotype consensualisation and the normalisation of implicit ideologies</i>	82
<i>Present research overview</i>	84
Study 3.1.....	84
Method	85
<i>Participants and design</i>	85
<i>Procedure</i>	85
<i>Dependent measures</i>	86
<i>Analytic strategy</i>	87
Results	89

<i>Discussion content</i>	89
<i>HLM analyses</i>	90
<i>Mediation</i>	90
Discussion	91
Study 3.2.....	92
Method	93
<i>Participants and design</i>	93
<i>Procedure</i>	94
<i>Dependent measures</i>	94
<i>Analytic strategy</i>	96
Results	96
<i>Mediation</i>	97
Discussion	99
Study 3.3.....	100
Method	101
<i>Participants and design</i>	101
<i>Procedure</i>	101
<i>Dependent measures</i>	101
Results	102
<i>Discussion content</i>	102
<i>HLM analyses</i>	104
Discussion	107
General discussion.....	108
<i>The power of talk</i>	109
<i>Limitations and future research</i>	110
<i>Conclusion</i>	111
4 The role of in-group consensualisation in identity (in)formation.....	112
<i>The crucial role of consensus</i>	113
<i>Alternative explanations and empirical challenges</i>	115
(1) <i>The consensual discussion of actions</i>	116
(2) <i>Discussion of extreme actions</i>	116
(3) <i>The prior inter-group context</i>	117
(4) <i>The process of consensualisation</i>	118
<i>Overview</i>	118
Study 4.1.....	119
Method	119
<i>Participants and design</i>	119
<i>Stimuli and procedure</i>	119
<i>Dependent Measures</i>	120
Results	123
<i>Mediation</i>	127
Discussion	128
Study 4.2.....	130
Method	131

<i>Participants and design</i>	131
<i>Stimuli and procedure</i>	131
<i>Dependent measures</i>	132
Results and discussion	133
<i>Mediation</i>	136
General discussion.....	137
<i>Implications and future directions</i>	139
<i>Conclusion</i>	142
5 Changing social identities: social consensus as a foundation of stereotype threat.....	144
<i>The negative effects of stereotype threat</i>	144
<i>Stereotype lift</i>	145
<i>Combating stereotype threat</i>	145
<i>The power of talk: transforming social identities</i>	145
<i>The present research</i>	147
Study 5.1.....	147
<i>Hypotheses</i>	148
Method	148
<i>Participants and design</i>	148
<i>Materials and procedure</i>	148
<i>Dependent measures</i>	149
<i>Analytic strategy</i>	150
<i>Qualitative data</i>	150
<i>Measure of consensus</i>	151
Results	152
<i>Stereotype consensus</i>	152
<i>Performance effects</i>	156
<i>Maths test score</i>	157
<i>Psychological effects</i>	159
<i>Evaluation apprehension</i>	159
<i>Anger</i>	159
<i>Gender identification</i>	160
<i>Mediation</i>	160
Discussion	160
<i>Transforming social identities</i>	161
Study 5.2.....	163
Method	163
<i>Participants and design</i>	163
<i>Materials and procedure</i>	164
<i>Dependent measures</i>	164
<i>Analytic strategy</i>	166
Results	166
Performance effects	166
<i>Maths test score</i>	166

Psychological effects	169
<i>Anxiety and evaluation apprehension</i>	169
<i>Social validation</i>	170
<i>Anger</i>	170
<i>Stereotype threat</i>	170
<i>Identification</i>	170
<i>Mediation</i>	170
Discussion	171
<i>Psychological effects</i>	171
General discussion.....	173
<i>Implications and future directions</i>	175
<i>Conclusion</i>	176
6 General discussion.....	177
The purpose of this thesis	179
Brief summary of the results	179
Theoretical and practical implications	184
<i>Towards an interactive model of social action</i>	184
Considering potential alternative explanations	188
<i>“Isn’t this still just individual cognition?”</i>	188
<i>“Is interaction a necessary condition for these effects to occur?”</i>	189
<i>“Isn’t this just polarisation?”</i>	191
Concluding remarks	193
References.....	195

LIST OF TABLES AND FIGURES

Table 1.1	The 3 stages of referent information influence	Page 24
Table 2.1	Coding of contrast variables for HLM, Study 2.2	64
Table 2.2	Mean scores at pre-consensus and consensus, Study 2.2	66
Figure 2.1	Mean discrimination scores, Study 2.2	67
Figure 2.2	Mean hostile norm scores, Study 2.2	69
Table 2.3	<i>N</i> Participants choosing to distribute resources equally between groups, Study 2.2	71
Table 3.1	Mean individual-level scores and scale intercorrelations, Study 3.1	88
Figure 3.1	The percentage of participants who mentioned each key code, Study 3.1	90
Figure 3.2	Gamma coefficients of the mediation model, Study 3.1	91
Table 3.2	Mean individual-level scores and scale intercorrelations, Study 3.2	98
Figure 3.3	Gamma coefficients of the mediation model, Study 3.2	99
Table 3.3	Coding of contrast variables, Study 3.3	105
Table 3.4.	Frequency of UKIP votes, mean individual-level scores and scale intercorrelations, Study 3.3	106
Figure 4.1.	Mean collective action intention scores, Study 4.1	125
Table 4.1.	Mean scores and scale intercorrelations, Study 4.1	126
Figure 4.2.	Beta coefficients of the mediation model, Study 4.1	127
Figure 4.3.	Mean collective action intention scores, Study 4.2	134
Table 4.2.	Mean scores and scale intercorrelations, Study 4.2	135
Figure 4.4.	Beta coefficients of the mediation model, Study 4.2	136
Figure 5.1	Stereotype content used to explain why men are better (A) or not better (B) than women at maths, Study 5.1	153

Table 5.1	Percentage of participants who mentioned each code in response to the question, “Why is it true that men are better than women at maths?” Study 5.1	155
Table 5.2	Percentage of participants who mentioned each code in response to the question, “Why is it not true that men are better than women at maths?” Study 5.1	156
Figure 5.2	Mean maths test scores, Study 5.1	157
Table 5.3	Means, standard deviations, scale reliabilities and intercorrelations of dependent variables, Study 5.1	158
Figure 5.3	Mean maths test scores, Study 5.2	167
Table 5.4	Means, standard deviations, scale reliabilities, and intercorrelations of dependent variables, Study 5.2	168
Figure 6.1	Models of social action: [A] Traditional model of social action, [B] Towards an interactive model of social action (IMSA)	180
Table 6.1	Summary of results	182
Appendix 3.1	Definitions of codes and example quotes, Studies 3.1 and 3.3	220
Appendix 5.1	Definitions of codes for responses to the question, “Why is it true that men are better than women at maths?” Study 5.1	223
Appendix 5.2	Definitions of codes for responses to the question, “Why is it not true that men are better than women at maths?” Study 5.1	225

ACKNOWLEDGEMENTS

There are so many people to whom I am indebted for their help, advice, and support in the writing of this thesis. The unvarying love and support of my parents, Ian and Gill Smith, and my brother and sister, Tim and Pip, has meant more to me than words can express. However, without my inspirational supervisor, Tom Postmes, this thesis simply could not have come about. Thanks must also go to my 2nd supervisor Alex Haslam, who has been a great mentor for many years now. I will always value the unique circle of support of members of SEORG and IPSIS, at the School of Psychology, University of Exeter. It was a really wonderful environment within which to spend my years as a research student, particularly with the Collective: Ivonne Hoeger, Andrea Day, and Aisling O'Donnell – true friends.

I was so lucky to be granted the opportunity to visit the University of Queensland (UQ), Brisbane during my PhD studies, and I would like to thank Winnifred Louis, Matthew Hornsey and Jolanda Jetten for their support, invaluable advice and hospitality there. Further thanks must go to Winnifred and Jolanda, along with Tom, for their excellent career advice, which helped me to obtain a post-doctoral fellowship at UQ. Thank you also to Craig McGarty and the other members of the identity formation research group at Murdoch University and the Australian National University for inviting me to visit, and for their feedback on my research.

I would like to thank Catherine Amiot and Kai Sassenberg for their advice on the outcome distributions used in Study 2.2, Marijtje van Duijn for her advice on mediation within multi-level models, and Lucy Johnston, Deborah Kashy, Aisling O'Donnell, and the anonymous reviewers for their comments on earlier versions of the manuscripts adapted for Chapters 2 and 3. Finally, I must thank the Economic and Social Research Council for the scholarship which supported my studies.

Laura Smith,
October 2008

CHAPTER 1

Introduction

“Whatever social unit we may be concerned with, this system of shared and socially confirmed notions is the basis from which behaviour can be described and predicted...As soon as anything changes in the environment, [...] new choices confront the individual. The processes which underlie these choices, and thus constitute the psychological aspects of social change at all levels are the proper subject matter of social psychology.” (Tajfel, 1972; p. 115).

“Faced with ethnic cleansing in Bosnia or genocide in Rwanda, we need to redress the evident mismatch between the effect that most research on inter-group bias has studied and most theories have sought to address and the most striking social problems that this research ought to be able to address.” (Hewstone, Rubin, & Willis, 2002; p. 594).

“Come to Cronulla this weekend to take revenge. This Sunday every Aussie in the Shire get down to North Cronulla to support the Leb and wog bashing day” (SMS circulated immediately before the riots in Cronulla, Australia; Jackson, 2006).

We can all relate to the experience of social interaction. It can be gratifying, helping us to bond with friends and colleagues. This type of interaction can help establish, inform, and bolster a sense of “us” and “who we are”. It also provides an arena in which we can verify and validate our beliefs. Discussion between “insiders”, of the attitudes and actions of “outsiders”, may function to vent social frustrations, informing “us” about “them”, i.e., validating out-group stereotypes. In this way, speaking about other groups can be poison for inter-group relations. At the extreme, hostile dialogue within societies can be a catalyst for war, sparking violence and helping

to perpetuate social conflict. For example, escalating social dialogue reportedly played a significant role in sparking the Australian riots of 2004 and 2005 (Jackson, 2006), in conflicts in Ambon (Erdentuğ & Colombijn, 2002) and in Rwanda (Amnesty International, 1996).

The role of mass communication as a catalyst for violence is widely acknowledged. Hutu extremists used the radio to start the 1994 genocide in Rwanda with the message: “It’s time to cut down the tall trees” (Harding, 2003). The literal result of this communication was that machetes were used to slaughter Tutsi men, women and children. A hundred days later, more than 800,000 people had been murdered, at a rate of one person every 10-12 seconds (BBC News, 2004). “Tall trees” is a reference to the stereotype that Tutsis were generally taller than Hutus, categorised as they were by physical characteristics. The radio broadcast used this stereotype of the Tutsis to make their actions seem more valid and legitimate. Furthermore, it was a signal of group (Hutu) consensus on a plan of action. The stereotypical and ideological content of the broadcast was the foundation of the subsequent hostile social action by Hutu extremists.

However, mass communication such as that used in the examples above is a relatively recent invention. According to this dissertation, “the word on the street” is potentially more powerful in its effects. Traditional means of communication and persuasion can be found in small groups, where people can create shared representations of their physical and social reality (Postmes, Baray, Haslam, Morton, & Swaab, 2006). This process of creating a shared sense of meaning through micro-conversations is the subject of this programme of research, with a particular emphasis on the role that interactions play in our understanding of inter-group relations.

AIMS AND OBJECTIVES OF THIS THESIS

Atrocities like the Rwandan genocide punctuate history, yet the significant role

of small-scale communications in helping to provoke and perpetuate inter-group violence has seldom been empirically studied. In light of the discrepancies between the reality and study of social perception and action, this thesis had a dual purpose: First, to examine, from a social psychological perspective, how social realities are established and transformed through interaction; and second, to investigate why the process of intra-group interaction can spark and exacerbate social conflict.

The message of the opening statements contrasts with the lack of ecological resonance of much of the research into social behaviour. As Hewstone et al., (2002) argued, prior research has often focused on phenomena such as in-group bias (see also Ellemers, Doosje, & Spears, 2004). In contrast to Tajfel's (1972) concern with social change in dynamic social systems, social psychology has tended to specialise in the study of isolated individuals' responses to a static social context. Orthodox research does not take into account the fact that the presence of – and psychological interdependence with – similar others is often *necessary* for social actions. Individuals may not even be able to *conceive* of thinking certain thoughts or undertaking certain acts alone, because the presence and validation of other in-group members are essential ingredients for reaching certain outcomes.

For this reason, the current research breaks with the social psychological tradition of studying inter-group behaviour in isolated individuals. It is proposed that significant progress in the understanding of inter-group behaviour can be made by investigating this behaviour in one of the contexts in which it originates, that is, in the *interactive group*. In this modest way, the present studies gave participants the opportunity to breach one key restriction imposed by modern experimental research, and discuss the context in which they were placed.

The individualistic methodology used in contemporary social psychological research is accompanied by a meta-theoretical perspective that social psychological

phenomena such as inter-group bias and stereotyping occur within, and are automatic social products of, individuals' perceptions of inter-group relations and particular out-groups. This knowledge is either assumed to exist *a priori*, or otherwise assumed to be readily inferred from features of the social context, in such a way that *individual cognitive* mechanisms are the only theoretical ingredients which are required to account for these social outcomes. From such a vantage point, the study of an individual reflecting on inter-group relations (or responding with key-presses to stimuli on a computer screen) is an undisturbed and “pure” method of studying the psychological processes behind these social phenomena. What generally remains implicit and empirically unexplored, is the pivotal assumption that these perceptions of social reality cannot emerge from individual cognition alone—they are, or are founded upon, knowledge structures that are consensually held among members of a particular group. One major disadvantage of the predominant paradigm for studying *social* cognition is that we circumvent the fact that individuals' knowledge and perceptions of the world are shaped or constructed socially by making up a “collective mind”. Therefore, we ignore the fact that the constituent individuals are establishing a “we” in the first place.

In addressing the aims above, I hope to re-conceptualise inter-group action as motivated (in part) by intra-group dialogue, and to understand one of the reasons why people who share particular social ideologies are motivated to advocate and engage in social action. We intend to demonstrate that social interaction serves to establish social reality by subjectively validating individual cognitions or observations. We argue that social communication has the power to influence identity content, including norms for social action.

Definition of Terms

In this thesis, I apply the theoretical framework of the *social identity approach* (comprising social identity theory, SIT; Tajfel & Turner, 1979, 1986; and self-

categorisation theory, SCT; Turner, 1985; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987) to investigate the social-interactive underpinnings of *social action*. According to this approach, the individual's self-concept comprises not only personal characteristics that define the person as a unique individual (their *personal identity*), but also characteristics that are shared with members of the various in-groups to which they may belong (their *social identities*). Tajfel and Turner (1979, p. 40) defined *social identity* as, "Those aspects of an individual's self-image that derive from the social categories to which he perceives himself as belonging". According to SCT, when people think of themselves in terms of a social identity (i.e., when their social identity is *salient*) their actions occur within the *social normative framework* associated with their group. In other words, social identity salience is the process that explains why social action occurs: If one's social identity as British is made salient, the norms and standards that are embedded in that identity are activated, and British people start behaving and thinking in a more British manner.

When defining social action, it may be useful to refer to a commonly cited definition of collective action: "A group member engages in collective action anytime that he or she is acting as a representative of the group and the action is directed at improving the condition of the entire group," (Wright, Taylor, & Moghaddam, 1990; p. 995). I define "*social action*" somewhat more broadly, as any behaviour undertaken by a person on behalf of their group. Although this remains to be empirically established, one can assume that these actions are dictated by the normative content of their social identity. This broader definition of *social action* includes actions undertaken for the benefit of the group as a whole, and actions that appear to be merely for the individual's own private benefit (see also Postmes & Smith, in press).

As Turner (1991, p. 2) eloquently expresses, "Our understanding of what a social norm is changes with our understanding of the influence process." According to

Turner (1991), “Influence relates to the processes whereby people agree or disagree about appropriate behaviour, form, maintain or change social norms and the social conditions that give rise to, and the effects of, such norms” (p. 2). This definition of social influence will serve as our starting point. Our definition of a social norm includes all in-group and out-group stereotypes, actions and attitudes that are deemed *acceptable, appropriate* and *legitimate* by members of a given in-group. A norm thus serves as an anchor for what is valid, and gives group members confidence in executing its prescribed and therefore socially desirable social actions – a concept known as “subjective validity” (Festinger, 1950, Turner, 1991).

Overview

This chapter begins with a review of the early research involving group discussion, which was not situated within the explicit terms of group identity. Next, an evaluation of the normative and informational influence accounts of group polarisation precedes a review of the social identity approach, and stereotype consensus, attitude and social action research. These topics have produced conceptually distinct literatures. Therefore, an attempt is made throughout this chapter to integrate the theories surrounding social interaction with those of identity and action. In sum, this chapter attempts to pinpoint some of the inconsistencies, and lack of theoretical integration of the purported definitions and processes underlying interaction, identity and action.

INTRA-GROUP INTERACTION

Early Research on Group Discussion

Group Decision

Research conducted during the Second World War marked a watershed in the understanding of social reality. Lewin (1947) suggested the group was an entity with properties distinct from those of the sum of its parts, seeing it as a “dynamic whole which is characterised by a close interdependence of [its] members,” (p. 8). Following

this, Lewin's (1953) studies on group decision-making were an emphatic and seminal demonstration of the power of group processes to influence individual action. By giving participants either persuasive lectures or the opportunity to have a discussion about that persuasive information, and reach a group decision, he showed that the discussion was far more influential for subsequent behaviour. He argued that listening to a lecture is passive and individualising; yet participating in a group discussion involves a higher degree of involvement, and offers the opportunity to express motivations. Therefore, the group discussion provided a more effective medium for action. Lewin (1953) stated that group changes are more easily effected than individual changes, because of the "unwillingness of the individual to depart too far from group standards" (p. 290).

However, Lewin's research relied upon self-reported behaviour, and is marked by several confounds (see also the *Critique* section). For example, there were differential expectations of follow-up, different experimenters per condition, and different levels of (public) commitment. A subsequent study by Coch and French (1948) neatly backs up the findings of Lewin, however. They discovered that a participatory group decision on an output goal increased production relative to the control condition. Furthermore, the effect of discussion on attitudes and behaviour has been replicated and elaborated by a multitude of further early research on group polarisation.

The Group Polarisation Phenomenon

Initial research on group decision making suggested that under certain circumstances, after interacting, group members will make a riskier decision together than would the individuals if making the decision on their own ("risky shift"; Stoner, 1968, Wallach & Kogan, 1965). It was subsequently discovered that such shifts did not just make groups more "risky"; group decisions could also be more cautious (Fraser,

Gouge, & Billig, 1971). Furthermore, these effects were not restricted to group decisions involving risk or caution, but were more general to any opinion or attitude. Indeed, a meaningful group discussion can shift individual group members' pre-discussion opinions and attitudes about a variety of issues (ranging from investment decisions to political viewpoints) to a more extreme position, in the direction in which they already tended (group polarisation: Moscovici & Zavalloni, 1969; see also Lamm & Myers, 1978, Myers, 1982, Myers & Bishop, 1970; Turner, 1991, Wetherell, 1987;).

These polarising effects of group discussion are anything but intuitive. As Wetherell (1987) asked, "Why do people not become more moderate or compromising through discussing an issue with their peers?" (p. 142). Indeed, on the basis of many theories of social influence (including Festinger's, 1954, social comparison theory) one would expect group discussions to moderate (or average out) the individual group members' perspectives. In order to address this question, two different classes of theoretical explanations for group polarisation have been suggested: (a) normative and (b) informational. In line with Deutsch and Gerard's (1955) classic work, the informational/normative distinction here refers to two distinct pathways to social influence. The term *normative influence* has been used to describe the pressure on individuals to conform to social expectations or group standards. *Informational influence* may be defined as the pressure to accept information obtained from another as evidence about reality.

Normative Influence

"It is [a] decisive fact about a person whether he has the freedom to act according to his beliefs or whether he has failed to develop (or has lost) the possibility of independence." (Asch, 1952; p. 451)

In terms of explaining group polarisation, social comparison processes (Festinger, 1954) are assumed to lead group members to infer the socially desirable (or

normative) positions within the group. In a group discussion which exposes a particular group preference for (or against) a particular issue or decision, group members therefore polarise in an effort to comply with the social norms of the group.

Normative influence is said to cause people to act in ways which they would not, if isolated and autonomous. As Nolan, Schultz, Cialdini, Goldstein, and Griskevicius (2008) point out, it has been accused of leading people to give falsehoods (Asch, 1958), use illicit drugs (Maxwell, 2002) or fail to help someone in need (Latané & Darley, 1970). Theories of normative influence developed against the backdrop of Sherif's (1935) analysis of norms as changing within an individual with his or her social "frame of reference", and Asch's (1952, 1958) classic conformity studies.

In Asch's (1952) initial study, participants were asked to judge the length of a line. In the experimental condition, the participant was placed in a group of confederates, who unanimously made incorrect judgements. Asch found that 33.2% of the errors in the experimental condition were identical with those of the incorrect majority, compared to only 7.4% in the control condition. Asch (1952) argued that this was because of a pressure to conform to the consensual expressions of the other group members (the confederates), rather than stand out by expressing a different perception. He framed this pressure as a "dynamic requirement" (p. 484) of the situation. However, the prevalent interpretation of these results has been in terms of straightforward social pressure exerted by a majority group against a lone deviant (e.g., Deutsch & Gerard, 1955)—the idea that the Asch experiments were demonstrations of the dynamic development of group norms has not generally been embraced (cf., Turner & Killian, 1972).

Social Comparison Theory

Festinger (1950) argued that it is "stating the obvious" (p. 272) that these pressures to conform in a group are exerted via a process of communication among

group members. According to the social comparison approach, two major sources of pressures toward uniformity of opinions or attitudes within the group are social reality and group locomotion. The group locomotion pressure occurs when uniformity of opinion is desirable for a group so they can move towards a common goal. The principle of social reality pressure assumes that there must be a basis for the validity of beliefs. If a fact cannot be checked as valid from physical reality, it must be checked by reference to whether other people share the opinion that the fact is valid. This is known as *subjective validity* and is achieved through social validation processes. A fact is therefore deemed valid if it is held consensually within a group. The implication is that individuals *depend upon* the perceptions of others (in a reference group deemed appropriate) for their understanding of reality: Consensus is valuable.

Building on this approach, social comparison theory (Festinger, 1954) outlined a process by which individuals alter their opinions according to the group norm. Basically, social comparison theory assumes that humans have a functional drive to evaluate their opinions and abilities. In the absence of a non-social way to evaluate our own opinions or abilities, we rely on those of others. We choose someone similar to ourselves for comparison purposes, and we are relatively attracted to situations where there is a close match between our abilities and opinions and those of others. If there is a discrepancy between opinions in a group, there is a drive to reduce that discrepancy. If a group is important, there will be pressure toward uniformity of opinion. Furthermore, the person who most typifies the opinion of the group will feel a greater urge to influence the opinions of others. The more relevant the opinion to the group, the stronger the pressure will be to reach uniformity.

The influence of social comparison theory on later theories of identity and identity change will become obvious in later sections of this chapter. However, although it provides insight into the social pressures and motivations of dynamic intra-

group life, social comparison theory does not consider the role of context in affecting intra-group dynamics, and group polarisation. Neither does it specify when and why the opinions of particular others become important or relevant.

Pluralistic Ignorance

The degree to which social norms anchor our everyday actions and thoughts is not generally visible to us. *Social* norms are also (by definition), widely shared perceptions of how things *ought* to be. Therefore, (much like the influence of our culture on our beliefs and customs) we are not ordinarily in a position to observe the social influence they exert on us. The phenomenon of pluralistic ignorance uses the unusual (although by no means rare) situation in which there is a discrepancy between prevailing attitudes and behaviour, to illustrate how social norms are inferred and how they regulate our behaviour.

Pluralistic ignorance is the tendency for individuals to change their attitudes in order to match an *erroneously* perceived norm. Like polarisation, it has been identified as a vehicle for social change (Breed & Ktsanes, 1961; Fields & Schuman, 1976; Miller & McFarland, 1991; Prentice & Miller, 1993). Research shows that people infer group norms from the public behaviour of others (e.g., Sherif, 1935; Asch, 1952). The state known as *pluralistic ignorance* is characterised by the conviction that one's private attitudes and judgements are different from those of others, even though one's public behaviour is identical (e.g., Miller & McFarland, 1987, 1991). However, there may be a common misrepresentation of the views of others. In other words, "Their own behaviour may be driven by social pressure, but they assume that other peoples' identical behaviour is an accurate reflection of their true feelings" (Prentice & Miller, 1993; p. 244). If this is the case, and an individual misrepresents the private views of others, the perceived norm will be incorrect (Prentice & Miller, 1993). The assumption that other group members endorse the norm (i.e., it is *socially shared*), is what gives the

norm its power of social influence over individuals' attitudes and behaviour (the "*importance of universality*"; Prentice & Miller, 1993, p. 244). Therefore, if "consensus (or the appearance of consensus) breaks down, the norm loses its influence", (Prentice & Miller, 1993; p. 244; see also Asch, 1952).

Therefore, normative theories assume that pressures of unanimity and conformity drive group polarisation. However, one problem with normative accounts of polarisation is that groups are not always driven by conformity (Postmes, Spears, & Cihangir, 2001) and are often characterised by passionate debate. Reaching a norm can be the result of a melting pot of ideas, information, and interpretations of events in context. For example, emergent norm theory (ENT; Turner & Killian, 1972) posited that norms are not imposed, but must organically develop through interaction because of a drive to make sense of an event which clashed with the old normative framework. A large part of the development or inference of group norms is the exchange of information. Indeed, norms themselves can be conceptualised as information about the group. This exchange of information during debate can have a significant influence on the attitudes and behaviour of individuals within groups.

Informational Influence

Persuasive Arguments Theory

Persuasive arguments theory (PAT; Vinokur & Burnstein, 1974) assumes that group polarisation occurs as group members become exposed to novel and persuasive arguments (i.e., information) during group discussion and modify their opinions accordingly. Specifically, "group-induced shifts" in choice after group discussion found in risky and cautious shift research are believed to be a consequence of "informational influence processes", and specifically "persuasive argumentation," (p. 305). PAT postulates that individuals are swayed towards the side of an argument for which there is the most novel, and therefore persuasive, information. Vinokur and Burnstein

(1974) argue that before discussion, if each member of the group knows a small proportion of the persuasive arguments, a considerable number of group members will be exposed to new persuasive arguments during the discussion. A shift would occur when those new arguments favour either side: thus, shifts towards caution occur if the new and persuasive arguments favour caution, whereas risky shifts occur in discussions where group members hear new arguments favouring risk.

PAT has received strong empirical support to suggest that the content of debate has a considerable influence on group members' post-discussion attitudes. For a number of reasons, however, it seems unlikely that information is the only (or even the main) factor that influences the choices of individuals within groups. For example, Lewin (1953) took great care to ensure that participants in all his experimental conditions were provided with the same information: Those who merely listened to lectures were given the same "facts" about offal that were given to those who were given the opportunity to discuss them. Indeed, Lewin argued that, "Lecturing may lead to a high degree of interest. It may affect the motivation of the listener. But it seldom brings about a definite decision on the part of the listener to take a certain action at a specific time. A lecture is not often conducive to a decision," (Lewin, 1953, p. 289). In other words, an additional psychological change (over and above informational influence) occurs as a result of processes during interaction within groups.

A further aspect of PAT, that seems unlikely on closer inspection, is the underlying assumption that rational processing of incoming information is the driving force behind these group effects, and in particular that *novel* information would be most persuasive. Research suggests that groups often tend to discuss information that all members already know, instead of unshared information that only a single member knows (Stasser & Titus, 1985). Moreover, group members value this shared information more highly *because* it is known by multiple group members (Postmes et al.,

2001; Wittenbaum, Hubbell, & Zuckerman, 1999). This finding is consistent with social comparison notions: upon discovering where the group stands, members tend to voice items of information which support the norm, and they value information which is consistent with the norm.

It is apparent from these findings, that group discussions are not just vehicles for “objective” information gathering, engaged in by autonomous individuals. The social dynamic in these groups may also play a significant role in the social influence process. Nevertheless, it is also clear that the content of group discussion is a key factor in determining the outcomes of the interaction.

Persuasion

The cognitive response approach (Greenwald, 1968) suggests that yielding to persuasive argumentation is associated with idiosyncratic cognitive responses to the message. Derived from this metatheory, it has been argued that attitudes are individually-held cognitive representations (e.g., Eagly & Chaiken, 1993), and can change with intra-individual elaboration processes, such as the extent to which people think about issue-relevant arguments (the elaboration likelihood model, or ELM; Petty & Cacioppo, 1981; 1986a; 1986b). The attitude literature does not postulate that information exchange is the sole influence on action, however.

The heuristic-systematic model (Chaiken, 1980, 1987; Chen & Chaiken, 1999) assumes that people have two ways of processing information: by systematically processing the content, or via heuristics, or cues peripheral to its content. For example, group consensus could be a short-cut to assessing the value of the information. In the elaboration likelihood model and the heuristic-systematic model, shared group membership is predicted to act as a heuristic, or “peripheral” cue for intra-individual attitude change, in terms of people who are liked and /or trusted, providing reality information. Other accounts of persuasion have highlighted how group identity may

impact upon attitude change (e.g., social judgement theory, Sherif & Hovland, 1961). According to social judgement theory, group members are more resistant to changing their attitude on a group-relevant issue because of the wider possibility of rejection and greater likelihood of fellow group members viewing incongruent messages as unacceptable.

Therefore, it seems that individual attitudes are influenced by identity concerns. Attitudes can be seen as “windows on identity,” (Hogg & Smith, 2007; p. 89), i.e., people assume that an individual’s expression of an attitude provides an insight into that person’s (or a group’s) ideology. However, the attitude change models focus on the individual level of analysis and therefore neglect the role of contextual variables, which often create the impetus for the “elaboration” of a relevant issue. The social contextual dimension is important for understanding why particular attitudes become the group norm, and why some do not (e.g., Hogg & Smith, 2007). As the aims of this thesis are to establish the interdependent element of social reality, which current methodological practice tends to prevent us from taking into account, it is important to consider how intra-group processes impact upon individuals’ psychological resources. A key characteristic of the informational explanation for group polarisation or social influence within groups is that it does not depend on group processes *per se*, merely on the processing of information. As the evidence for normative influence betrays, information processing is not the only factor driving group polarisation (De Dreu, Nijstad, & Van Knippenberg, 2008).

Critique

The normative and informational influence paths to polarisation are unlikely to be mutually exclusive, and arguably should not even be seen as distinct concepts (Turner, 1991). For example, to some extent, the information transferred during a group discussion could be described as “the group norm”. Indeed, a meta-analysis

suggested that both pathways to influence combine to produce polarisation (Isenberg, 1986). Wetherell (1987) reviewed the various accounts that were forwarded for the group polarisation findings, and highlighted several unresolved theoretical issues.

First, it is not useful, on the balance of evidence, to base polarisation models on externally defined values. Importantly, Wetherell (1987) asked, what makes a persuasive argument persuasive? Mackie and colleagues have shown that whether the source of information is in-group or out-group will affect the persuasiveness of the information (Mackie, 1986; Mackie & Cooper, 1984; Mackie, Gastardo-Conaco, & Skelly, 1992; Mackie, Worth, & Asuncion, 1990; Wilder, 1990; see also Fleming & Petty, 2000). Later research showed that the salience of psychological group memberships can affect which information is transferred during discussions and the extent of its influence on the individual (see Postmes, Haslam, & Swaab, 2005). Furthermore, Wetherell asks, in which situations will polarisation occur? None of the accounts of social influence above provide any comprehensive explanation by which socio-structural factors influence normative content.

On the other hand, we may also consider that what is “persuasive” is open to intra-group debate, rather than being dictated solely by the structural context in which the information is transmitted. In addition, why are some kinds of extremity desirable? Informational influence accounts do not explain why people value consensus / unanimity, or the conditions under which they would prefer to have a shared or a more nuanced perspective (Postmes, Haslam, et al., 2005; Postmes, Spears, & Cihangir, 2001). Wetherell (1987, p. 151) concludes that:

“It seems likely that certain kinds of extremity are not desirable generally but only because they become the tendency (norm) for the discussion group. If so, some account must be given of group norm formation and conformity processes in polarising groups. Social comparison and persuasive arguments models thus reach a similar impasse. The character of group interaction seems

likely to [...] influence what is seen as persuasive and what is valued [...] yet both theories tend to treat group discussion as a neutral environment [...] Little attention is paid as a result to factors or psychological states which might emerge from the group process itself.”

The assumption that group discussion is a neutral environment for information exchange is often made in research. The widespread use of focus groups to gather “information” is one example. However, small groups can develop their own ideology (e.g., Turner & Killian, 1972), which is related to the wider socio-structural political climate. For example, Lewin (1953) failed to situate the “group standards” in his studies within the socio-political environment of the time. It seems plausible that the direction of the group decision reached was influenced by the norms of the post-war context. He also did not have evidence or theory of a coherent process underlying his findings. It is possible that the group discussion condition allowed an identity to form. In this way, the consensual social change occurring within small interactive groups can be conceptualised as the formation or change of the group’s *social identity* (Postmes, Haslam, & Swaab, 2005; Postmes, Spears, Lee, & Novak, 2005). In terms of the aims of this thesis, this speaks to the establishment of social reality through interaction.

IDENTITY

The idea of an organically emerging sense of shared identity appears logical (at least if one accepts that the facts of inter-group life and social relations are socially mediated and therefore negotiated with other in-group members). However, it is not that straightforward to find a snug place for this idea within the social identity perspective. Using Tajfel and colleagues’ (1971) early group research as an empirical basis for his arguments, Turner (1982) chose to define the group in cognitive terms rather than by interdependence of the group members, and social cohesion. Instead of integrating elements of the (traditional) social-cohesive models of the group with the

cognitive redefinition he advanced, he chose to fully subject any alternative theories of social cohesion (e.g., as advanced by interdependence theory; Kelley & Thibaut, 1978; Thibaut & Kelley, 1959) to an overarching cognitive framework. This approach had very fortunate theoretical consequences (e.g., the construction of a grand theory for group processes). However, one consequence that this dissertation questions is that, in doing so, Turner also situated the group firmly within the mind of the individual. Turner (1982) pointed out, that rather than individuals being attracted to each other as idiosyncratic individuals, they are attracted as group members. Social cohesion is thus conceptualised as a product of self-categorisation (rather than a pre-cursor), which results in mutually perceived similarity between the self and other in-group members (e.g., Hogg, Coopershaw, & Holzworth, 1993). Turner marked his thesis on the group with the following caveat, however, suggesting that the two conceptualisations of group identity (interdependence and categorisation) are not necessarily opposing camps:

“There is some polemical value in stressing single-mindedly the virtues of a new idea and playing down those of the old. It may well be, indeed it is almost certain, that at some stage a theoretical conception of the social group will emerge which integrates the truth contained in both definitions. A pendulum, however, must swing in both directions before it can come to rest.” (Turner, 1982; p. 17)

The Social Identity Approach

The extent to which a person’s identity is constructed by the contexts in which they find themselves has been a subject of large-scale social psychological enquiry. The social identity approach posits that the socio-structural context can have a profound impact on the salience and content of social categories. However, the popularity of the social identity paradigm has somewhat sidelined the investigation of the impact of intra-group dynamics, characterised by inter-personal exchanges, on identity content. In this section, first I outline the tenets of the social identity approach, and then I establish its

usefulness in understanding the impact of interaction on identity content, as demonstrated by the research described in the previous section.

Social Identity Theory

In Sherif's (1956) classic boy-camp studies, competition between groups for scarce resources was found to promote competition and inter-group conflict, and simultaneously to enhance intra-group cohesion, identification with, and positive attachment to, the in-group. This was one important impetus for Tajfel and Turner (1979) to develop *social identity theory* (SIT). SIT provides a framework for understanding the behaviour of individuals when they categorise themselves as group members. Tajfel and Turner (1979) proposed that behaviour rests on a continuum, of which the extreme ends correspond to inter-personal and inter-group behaviour. Social identity refers to the part of a person's self-concept that is derived from their psychological group memberships. When a social identity is salient, individuals will see the world through the lens of this identity, and subscribe to the *socially shared* identity content (group norms, ideology, stereotypes and perceptions of the socio-structural context) associated with that identity. According to this approach, individuals will strive to maintain or enhance a positive social self-image. When social identity is salient, self-evaluation is determined by inter-group comparisons.

Alongside their self-categorisation along the interpersonal-intergroup continuum, individuals are also said to hold beliefs about the nature of structural relations between groups in their society. The two extremes of these beliefs are that the social system is characterised either by social mobility or social change. Social mobility refers to the belief that the society is permeable, and therefore individuals can move between groups. Social change beliefs indicate that the groups within the social system are markedly stratified and there is no possibility to move between groups. When the evaluation of the in-group relative to the out-group is not positive, individuals will

embark upon strategies to leave their group for one that is more positively distinct, or to make the in-group more positively distinct. The choice of strategy depends upon the aforementioned social belief systems.

In the hierarchical social change belief systems, individuals are likely to act as in-group members relative to defined relationships with out-groups. Socio-structural conditions, such as inter-group conflicts of interest (cf. realistic group conflict theory, RCT; Campbell, 1965; Sherif, 1956, 1967; Sherif & Sherif, 1956) are highly salient in these social change systems of beliefs. Therefore, these systems are associated with inter-group conflict and behaviour in line with the extreme “social identity” end of the interpersonal-inter-group continuum.

The minimal group studies (Tajfel, 1978; Tajfel, Billig, Bundy, & Flament, 1971) demonstrated that mere categorisation of isolated individuals into ad-hoc, arbitrary groups was sufficient for inter-group discrimination, exhibited as in-group favouritism or bias. In this way, minimal group research is of much value for our understanding of group dynamics. Indeed, in some respects, taking into account the caveat that minimal groups were never meant to mirror the real world; minimal groups can be maximal. They demonstrate the instinct of humans to discriminate in favour of their group when stripped of all social cues, except group membership. However, in doing so, these studies may be interpreted as negating the argument that explicit conflicts of interest, social cohesion or interdependence of individuals is necessary for the occurrence of social action. Tajfel and Turner (1979) argued that the “essential criteria” for group membership are that the individuals define themselves and are defined by others as group members, rather than inter-member interaction or interdependent goals (p. 40). This blow to theories of interdependence is softened however, by the very minimal nature of these laboratory studies. These minimal contexts were not intended to tell the whole story,. Indeed, by stripping away the meaning or content of the group

membership, research with minimal groups cannot contribute to our understanding of the social processes that are driven by interdependence, cohesion or social exchange.

Whilst SIT aimed to provide a framework for the understanding of inter-group discrimination, it did not take into account the valence of the outcomes with which individuals discriminated between the in-group and an out-group (Buhl, 1999; Mummendey & Otten, 1998). Later research showed that while in-group members may display in-group favouritism under these minimal conditions by giving the in-group more “rewards” than they gave to the out-group, they are unlikely to endorse out-group harm by “fining” the out-group more than the in-group. Thus, it has been argued that the psychological processes specified by SIT would be applicable to in-group bias, but not out-group derogation (Brown, 2000). Indeed, it has been suggested that variations in in-group positivity do not systematically correlate with negativity towards out-groups (e.g., Hinkle & Brown, 1990; but see Turner, 1999). This is also suggested by research on the positive-negative asymmetry effect of in-group bias (PNAE; Mummendey et al., 1992; Reynolds, Turner, & Haslam, 2000), explored in more depth in Chapter 2.

Self-Categorisation Theory

Self-categorisation theory (SCT; Turner, 1982, 1985, 1991; Turner et al., 1987) examines the relations between the self, social norms and social context. The central tenets are as follows. The self is comprised of a series of self-concepts or *categorisations*. These categorisations vary in their level of abstraction, from personal to social identity, and in their *content*. This content can change with different social contexts. An example may be useful to illustrate this idea. Imagine an individual, Bess. Bess is an illustrator, and in her spare time, a member of her local football team. When illustrating, she works in isolation. This is a context in which her personal identity would be salient. At this time, she considers herself an introspective, quiet and creative person. When

practising football with her team mates however, her “football” social identity is salient. It is normal for her and her team mates to act in similar ways to each other: loud, boisterous and competitive. When competing against another team – i.e., a slightly different context to club training- the individuals within Bess’ team, including herself, become more aggressive. In other words, the content of the “football” identity changes slightly.

According to SCT, whichever self-category is salient at a particular time is a function of perceiver readiness to use the category, its relative accessibility and its *fit* with stimuli. *Comparative fit* is defined by the metacontrast principle. The metacontrast principle states that, “A group of stimuli is more likely to be categorised as a single entity to the degree that the differences within that group (on relevant dimensions) are smaller than the differences between that group and other stimuli” (Turner, 1991; p. 156). *Normative fit* is defined by the congruence of the content of the category with behaviour.

The salience of group membership enhances the perceived similarities within, and the differences between groups (McGarty & Penny, 1988). SCT argues that as social identity becomes salient, the individual-level of abstraction considerably diminishes (depersonalisation). Therefore, individuals become aware of their identity either as an individual *or* as a group member (Turner, 1991; p. 155). Thus personal and social identity are seen as mutually exclusive (Spears, Ellemers, Doosje & Branscombe, 2006) – a principle known as “functional antagonism” (but see Postmes, Baray et al., 2006; Turner, Reynolds, Haslam & Veenstra, 2006; The *Inductive Model of Identity Formation* section).

According to SCT, this is how the process of social influence occurs. In this state, the self is defined and stereotyped (also known as “self-stereotyping”) in line with the content of the relevant social identity (Hogg & Turner, 1987b).

SCT and Social Influence

Therefore, according to SCT, when a social identity is salient, an individual's actions and behaviour should be guided by the content of that identity. It is assumed that if the identity is salient, then its content is perceived as valid. SCT argues that identity content is not fixed, but dependent upon inferences made from the socio-structural context (Doosje, Haslam, Spears, Oakes, & Koomen, 1998; Haslam & Turner, 1992; Turner, Oakes, Haslam, & McGarty, 1994). This has been labelled a "top-down" or "deductive" route to identity formation (Postmes, Haslam, et al., 2005; Postmes, Spears et al., 2005; see also Turner, 1982). It is expected that other in-group members will act in the same way, as a function of a shared social identity in a shared context:

"The perceived, expected or believed agreement of similar others in the same situation implies that our behaviour is a function of the *objective world* rather than our personal biases, prejudices and idiosyncrasies," (Turner, 1991; p. 161, italics added).

This will increase with the extent to which they see each other as in-group members, share the same context, are uncertain about the response and a correct response is important for the group (Turner, 1991). From this idea, grew a self-categorisation theory explanation for the group polarisation phenomenon.

The SCT Account of Group Polarisation

Referent informational influence (RII; Turner, 1982; Table 1.1) is a top-down, deductive process of social influence deriving from an individual's self-definition as a group member. This theory was created as a combination of Deutsch and Gerard's (1955) informational influence with Kelman's (1961) process of identification and French and Raven's (1959) referent power. RII differs from the other main theories of

influence by arguing that people conform to a category norm, not to the behaviour of specific others, or to validate reality or avoid social disapproval (Hogg & Smith, 2007).

Based on the ideas behind SCT, RII rejects the division of normative and informational influence, and attempts to answer Wetherell's (1987) theoretical questions (see *Critique* section) by reconciling the two accounts and pointing to the critical role of social self-categorisation in mediating group polarisation (Abrams, Wetherell, Cochrane, Hogg, & Turner, 1990; Turner, 1987b). This approach re-conceptualises the informational value (or persuasiveness) of a communication as the degree to which it is perceived to be prototypical of an initial in-group norm or consensus. In other words, what is seen as informative is normative. It is argued that the most persuasive in-group members will be those who are most prototypical. What is perceived as prototypical is defined by the metacontrast principle, and therefore varies with the socio-structural context. Thus, unlike the original normative and informational influence accounts, RII takes into account the effects of the inter-group context within which arguments are voiced (McGarty, Turner, Hogg, David, & Wetherell, 1992).

Table 1.1

The 3 Stages of Referent Information Influence (Turner, 1982)

-
1. Individuals define themselves as members of a distinct social category.
 2. Individuals form or learn the stereotypic norms of that category. They ascertain that certain ways of behaving are criterial attributes of category membership. Certain appropriate, expected or desirable behaviours are used to define the category as different from other categories.
 3. Individuals assign these norms to themselves in the same way that they assign other stereotypic characteristics of the category to themselves when their category becomes psychologically salient. Thus their behaviour becomes more normative (conformist) as their category membership becomes salient.
-

According to SCT's principle of metacontrast, in an inter-group context, individuals' judgement should polarise towards the position that accentuates the differences between the in-group and the out-group, whilst minimising intra-group differences (see also Mackie & Cooper, 1984; Turner, 1991). In a nutshell, RII theory explains group polarisation as, "Conformity, though self-categorisation, to a local in-group norm which is polarised as a result of the in-group being located towards an extreme of the salient comparative context or frame of reference," (Turner, Wetherell, & Hogg, 1989; p. 135). From this perspective, in-group members will move toward a more extreme position in order to maximally differentiate their in-group from the out-group, whilst doing justice to the different viewpoints that may exist within the in-group (McGarty et al., 1992; Turner et al., 1989). Thus, in contrast to the predictions of PAT, attitude shift should depend on the group membership of the information's source and not on the information *per se* (Mackie, 1986; Mackie & Cooper, 1984).

The most straightforward evidence for this theory would be bipolarisation, i.e., attitude divergence of groups in relation to each other. However, the evidence for this has not been forthcoming. Perhaps the strongest evidence for an SCT account of group polarisation comes from research on the social identity model of deindividuation effects (SIDE; Reicher, Spears, & Postmes, 1995), with some evidence of bipolarisation (Postmes, Spears, & Lea, 2002; Study 1). However, other research in the same line leads to the suggestion that categorisation alone could not account for the effects (Postmes, Spears, Sakhel, & de Groot, 2001; see *Theories of Collective Action* section). In addition to the empirical difficulties that have marred attempts to find strong confirmation of SCT's predictions for group polarisation effects, there are a number of theoretical issues that should also be considered.

For example, in RII, social identity is seen as a focusing lens, the output of which affects intra- and inter-group behaviour (Turner, 1982; p. 21). Therefore, SCT

requires self-categorisation as a group member (i.e., social identity salience) *before* any social influence can occur. While RII assumes that interaction impacts upon the content of identity, the dynamic nature of this process has not been explored. Instead, RII specifies that self-categorisation as a group member must come first (see Table 1.1, [1]). This could be interpreted as not allowing for intra-group processes to define identity *before* individuals become identified. Turner and Oakes (1997) specified that a social bond is a “*pre-condition* [...] for the validation of cognition through mutual social influence” (p. 360; italics added). Therefore, this approach depicts group polarisation as an effect of social category salience (although SCT argues that social identity content is dynamic, it focuses largely on contextual variations). But consider, for example, a collection of strangers in a doctor’s waiting room. They are largely uninterested in each other and do not share a salient self-category. Upon striking up a conversation, commonality may be discovered and defined, and an identity may be formed.

Yet according to SCT and RII, the changes occurring *within* groups which result in polarisation cannot feed back into the identity process – except by increasing salience. This perspective ignores the dynamics of small groups and their multiple functions to the individuals within them. However, there is evidence, conducted within the social identity tradition, that intra-group interaction feeds perceptions of in-group homogeneity over time (Oakes, Haslam, Morrison & Grace, 1995). Yet traditional social identity theorising ignores the fact that one of the subjects of group discussion may be the group itself. It seems plausible however, (and this thesis will demonstrate) that groups have the capacity to *transform* social identities by debating them.

Therefore, one limitation of SCT and SIT lies in the lack of attention to social identity contents and, in particular, to how they originate. Of course, social stereotypes are influenced by the comparative context within which they are considered (Haslam, Turner, Oakes, McGarty, & Reynolds, 1998). For example, the English drink a lot

compared with the French. Compared with the Scots, other national attributes are more prominent. These are essentially meta-contrasts which each of us can individually compute on the basis of the known features of the various groups in question. However, the origin of these known features is profoundly social. The implicit assumption in this approach is that each individual within the group perceives this reality in the same way. This is not as straightforward as it sounds (see the *Pluralistic Ignorance* section). This brings us to the main outstanding issue with the RII account of group polarisation: it states that, “Individuals form or learn the stereotypic norms of that category” (Table 1.1, [2]), but does not specify how this is achieved. The SCT account is quite explicit about how people respond to a given social reality, but it remains silent as to how that reality is established. We suggest that “fit” is a matter for debate, and that “accessibility” more often than not is “primed”, if not actively instantiated through conversation.

In sum, it is argued here that social identities are very rarely (if ever) entirely pre-determined outside of the individual, nor is their salience entirely a product of influences of the social context on that individual. For most, if not all social identities, we need to form, modify and inform our perceptions of them *socially*, i.e., through reaching a *social consensus* upon which these perceived social realities are predicated. This fact necessitates logically that the individual is an active participant in the definition of social identity contents. In creating the conditions under which they can become more or less salient, we construct identities and their importance through interaction with similar others.

Theories of Interdependence

The study of social influence is characterised by a dichotomy between group (social-normative) accounts like the social identity approach and inter-personal theories (e.g., interdependence; Postmes, Spears, Lee, & Novak, 2005). Theories of

interdependence were influenced by Festinger's (1950) group cohesiveness approach, and begin with the idea that groups are created by an attraction (Lott & Lott, 1965) or reliance between the members for an understanding of social reality.

Whilst recognising that the out-group is an important source of reference and that its effect on in-group bias is well documented, interdependence researchers have examined how group entitativity emerges from intra-group comparisons (Gaertner, Iuzzini, Guerrero, Witt, & Oriña, 2006; Gaertner & Schopler, 1998). This theoretical angle has its roots in Lewin's (1947) definition of the group as being "a dynamic whole based on interdependence" (p. 131). Ultimately, theories of interdependence allow individual group members to have agency over their social perceptions. They are not passive observers of an "objective" social reality but able to interact within it, and therefore have the power to socially construct and change that reality. Furthermore, interaction focuses not just on where the in-group is in relation to the social context, but where the group wishes to be in the future. Intra-group interaction thus can drive social change.

However, the social cohesion account of psychological group formation (Lott & Lott, 1965) has been criticised for relying on a process of inter-personal attraction (Turner, 1982; Hogg & Turner, 1987). The important lesson to take from the interdependence literature is the conceptualisation of the interaction between individuals and reality as *dynamic*, i.e., individuals can influence their perception of reality together. However, at the same time, this research takes away the concept of identity—the idea that there is a connection between self and group that is self-defining.

The Inductive Model of Identity Formation

Contrary to social comparison theory and theories of interdependence, both SIT and SCT focus largely on ways in which contextual comparisons with the out-

group engender social influence, group polarisation and in-group bias. However, it is clear that the actions and attitudes of individuals can be influenced by group members without reference to the wider social context (e.g., Lewin, 1947; Asch, 1956; see also Fielding & Hogg, 1997; Oakes et al., 1995). Indeed, Reicher, Haslam and Hopkins (2005) research on entrepreneurial leadership suggests that leaders can shape in-group members' social perceptions. Turner (1982) hints at such a "bottom-up" process of identity formation, building on Tajfel's (1959, 1972; as cited in Turner, 1982) distinction between "deductive" and "inductive" aspects of categorising. Turner (1982) defined "induction" as, "the means by which the criterial attributes of some category are inferred from one or more individual members" (p. 28). However, later conceptualisations of SCT focus on the top-down deduction of attributes from the basis of category membership, rather than this inductive route.

The original formulations of the interpersonal-intergroup continuum (Tajfel, 1978) and functional antagonism, appeared to imply that interaction between *individuals* cannot be reconciled with interaction in terms of a common or shared social identity (p. 34; but see Postmes, Baray et al., 2006;). This creates problems for the dynamics within small groups, as it implies that the interaction within such groups and its subsequent effects should be largely or entirely restricted to personal levels of identity. In other words, it seems that there can be no interaction between levels of identity (Spears, Ellemers, Doojse & Branscombe, 2006). However, a recent conceptualisation of functional antagonism by Turner, Reynolds, Haslam and Veenstra (2006) suggests that social and personal identity can be simultaneously salient. This allows for the possibility that inter-personal interaction can create or "induce" social identities for those people involved in a discussion (i.e., a possibility that interaction can be inter-personal, inter-group or *intra-group*). Indeed, several studies now suggest that group members do not need to perceive themselves as members of a homogenous group in order to

influence each other (Festinger, 1950; Festinger & Thibaut, 1951; Haslam, Eggins & Reynolds, 2003; Postmes, Haslam, & Swaab, 2005). Postmes, Haslam et al., (2005) concluded from their review of the evidence that the distinction between interpersonal and social influences violates the phenomenology of social influence, and therefore argued that the interpersonal and social identity explanations of group processes need not be mutually exclusive (cf., Gaertner & Schopler, 1998; Hogg & Hains, 1998). In other words, “Social identities in interactive groups can be based on idiosyncratic contributions of its members as well as on preexisting social categorisations,” (Postmes, Haslam et al., 2005; p. 747; see also Haslam, Eggins & Reynolds, 2003).

Theories of interdependence / social cohesion, and SCT differ as to whether or not they assume a linear causal path in the mechanism of group formation. In other words, SCT assumes that group membership evinces social attraction. In contrast, interdependence theories argue that social attraction leads to the perception that the individuals comprise a group. Do individuals need social attraction to want to self-categorise as a group member, or does group membership evince social attraction? Or, in the language of age-old philosophy, which came first, the chicken or the egg?

“Interaction is important not just as a medium to study the influence of self-categorisation processes (i.e., as dependent variable), but also as a theoretical process in self-categorisation [...]. This is partly (a) because self-categorisation depends on social consensus, (b) because interaction is pivotal to the mutual formation of personal and social identity in relation to each other, and (c) because interaction informs the content of those identities.” (Postmes, Baray et al., 2006; p. 216)

In other words, polarisation should not be only treated as a *result* of self-categorisation, but as a tool for identity (or “category”) formation. If seen in this way, the process of group polarisation establishes novel group norms, or identity content. With all of these outstanding conceptual issues, it appears an appropriate time for

Turner's (1982) pendulum to rest at the mid-point of the interdependent and cognitive theories, within a new, interactive metatheory.

The key limitation of the literatures outlined above, is that they are often focused on the effect of cognitive input on individuals' decisions, attitudes, and actions. What the majority of this research ignores (and what the aims of this dissertation imply) is that first, *we* go through the process of making up *our* minds, and then this provides the *individual cognitive resources* that are required to act within that social reality. In other words, our social actions are not merely the consequence of our individual cognitions, they are an essential precondition for those individual cognitions and even individual cognitive processes to occur in the first place.

At the heart of the inductive model of identity formation, (IMIF; Postmes, Haslam et al., 2005; Postmes, Spears et al., 2005) is the assumption that social identities, although represented in individual cognition, are socially shared and *socially constructed* entities. The IMIF recognises that fellow in-group members constitute a social and ideological resource for socially-identified individuals to discover, negotiate and validate their belief systems (see Correll & Park, 2005; Festinger, 1950, 1954). Therefore, this model outlines a "bottom-up" or "inductive" route to identity formation via communication at the intra-group level (Drury & Reicher, 2000; Postmes, Haslam et al., 2005; Postmes, Spears et al., 2005; Reicher, 1987; Turner, 1982).

This is important to how we conceptualise motivations for social action: Often, what is morally justified and legitimate for the group to do will not follow immediately from the cognitive perception or "deduction" of inter-group relations (where *we* stand vis-à-vis *them*). Especially when it comes to engaging in a confrontation, any successful collective action demands that actions and objectives be negotiated with, consensualised upon, and validated by other in-group members, at least to a certain degree. A "united front" is critical for the success of social action.

The Influence of Identity Content on Inter-group Relations

The literature reviewed above leads to the suggestion that individuals do not engage in reality testing in isolation, but through the processes of collective discussions and interpretations. SCT and RII are premised on the assumption that social influence is driven by the expectancy of consensus. It has been argued that when there is a perception of consensus on social stereotypes, they are transformed into *subjectively valid* beliefs (Haslam, Turner, Oakes, McGarty, & Reynolds, 1998) and this can influence attitudes and behaviour (Sechrist & Stangor, 2001).

Once they are considered subjectively valid representations of reality, socially shared stereotypes become not simply a set of traits, but powerful, functional, dynamic and value-laden constructs which define and justify the position of groups in the socio-structural context, reflecting ideologies and a particular world view (Haslam, 1997; Jost & Banaji, 1994; Reicher, Hopkins, & Condor, 1997; Reynolds, Oakes, Haslam, Nolan, & Dolnik, 2000; Tajfel, 1981). This understanding of social reality is necessary for individuals to be able to plan collective action effectively (Reicher, et al., 1997). Therefore, stereotypes can be functional: They may include implicit prescriptions for appropriate collective behaviour toward the stereotyped group (Haslam, 1997; Reicher et al., 1997). Within conducive socio-structural conditions, their content can be used to challenge social hierarchies (Reynolds et al., 2000) and propel social protest (Stott & Drury, 2004).

Stereotype Consensus

A series of studies have shown that intra-group interaction can increase stereotype consensus in the group normative direction (*consensualisation*; e.g., Haslam, Oakes, Reynolds, & Turner, 1999). The extent to which interaction was successful in doing this however, was crucially dependent upon social identity salience (Haslam, 1997; Haslam, Turner, Oakes, McGarty et al., 1998; Haslam, Turner, Oakes, Reynolds

et al., 1998; Haslam et al., 1999). Following SCT, this is arguably because social identity salience should enhance in-group members' influence on each other by promoting perceptions of their homogeneity (via self-stereotyping) and thus increasing their capacity for mutual validation (Haslam, Turner, Oakes, McGarty et al., 1998). This is true for the in-group's self-stereotype as well as the in-group's stereotype of the out-group (Haslam, Turner, Oakes, Reynolds et al., 1998). It could be concluded from these studies that identity content can be shaped simultaneously by "top-down" inferences from social categorisations and the socio-structural context, and in an emergent, "bottom-up" manner through intra-group interaction. However, Haslam and colleagues interpreted their results as stemming largely from increased social identity salience, which impacted upon identity content and enhanced intra-group consensus – a "top-down", unidirectional path.

Stott and Drury (2004) gathered evidence that the stereotype consensus reached through intra-group interaction can impact upon inter-group action. One conclusion that may be reached is that consensual stereotypes can operate as psychological tools which groups use to mobilise themselves (Haslam, Turner, Oakes, Reynolds, & Doosje, 2002). Indeed, Tajfel (1981) argued for the integration of the social and individual functions of stereotypes. He believed that groups can use stereotypes to justify their actions against other groups. Therefore it seems, as posited at the beginning of this section, that stereotypes could contain shared behavioural norms – prescriptions for (potentially hostile) group-based behaviour. However, sharing and validating the stereotype is critical.

In-group consensus surrounding the out-group stereotype can be used as a platform to justify and legitimise inter-group violence (Haslam, Turner, Oakes, McGarty et al., 1998; Tajfel, 1981). For example, during the Wannsee conference, on January 20th, 1942, several high ranking Nazi officials met to discuss the "solution" to

the “Jewish question”. They entered the discussion with shared stereotypes, and exited with a consensual plan for hostile action against the out-group – the infamous “Final Solution”.

“In order for perceivers to have confidence in the correctness of their beliefs – and therefore use them as a basis for coordinated and subjectively meaningful social behaviour – those beliefs need to be validated by other people [...] Consensus is an outcome and expression of an identification process that is capable of mobilising and motivating group members in a way that independent activity and an individual product never could.” (Haslam, Turner, Oakes, McGarty et al., 1998; p. 232)

However, the image of consensualisation (or interaction) as a process through which *existing* stereotypes become more shared (or are, in other words, disambiguated) is a limited conception. Consensualisation must also be the process through which social realities (not just stereotypes, but also their validity) are created and transformed. As realities, they can become the foundations for action, in a way that is not always possible with purely individually held cognitions.

Recent Advances. Paralleling the ideas that the content of identities can be constructed through intra-group interaction (the IMIF), McGarty and colleagues have demonstrated that comparable bottom-up processes of identity formation (or at least elaboration) may increase commitment to *positive* social action in opinion-based groups (the opinion-based group interaction method, or OBGIM; Blüch, McGarty, Reynolds, & Muntele, 2007; Thomas & McGarty, in press).

McGarty and colleagues have demonstrated that intra-group interaction can increase commitment to a cause for which in-group members have already indicated support. However, if the intra-group interaction is truly capable in theory of causing identity formation and / or identity content *change*, it should be possible to *qualitatively transform* in-group norms. In other words, rather than making individuals’ views more

committed or strengthened in the direction in which they originated, (akin to group polarisation; Moscovici & Zavalloni, 1969), an *alternative* identity content could potentially emerge.

ACTION

“The prevalent theory in psychology assumes action to be the direct result of motivation. I am inclined to think that we shall have to modify this theory. We shall have to study the particular conditions under which a motivating constellation leads or does not lead to a decision, or to an equivalent process through which a state of ‘considerations’ (indecisiveness) is changed into a state where the individual has ‘made up his mind’ and is ready for action, although he may not act at that moment.” (Lewin, 1953; pp. 288-289)

At this juncture, it is helpful to return to the aims of this thesis. At the beginning of this chapter, I made reference to atrocities such as the Rwandan genocide, and intimated at the role of communication in helping to provoke and perpetuate inter-group violence. We have already seen that group discussion could qualitatively transform individuals’ conceptions of social reality, and that consensualisation of out-group stereotypes may help to justify inter-group behaviour. In this section, I review prior research which investigated the other possible antecedents of social conflict, and then outline how social interaction may influence the decisions of individuals to engage in such collective action.

As Lewin’s quote above illustrates, in the social psychological literature there has been a lack of consensus, or at least a lack of theoretical integration, of the different theories of why people choose to act. This has caused divergence of the attitudes, social influence, and collective action literatures. Each of these is outlined below.

Attitudes in Action

One of the central areas with which attitude research has been concerned is the relationship between attitudes and behaviour, i.e., can behaviour be predicted from

attitudes? Do attitudes inform behaviour? The theory of reasoned action (TRA; Fishbein & Ajzen, 1975) was built on the assumption that social norms, as well as personal attitudes (i.e., intention to act) contribute to action. According to the TRA, personal attitudes (towards the behaviour) plus the subjective norm (the person's belief about whether salient others believe he or she *ought* to engage in the behaviour) influence intentions to act. However, although normative support does improve the correlation between attitudes and behaviour, this effect is “surprisingly small” (Hogg & Smith, 2007; p. 109). Does this necessarily imply that knowledge of the subjective validity of a norm is not important for social action? Or do individuals require something different / additional in order to act as a group member? The research reviewed in the previous section would certainly imply that self-categorisation as a group member is a necessary condition for the influence of the norm on the individual.

Following from Festinger's (1950) description of the subjective validity of a norm (see *Definition of Terms*), the theory of reasoned action (TRA: Fishbein & Ajzen, 1975) and the theory of planned behaviour (TPB; Ajzen, 1985) define the subjective norm as the amount of pressure that people feel from similar / significant others to act in a certain way. More recently, there has been a distinction drawn between injunctive (how people should behave in a given situation) and descriptive (how people actually behave in a given situation) norms (Cialdini, Kallgren, & Reno, 1991; Cialdini, Reno, & Kallgren, 1990; Smith & Louis, in press). These types of norms have differential influence on behaviour: Injunctive norms arguably promote social change, while descriptive norms promote the status quo.

In the TRA, only a superficial distinction is made between personal and social (normative) beliefs about a behaviour (Eagly & Chaiken, 1993; Terry & Hogg, 1996). The TRA, as it stands, is really only relevant to predicting the behaviour of the individual, and is less suited to explaining collective, or cooperative action (Liska, 1984).

In order for this to improve, social factors need consideration. As the attitude literature stands, it harps back to the individualistic metatheory of much of the social psychological research into group behaviour. Furthermore, '*intention*' is not a well-articulated concept. Intentions vary on a continuum from vaguely formulated thoughts to clear-cut plans to engage in action at a particular time (Sternberg, 1990).

The theory of planned behaviour (TPB; Ajzen, 1985, 1991), as an alternative to the theory of reasoned action, includes consideration of the amount of "control" one has over their behaviour – or rather, in terms of self-efficacy (Bandura, 1977), how "possible" the behaviour is. The research reviewed above clearly shows that this can be informed through interaction, i.e., when an individual's attitude becomes based on/rooted in group norms and is socially delineated, it is qualitatively transformed into *social behaviour*. It has been shown that consensualisation processes can inform private judgments, shaping normative attitudes and importantly, what is considered *possible*.

Interaction in Action

Although the collective action literature has been treated independently to stereotyping and polarisation research, this fracture of topics has done the study of action a disservice. Given the finding (from group polarisation research), that attitudes can become more polarised and hostile through social interaction (e.g., Myers & Bishop, 1970), it seems logical that similar processes may play a role in collective actions, such as the escalating behaviour of an angry crowd (Reicher, 1987). It seems there is no clear reason why the processes discovered in small interactive groups may not be applicable to action that occurs in larger psychological groups – and the hypothesised processes theoretically integrated.

For example, it may be that overt inter-group hostility is perpetrated in the assumption of an in-group consensus behind it. In comparison, overt in-group favouritism can seemingly exist without such assumptions (i.e., it can be individually-

based as much as is inter-group behaviour) and therefore, was displayed in a minimal group context (while out-group punishment was not). This qualitative difference between out-group punishment and in-group favouritism (e.g., Struch & Schwartz, 1989) underlines the importance and relevance of collective action for this research. I recognise that individuals can only challenge other groups in a very limited way if on their own, and need the support and validation of other in-group members in order to derogate and punish out-groups. Thus, hostility towards another group becomes much more likely when such actions can be taken collectively (or at a minimum, are perceived to be consensually valid). In the absence of such tacit support, the hostile actor risks the wrath of both in-group and out-group members. Thus, hostile social action may become *aggressive collective action*, through a process in which the emergence of social norms may play a key role (e.g., Turner & Killian, 1972; Reicher, 1984, 1987). In contrast, in-group favouritism could be perpetrated with minimum fuss: in-group members would tend to favour this kind of behaviour, and this is not easily challenged by out-group members.

Wright, Taylor, and Moghaddam (1990) saw stereotyping as integral to the process of social change through collective action, an argument that was supported by the evidence of Reynolds, Oakes, Haslam, Nolan, & Dolnik (2000) and that is described in the *Stereotype Consensus* section. After all, as described in the opening quote of this chapter, a sense of collective purpose does not arise in a vacuum (Tajfel, 1972), but through the development of a common ideology and shared understanding of reality (Postmes, Baray et al., 2006). Historically, the only way in which this could be developed was through intra-group communication (cf. Festinger, 1950). Therefore, it is argued that it is useful to study inter-group action by looking at processes within the in-group.

*Theories of Oppression*¹

Theories of collective action have focused mainly on the behaviour and motivations of lower status groups, who feel a sense of injustice (Postmes & Smith, in press; Van Zomeren, Postmes, & Spears, 2008). At least part of the reason for this relative neglect for the motives of oppression (that is, collective action by higher status groups) is that it is often considered a *normal* or even *natural* phenomenon: where groups compete for scarce resources, oppression appears functional on both evolutionary and economic grounds. Indeed, the idea is common to perspectives in sociology, philosophy, political science, history, economics, and social psychology (Bobo, 1999; Sidanius & Pratto, 1999, for reviews). Of course, the dominant group's position provides them with privileged access to the necessary means, ensuring the efficacy of their enterprise.

Realistic group conflict theory (RCT; Campbell, 1965; Sherif, 1956, 1967; Sherif & Sherif, 1956) states that conflicts of interest between groups result in group conflict (Campbell, 1965). Competition for scarce resources was found to promote competition and inter-group conflict (Sherif, 1956). Three contemporary perspectives to explain oppressive social action were built upon RCT: social dominance theory (SDT; Sidanius & Pratto, 1999), systems justifications theory (SJT; Jost & Banaji, 1994), and SIT. They agree that inter-group threat produces oppression, but advance different ideas about the processes involved. Briefly, SDT (Sidanius & Pratto, 1999) suggests that some people naturally hold the view that status differences are good—they have a “social dominance orientation.” Individuals with such views would be more likely to endorse actions to preserve status differences when under threat. SJT is not incompatible with this idea. It proposes that humans tend to justify the current social

¹ This sub-section on theories of oppression is adapted from Postmes & Smith (in press)

system, including the position of their group (Jost & Banaji, 1994). For high status groups, SJT implies that oppression is a straightforward way of satisfying a need to maintain the status quo (see also Blumer, 1958). Jost and Banaji (1994) argued that stereotypes fulfil an ideological role in the domination of lower status groups by higher status groups, by justifying inter-group exploitation (see *Stereotype Consensus* section).

Finally, SIT sees inter-group threat as the main precursor to oppression (Haslam, 2004; Tajfel & Turner, 1979). SIT is more sophisticated in acknowledging that outright social competition is rare, and is but one strategy for combating (or presumably maintaining) inequality. Nevertheless, it proposes that threats to the status quo should result in a response from the high status group to restore or preserve positive in-group distinctiveness. In sum, although all these theories emphasise very different processes and variables, they all agree that oppressive social action results mainly as a functional response to inter-group threat.

One problem for theories of oppression is that the relationship between status threats and oppression is more elusive than is assumed by these theories. The complex relationship between threat and prejudice is evident in recent experimental research. High status group members who are threatened with future deprivation show some oppressive tendencies, but when they look forward to future gratification (status increases), oppression increases more sharply (Dambrun, Taylor, McDonald, Crush, & Meot, 2006; Guimond & Dambrun, 2002). This finding resonates with historical research suggesting that inter-group struggles are particularly acute when the economic tide is rising after having been low (Rudé, 1964; Thompson, 1971; Tilly, Tilly, & Tilly, 1975). In sum, the relation between inter-group threat and hostile social action is not as straightforward as is often assumed. One reason for this, I suggest, is that there are multiple motives for oppression. One of these is individual group members' orientation to local in-group norms (Postmes & Smith, in press). If these norms

support oppression, a group may oppress, relatively independent of the level of threat experienced.

Theories of Collective Action

“No mere human can stand in a fire and not be consumed.”

(Christabel LaMotte in *Possession: A Romance*; Byatt, 1991)

Theories of collective action have begun to integrate the dynamic view forwarded above, of why individuals are motivated to act. Early theories of action in social psychology were heavily influenced by Le Bon (1895/1947), who coloured perceptions of the revolutionary crowd as a primitive force comprised of individuals who had lost their sense of self through “submergence” in the crowd, and were thus freed from the shackles of personal responsibility:

“By various processes an individual may be brought into such a condition that, having entirely lost his conscious personality, he obeys all the suggestions of the operator who has deprived him of it, and commits acts in utter contradiction with his character and habits,” (Le Bon, 1895; p. 32).

This concept underpins later theories of deindividuation, which argue that revolutionary social action results from a “psychological state of decreased self-evaluation and decreased evaluation apprehension causing anti-normative and disinhibited behaviour,” (Postmes & Spears, 1998, p. 238; Diener, 1976, 1979, 1980; Festinger, Pepitone, & Newcomb, 1952; Zimbardo, 1969).

Deindividuation theory has often been used to explain social atrocities such as terrorism or genocide (Staub, 1996; Staub & Rosenthal, 1994). However, in Staub’s (1997) classic thesis on genocide and group violence, he observes that group members often willingly accept sinister group norms, goals, views and ideology: people *join* rather than simply obey out of respect. This implies that the out-group hostility that ensues is a choice, rather than the result of a mere loss of self. Therefore, it appears that rather

than being disinhibited, the behaviour could be oriented to the local norms of the in-group. Therefore, like with research into oppression (Postmes & Smith, in press), it is important to consider how the motivations of the individual group members evolve in relation to the intra-group dynamic.

In this vein, emergent norm theory (Turner & Killian, 1972) posited that novel collective behaviour emerges from a normative crisis, depending upon how a precipitating event is collectively interpreted by participants. The crisis makes traditional normative guidelines for behaviour obsolete, and the sense of uncertainty and ambiguity that follows compels people to act to create meaning through interaction. This enables a new, emergent normative structure.

The idea that norms can explain collective action is a key assumption of the social identity model of deindividuation effects (SIDE; Reicher, Spears, & Postmes, 1995), which focuses on the cognitive and strategic factors involved in identity definition and enactment (see especially Drury & Reicher, 2000). According to SIDE, individuals within crowds act in normative and controlled ways. Different to emergent norm theory, it specifies that these norms are specific to the social identity of the crowd. Thus, rather than experiencing a loss of self (as deindividuation theory and Le Bon proposed), the individual re-categorises from personal identity to a common social identity with the crowd.

The empirical tests of SIDE in computer-mediated communication (CMC) environments provide confirmation for SIDE's prediction that anonymous individuals experience higher social identity salience and enhanced conformity to group norms (in conducive conditions; see Reicher et al., 1995). This finding demonstrates that the co-presence of group members may be unnecessary for the occurrence of social influence. It is also much more consistent with the cognitive underpinnings of social influence as proposed by self-categorisation theory (e.g., Turner, 1982), and much less consistent

with the proposals of interactive and / or interdependence models (e.g., Asch, 1952; Homans, 1961; Shaw, 1976; Zander, 1979).

However, later SIDE research demonstrated that social interaction in anonymous groups could produce changes in social identity content (e.g., Sassenberg & Postmes, 2002; Postmes, Spears, & Sakhel, & de Groot, 2001). This is subtly different to earlier SIDE research (Spears, Lea, & Lee, 1990), which focused mainly on the social influence stemming from salient social identities whose content had been defined *a priori*, e.g., by means of false feedback about group norms.

“Our results also make clear that the group is not just a mere cognitive concept [...] Interaction and communication between individuals serves to establish and strengthen a local group norm (Haslam, 1997). In this sense, the group is both cognitive and social at the same time. We argue that social norms can be induced from social interaction and that identification with the group is a prerequisite for such norm construction to occur,” (Postmes et al., 2001).

Therefore, rather than being entirely based upon self-categorisation processes, later SIDE research (particularly that focusing on the strategic SIDE) increasingly acknowledged the relevance of intra-group processes. For example, research showed that the perception of social support engendered through CMC can produce social action (Spears, Lea, Corneliussen, Postmes, & Ter Haar, 2002).

A similar shift to considering the dynamic development of social identity contents through inter-group interaction can be seen in the formulation of the Elaborated Social Identity Model of Crowd Behaviour (ESIM; Drury & Reicher, 2000; Reicher, 1996, 1997a, 1997b; Stott & Reicher, 1998). ESIM argues that by analyzing crowd events as developing interactions, one can account for both social determination and social change in collective action. Although the focus in the ESIM is on the interactions between groups, the fact that interactions are crucial for change to occur, and that social identity content is dynamic and dependent upon social interaction, is

complementary to the arguments forwarded in this thesis. For example, Drury & Reicher (2000) argue that an unexpected inter-group dynamic can emerge if crowd members hold a different understanding of their social position to that held by an out-group (e.g., the police), and if the out-group has power over crowd members and can wield this power to enact that understanding. This new inter-group dynamic can then affect the identity content of crowd members. Therefore, identity is conceptualised as a process, rather than a static given entity. Importantly, the perceived legitimacy of certain actions is a fluid part of this process.

SUMMARY

Early research showed that group discussion can increase commitment to a group decision, and polarise opinions. Two explanations were forwarded for these effects: normative and informational influence. However, these theories do not account for contextual variance, or explain why certain information is considered valuable or persuasive for particular groups. Rather than being seen as distinct paths to polarisation, the two theories of influence have been combined within theories of group behaviour.

Theories of group identity and interdependence were then discussed. Theories of interdependence posit that social influence occurs because individuals look to others for validation of their understanding of reality. This approach has been criticised as it is based on inter-personal attraction, which fails to take the social context – or in fact group identity – into account. Instead, the social identity approach assumes that the presence of an in-group necessitates the presence of an out-group, and therefore the processes about which it theorises are based within the socio-structural context. According to this approach, the extent to which interaction is persuasive or catalyses

identity change therefore, depends upon socio-structural conditions. This reduces interaction to a variable whose impact depends upon self-categorisation.

However, the inductive model of identity formation and findings from small-group research suggests that intra-group interaction can also shape the content of self-stereotypes, including out-group stereotypes. These stereotypes have the power to legitimise social action as they contain ideologies about the social system and therefore, implicit norms for action. Therefore, shaping identity content, including stereotypes, through intra-group interaction, may have significant consequences for social action. Both theories of oppressive and collective action are beginning to recognise the dynamic evolution of action.

THE CURRENT RESEARCH

Let me take you back to the opening quotes. Can existing polarisation accounts really explain the power of an SMS to turn an “average” person into a rioter? Or the power of a radio broadcast to legitimise the start of genocide? Is this really because of an opinion shift? Or have changes occurred at an identity level, involving longer term behavioural shifts, and permeating many contexts? I argue that an integrated approach to social action, which takes into account intra-group interaction and social identity, would more satisfactorily explain these social phenomena.

Let us consider the examples at the beginning of this chapter. The stereotypical and ideological content of the Hutu extremists’ radio broadcast was the foundation of their subsequent hostile social action. The White Australian riot protagonists in Cronulla used a widely circulated SMS to provoke, justify and *make normative* supporting and engaging in violence against Lebanese Australians. It has been argued that the protagonists attempted to use the SMS to align the norms of mainstream White Australians with their own far-right views (Bliuc, Hartley & McGarty, 2008). In both examples, I argue that the intra-group communication had the potential to impact upon

in-group members' perception of social reality, and transform the way in which they understood what was legitimate, valid, and justifiable social behaviour.

The unifying theme of the empirical chapters presented in this thesis is that intra-group interaction impacts on in-group identity content, and this content is the foundation of social action and social behaviour. It is by studying the processes by which social identities emerge and evolve through intra-group interaction that we seek to advance how in-group norms for inter-group behaviour come into being, and why they perpetuate or transform. Through this, the research hopes to contribute to our understanding of inter-group behaviour in general.

This introductory chapter has suggested that intra-group interaction has the power to transform in-group identity content from the "bottom-up", and in doing so, aids the development of norms for inter-group action. It is suggested that it would be conceptually insufficient to theorise about social action without studying in-group identity content. Chapter 2 is concerned with the discrepancy in traditional research, which suggests that the action of favouring the in-group (or "in-group love") does not necessarily lead to action which punishes the out-group ("out-group hate"; Brewer, 1999; Struch & Schwartz, 1989). This chapter outlines contrasting research plus events which suggest that under certain conditions, in-group ideologies clearly do come to legitimise and justify out-group punishment (Amiot & Bourhis, 2003). Against this backdrop, the empirical research in Chapter 2 suggests that intra-group interaction can lead to the formation of in-group norms which legitimise such punishment.

Chapter 3 continues to study hostile action towards out-groups as a result of intra-group interaction, by comparing the effects of intra-group interaction to those of individual reflection and measuring the development of consensual in-group norms for action. Findings suggest that consensualising within groups on out-group stereotypes (rather than on irrelevant topics or action plans) is most effective at qualitatively

transforming the in-group norm for action through social validation processes. As the establishment of consensus on social reality through interaction is suggested to play a crucial role for individuals' decisions on how to think and behave, this is investigated in Chapter 4. Through controlling the content and degree of consensus reached in a simulated interaction, the studies reported in Chapter 4 suggest that high consensus about norms for action can legitimise and increase action intentions.

In the last empirical chapter (Chapter 5), we investigate the impact of intra-group interaction (versus individual reflection) on a different form of social behaviour (performance), within the stereotype threat paradigm. This aimed to extend the scope of the investigation of the impact of interaction in transforming in-group stereotypes, and on both in-group and out-group behaviour, relative to each other. This chapter continues to provide evidence that it is stereotype content and the associated implicit behavioural norms that are challenged, transformed and consensualised upon through interaction.

In the final chapter (Chapter 6), I draw some conclusions from the empirical research presented in this thesis, discuss their theoretical implications, and answer pertinent questions raised by this work. It is important to note that Chapters 2, 3, and 5 were originally written as stand-alone manuscripts for publication. Whilst this has the advantage that each chapter can be read on its own, there is a risk of some repetition of the central theoretical argument, literature reviewed and definitions. The manuscripts have been adapted to keep this overlap to a minimum.

CHAPTER 2²

Intra-group Interaction and the Development of Norms which Promote Inter-group Hostility

Social psychological research has demonstrated a robust tendency for humans to favour the in-group over out-group(s) (Allport, 1954; Tajfel, Billig, Bundy, & Flament, 1971). However, although in-group favouritism is commonplace, action to deliberately harm or disadvantage an out-group in laboratory research is relatively rare (Bettencourt, Dorr, Charlton, & Hume, 2001). This is corroborated by research on the positive-negative asymmetry effect (PNAE), which demonstrates that the tendency to favour one's in-group over an out-group when making "positive" allocations (such as distributing rewards) is strongly attenuated when people are asked to make "negative" allocations (such as punishments or fines, Struch & Schwartz, 1989; Mummendey et al., 2002). At the same time, it is undeniable that out-group punishment does occur, and is not uncommon outside of the laboratory. In order to address this incongruity, the present chapter investigates the conditions under which a tendency to favour one's in-group can be transformed into "out-group hate" (Brewer, 1999). Specifically, the research seeks to identify conditions and processes that foster the development of

² This chapter is an adapted version of Smith and Postmes (in press). Study 2.2 was included in my dissertation submitted to the University for the degree of Master of Science in Social and Organisational Psychology, completed as part of a four-year Economic and Social Research Council PhD scholarship. The manuscript has been re-written however, new data collected (Study 2.1) and the data from Study 2.2 completely re-analysed using a different statistical method. This is compatible with the University's Statement of Procedures for the Presentation of Theses/Dissertations for the Degree of PhD, [2.1].

socially shared ideas that it is acceptable or even desirable (i.e., normative) to punish the out-group.

The PNAE and its Causes

The asymmetry in allocations of positive and negative resources has been widely replicated and several explanations for it have been advanced (Mummendey & Otten, 1998, for a review). One explanation is that divisions of rewards and fines are not likely to be psychologically equivalent acts. Psychologically, favouring the in-group is clearly very different to punishing the out-group: Punishment is a hostile act, and requires considerable justification. By comparison, favouritism is generally less frowned upon, and is more easily justified. As a result, punishment may be undesirable according to generic social norms, and in-group favouritism relatively more appropriate and normative (Mummendey & Otten, 1998).

However, in light of the blood-ridden history of the human race, the concept of a generic social norm against out-group punishment appears rather questionable. At various points in history, and perhaps most so in recent history, the “fundamental unwillingness” to harm others seems to have been replaced by a profound and widely shared willingness to inflict harm (e.g., Hobsbawm, 1995; Goldhagen, 1997). Indeed, the traditional PNAE has not always been replicated (Amiot & Bourhis, 2003). In fact, a *reversal* of the PNAE has been demonstrated through a simple re-framing of the allocation task (Sassenberg et al., 2003).

Therefore, it appears there are certain conditions under which group ideologies may come to *legitimise* and *justify* out-group punishment (Amiot & Bourhis, 2003). What we do not know is *when* and *why* it can become acceptable or desirable to harm the out-group. Against this backdrop, the present research examines how norms develop which legitimise such punishment.

Consensualisation and Norm Formation

Although there may be norms that support a fundamental unwillingness to harm out-groups (e.g., Mummendey & Otten, 1998), it is also possible that norms could develop which are then *used* by groups to legitimise and coordinate behaviour in a particular inter-group context (see also Amiot & Bourhis, 2005). In normal cases, direct negative action towards an out-group may not be required, and may even be counterproductive (see also Turner, 2005). However in some cases, such as when the out-group explicitly obstructs the advancement of perceived in-group interests (Sherif, 1967), there can be an emergent intra-group consensus which justifies that the in-group protect their interests “at all costs” (M. E. Turner, Pratkanis, & Samuels, 2003; Haslam, 2004, p. 111). This process of reaching a shared group perspective appears to be more than a process of cognitive inference (or "deduction") by individuals. Many studies suggest that social interaction plays a central part in the process of reaching consensus, and that it is related to the formation (or articulation) of a sense of shared identity and in-group norms (Postmes, Haslam, & Swaab, 2005; Postmes, Spears, Novak, & Lee, 2005; Swaab, Postmes, van Beest, & Spears, 2007; see Chapter 1).

So the question is, what occurs during this “consensualisation” process to mobilise hostile inter-group action? There is evidence that in-group and out-group stereotypes become more consensual through interaction (Haslam, Turner, Oakes, McGarty, & Reynolds, 1998), and this has been shown to impact upon inter-group behaviour (Haslam et al., 1998; Haslam, Turner, Oakes, Reynolds, & Doosje, 2002). Furthermore, Stott and Drury (2004) demonstrated that consensualisation around negative stereotypes fed preferences for conflict. This may be related to the phenomenon that intra-group interaction tends to make individuals’ views more extreme in whichever direction they are already tending (group polarisation; Moscovici & Zavalloni, 1969).

Intra-group interaction is likely to bring about more than the development of individuals' perceptions of inter-group stereotypes, however. For example, if the group reaches consensus on a viewpoint, this carries with it a sense of social validation (see Chapter 1). Social validation has normative implications for group members: If consensus can be attained about how to, say, punish an out-group member, the very fact that agreement can be reached about this action can be perceived as a signal that it is justifiable (Postmes, Haslam et al., 2005; Turner, 1991). Furthermore, consensus and social validation serve as strong indicators that other in-group members would support the action, which is an important consideration when it comes to deciding to engage in inter-group hostility (van Zomeren, Spears, Leach, & Fischer, 2004). Through these processes, intra-group consensualisation could mobilise group members above and beyond the extent to which individual thought processes could do this alone.

Therefore, group discussion could be a vehicle through which in-group members not only gain the knowledge and validation that they are experiencing a situation in the same way, but can explicitly develop an articulate, consensual in-group norm for inter-group behaviour. This level of consciously validated, shared in-group meaning is achievable through the sharing of views in discussion, but of course, intra-group interaction may not be the only way in which consensualisation could occur. Alternatives include exposure to mass media messages, consensus feedback, exposure to behavioural models, and so on. This is not inconsistent with the argument that interaction is a key process in the formation of norms that promote hostility. After all, these alternative ways of exerting social influence are not unrelated to the interactive processes in small dynamic groups, although the functional origins of the process must lie in the simple exchange of viewpoints within a group (see also Caporael, 2001; Postmes et al., 2006).

This focus on interaction within small groups is consistent with metatheoretical assumptions made by Asch (1958) and Turner (1987), and resonates with the arguments made by Reicher and colleagues concerning the relationships between (inter)group dynamics and social identity (e.g., Reicher, 1996; Stott & Drury, 2000) and social stereotypes (Stott & Drury, 2004). However, it also emphasises that in-group members orient their actions towards other in-group members as much as towards the out-group. Only in the rare case of complete social isolation do people orient themselves only towards generic social norms. More commonly, people attend to more local (emergent) in-group norms, which may on occasion deviate sharply from general social standards (Postmes & Spears, 1998). Removing this localised intra-group dynamic can seriously alter the social conditions which can give rise to discrimination. It is argued therefore, that hostile group behaviour, and therefore phenomena such as the PNAE cannot be understood or adequately theorised outside the intra- (and inter-) group context within which they occur, i.e., inter-group behaviour becomes dependent on the intra-group dynamic.

This theoretical background leads us to the central prediction in this chapter: That intra-group interaction has the potential to overcome generic norms against out-group punishment in minimal group settings. More specifically, if there is a reluctance to administer punishments in inter-group allocation tasks, then intra-group discussion should be a vehicle that can help group members create a normative climate in which such punishments are justified.

In-group Obstruction as an Aggravating Condition

It should be stressed that intra-group processes are contingent on the inter-group environment. When groups discuss inter-group issues, socio-structural conditions will mark the boundaries of what social constructions are possible (i.e., as

reality constraints). Inter-group relations will thus influence the norms that emerge within groups during intra-group interaction.

The design of this research therefore, took into account that an antagonistic inter-group setting may be required before punishments could begin to be legitimised. Thus we created an experimental situation in which the out-group would be perceived to obstruct the advancement of in-group interest (i.e., similar to negative interdependence, see Sherif 1967; Wright, Taylor, & Moghaddam, 1990; Ellemers, 1993). Obstruction is about the clashing of group trajectories – not a static social context but one in which both groups are on a path of action. Implicit in the manipulation of obstruction is that action is necessary to restore the inter-group balance. This inter-group dynamic would provide arguments to justify, and therefore legitimise, the out-group punishment, i.e., provide “fodder” for intra-group consensualisation around an aggressive norm. Thus, we reasoned that if the out-group were perceived as an obstacle, the process of consensualisation around a hostile norm would be accelerated.

In sum, we predicted that if the in-group is obstructed by the out-group, they would display discrimination to redress the balance of power. In the context of this research therefore, obstruction elaborates the concept of negative interdependence (cf. Sherif, 1967). It guides “who we are” and “who we want to be” in relation to an out-group. In this way, obstruction may shape in-group identity via intra-group consensualisation, and thus also normative group behaviour.

STUDY 2.1

The significance of the current research hinges on the psychological perceptions of positive and negative allocations. Our assumption (and that in the PNAE) was that these are perceived as *rewards* and *punishments*, respectively. In order to test this base-rate assumption, a pilot study was conducted. It was hypothesised that

negative allocations would be perceived as having a more negative impact on the recipient than positive allocations, and they would be less pleasant to allocate in a minimal group context.

Method

Participants were 20 volunteers (Mean age=24.14, 11 females). The pilot had a 2 condition (Valence of monetary allocation: Positive vs. negative) within-subjects design. Participants were asked to imagine that they were members of group A, and were given a sum of money to allocate to members of group A and members of another group (group B). Monetary allocations were made through 4 examples of the Tajfel matrices (Turner, 1983; Bourhis, Sachdev, & Gagnon, 1994). Each matrix contained seven different pairs of numbers. Participants were asked to choose one pair of numbers per matrix to represent the amount they would like to allocate as a reward to an anonymous member of the in-group (group A) and an anonymous member of the out-group (group B). Participants then completed the questionnaire, which measured how they felt about the task. Participants were subsequently presented with the matrices again, but asked to allocate the money as a fine, instead of a reward. They were then asked to complete the questionnaire again.

Dependent Measures

The total amount of money allocated to the in-group was collapsed across the Tajfel matrices to provide an overall measure of inter-group discrimination (negative allocations were reverse-coded). However, for the purposes of this pilot study, the primary dependent variables measured participants' emotions resulting from the two allocation tasks and their perceptions of the tasks. Items measured the extent to which allocating rewards was a pleasant act, "The allocation task was pleasant" and "It was a nice thing to do", and the perceived impact of allocations on recipients, "The allocation would have a positive impact on recipients" and "The allocation would have a negative

effect on recipients”. A single item was included to measure the perceived legitimacy of the allocation task: “I felt I was taking legitimate action whilst allocating the money”. Each item was answered on a 7-point scale, where 1=strongly disagree, 4=neither agree nor disagree, 7=strongly agree.

In order to measure the level of personal satisfaction and/or distress the tasks caused, adapted versions of the positive and negative affect scales (PANAS) were included (Watson, Clark, & Tellegen, 1988), in which participants rated the extent to which they experienced emotions that were positive (8 items, $\alpha=.81$ for reward allocations and $\alpha=.79$ for fine allocations) and negative (16 items, $\alpha=.93$ for reward allocations and $\alpha=.96$ for fine allocations) on a scale of 1=not at all to 5=extremely. Listed positive emotion words were: determined, excited, comfortable, enthusiastic, pleased, legitimate, happy and interested. Negative emotion words were: distressed, guilty, hostile, illegitimate, ashamed, jittery, upset, nervous, afraid, uncomfortable, tense, anxious, angry, worried, awkward and uneasy.

Results and Discussion

Repeated measures ANOVA showed that participants displayed more inter-group discrimination when allocating rewards ($M=176.00$, $SD=29.02$) than fines ($M=138.00$, $SD=37.34$), $F(1, 19)=8.16$, $p=.01$, $\eta^2=.30$, replicating the traditional positive-negative asymmetry effect. Participants also felt that allocating rewards was marginally more legitimate ($M=4.65$, $SD=1.63$) than allocating fines ($M=3.95$, $SD=1.67$) $F(1, 19)=3.86$, $p=.06$, $\eta^2=.17$.

Participants agreed significantly more with the item, “The allocation would have a positive impact on recipients” in the reward allocation condition ($M=5.45$, $SD=1.54$) compared to the fine allocation condition ($M=2.60$, $SD=1.82$) $F(1, 19)=21.06$, $p<.01$, $\eta^2=.53$. Conversely, participants agreed that, “The allocation would have a negative

effect on recipients” more when allocating fines ($M=4.50$, $SD=2.09$) than allocating rewards ($M=2.80$, $SD=1.40$) $F(1, 19)=8.06$, $p=.01$, $\eta^2=.30$. They also agreed that “It was a nice thing to do” more in the reward condition ($M=5.35$, $SD=1.60$) than the fine condition ($M=2.60$, $SD=1.93$) $F(1, 19)=14.68$, $p<.01$, $\eta^2=.44$. Agreement with the item, “The allocation task was pleasant” was also significantly higher when allocating rewards ($M=4.50$, $SD=1.32$) compared to allocating fines ($M=3.10$, $SD=1.59$) $F(1, 19)=9.00$, $p<.01$, $\eta^2=.32$.

Results for the experienced emotions also corroborated the assumptions. Participants felt significantly more positive about allocating rewards ($M=2.83$, $SD=.71$) than allocating fines ($M=2.08$, $SD=.66$) $F(1, 19)=21.37$, $p<.01$, $\eta^2=.53$. Conversely, participants felt significantly more negative about allocating fines ($M=1.88$, $SD=.74$) than allocating rewards ($M=1.47$, $SD=.51$) $F(1, 19)=9.79$, $p<.01$, $\eta^2=.34$.

In sum, results confirm that participants were more likely to discriminate on rewards than fines, find allocating rewards to be a more legitimate and pleasant act than allocating fines and believe that reward allocation would have a more positive impact on recipients than fine allocation. Conversely, participants thought allocating fines would have a more negative effect on recipients and felt more negative about allocating fines than rewards. Therefore, it can be concluded that the allocation of fines and rewards are not equivalent. Psychologically, allocation of positive resources is experienced as administering a *reward*, allocation of negative resources a *punishment*.

STUDY 2.2

Having established that allocating negative resources was considered to have a negative impact on recipients, the main study was designed to test four hypotheses, that: (a) Perceived obstruction would reduce or remove the illegitimacy of punishing the out-group in this way; (b) The PNAE would be replicated before interaction took place; (c) Intra-group interaction should have a transformative effect on discrimination,

and (d) Discrimination would be mediated by a hostile group norm. It was thus predicted that when groups felt unobstructed by the out-group, the PNAE would be replicated—discrimination only being displayed when positive outcomes were allocated. In groups that were obstructed, however, discrimination should be displayed on both positive and negative allocations. Furthermore, interaction would serve to formulate and focus in-group norms that guide future behaviour.

Method

Participants and Design

Participants were 102 volunteers from a sixth-form college. Participants were recruited as part of an educational visit to the institution. The mean age was 16.7 years old and 53 participants were male. All participants were unpaid and naive as to the purposes of the study. The design was a 2 (valence of monetary allocation: reward vs. fine) x 2 (group advancement: obstructed vs. not-obstructed) x 2 (phase: pre-consensus, consensus) mixed factorial design, with repeated measures on the phase factor. Participants were randomly allocated to conditions, and within that to small 3-person discussion groups ($N=34$).

Independent Variables

Participants were asked to decide upon a suitable monetary allocation for the out-group. This task was carried out twice – once individually (pre-consensus) and once after a group decision (consensus). Measures taken before and after group discussion constitute the repeated measures factor phase in the design. One between-subjects independent variable was the valence of the allocations. Participants were asked to allocate monetary rewards (positive) or to administer fines (negative “punishments”) to the in-group and out-group. The second between-subjects independent variable was whether the out-group obstructed the future advancement of the in-group, or not. In one condition, participants were given feedback that suggested that the out-group

deliberately obstructed the in-group's opportunities to carry out a subsequent "survival" task successfully. In the other condition, feedback suggested that the out-group did not obstruct the in-group.

Dependent Measures

The main dependent variable was discrimination. As before, monetary allocations were made through the Tajfel matrices³ (Turner, 1983; Bourhis et al., 1994). Measures used were identical to those of Amiot and Bourhis (2005). There were seven matrices, each containing seven different pairs of numbers. As before, for each matrix participants were asked to choose one pair of numbers, which represented the amount allocated (or fine awarded) to an anonymous member of the in-group and an anonymous member of the out-group. The total amount allocated (or fined) to the in-group was used as the main dependent measure. The total allocations were collapsed across matrices and standardised. Negative allocations (fines) were reverse-coded. The inter-group discrimination measure reflects the total monetary allocation to the in-group over the out-group (i.e., the net in-group gain), such that zero indicates that money was distributed equally between the in-group and out-group.

Two further measures were included as process variables. These scales consisted of statements to which participants indicated agreement on 7-point scales (1= "strongly disagree," 7="strongly agree"). A 9-item "hostile norm" scale was designed to measure the extent to which participants felt the group norm legitimised hostility, which had good reliability ($\alpha=.89$). This scale was included as a potential mediator of allocations. Example items in this scale were, "I feel it would be justified to take money away from the other group"; "I think it would be ok to favour my group"; "I think it would be

³ The full range of allocation strategies detectable in the matrices were measured, but were not considered to add anything of value in terms of the analyses reported here. The results are available from the author on request.

OK to punish the other group” and “Disadvantaging the other group is the right thing to do.” The hostile norm measure includes statements both about the legitimacy of favouring the in-group and punishing the out-group in order to be relevant to both reward and fine allocation. This was included at both the pre-consensus and consensus phases to measure what the in-group considered legitimate and appropriate inter-group behaviour before and after discussion.

Ellemers, Kortekaas and Ouwerkerk’s (1999) 3-component measure of identification was included at the consensus phase (10 items, $\alpha=.86$). Within the overall measure of identification, the 3 sub-scale components differentiated between social self-categorisation (3 items, $\alpha=.66$), “I identify with the other members of group A”; “I am like other members of group A” and “Group A is a reflection of who I am”; group commitment(3 items, $\alpha=.78$), “I would like to continue working with group A”; “I dislike being a member of group A”(reverse-coded) and “I would rather belong to group B” (reverse-coded); and group self-esteem (4 items, $\alpha=.79$), “I think group A has little to be proud of” (reverse-coded); “I feel good about group A”; “I have little respect for group A” (reverse-coded) and “I would rather not tell that I belong to group A” (reverse-coded). The measure of identification was included to counter alternative explanations of the results in terms of group discussion galvanising social identity.

Procedure

All participants were presented with a consent form, which stated that this was an experiment on group survival. The experiment consisted of two phases. During the pre-consensus phase, participants were informed that they would carry out a (theoretical) survival task, and in that context they divided money between in- and out-group. The second (consensus) phase consisted of a group discussion during which

groups were asked to reach consensus about the allocations and the other dependent variables.

Pre-consensus phase. This phase started with allocation to conditions and groups. Participants were seated at classroom desks and read consent forms. The experimenter presented herself as a psychologist conducting research into the psychology of group survival. All consenting participants were randomly and equally allocated to one of 2 groups (ostensibly group A or B). Participants were placed in an isolated room with their team, and from that moment on treatments were identical except for the manipulations of independent variables. All participants were informed that they were allocated to group A, and that group B were in an adjacent room. Participants were then seated at individual desks. They were asked to read feedback about the tasks they were going to carry out, and fill out some matrices. This was all done in silence to ensure that all questions were answered individually at this phase.

The feedback then informed participants that their group, A, was going to compete against group B in a theoretical “jungle-survival competition.” In order to beat group B, they were told that they must perform better than them at a survival task. Furthermore, in order to complete the survival task, it was essential that they had certain named items of equipment: a torch, matches, a tent, food, water, a knife and a first aid kit. Participants were not given any specific information about the nature of the task yet (although it was clear to them that this was an exercise rather than life or death situation).

Participants were then asked to indicate which items they would like in their group’s kit by ticking boxes that corresponded to those items. They were also told that group B were given the opportunity to divide the kits between groups A and B. After a brief wait, during which the experimenter ostensibly collected group B’s proposed division, participants in the obstructed condition were told that group B did not

allocate certain named essential items to group A, but allocated these items to themselves instead. This was done to give participants the impression that the out-group were an obstruction to their group's progress in the survival task. In the non-obstructed condition, the feedback was that group B allocated items equally between A and B, and so appeared fair.

Next, the distribution of money was justified by informing participants that such distributions took place within the context of the jungle survival task. Participants were told that in order to supplement the jungle survival kit, they had £500 to buy extra items. Critically, they were made aware that they would not be able to use this money to buy essential items (i.e., they could not use the money to undo the out-group's obstruction). Participants were instructed to divide an unspecified amount of this £500 between members of their group and the out-group, either by a reward system or fining system, using the Tajfel matrices. Thus, if money was allocated as a fine, participants would have X amount less to spend on extra items of kit. This design ensured that the operationalisation of the monetary allocation and the manipulation of obstruction were unconfounded. It was assumed that fining the out-group was judged to be a more negative behaviour than was rewarding the in-group. After making the allocations, participants were asked to complete the hostile norm scale.

Consensus phase. Following instructions from the experimenter, participants were placed into smaller discussion groups of three participants, within the superordinate "Group A" category. Participants were asked first to discuss their "jungle survival plan," with particular reference to how they would use their survival kit in order to be more successful than Group B. This was done to make the obstruction salient. They were then instructed by the experimenter to discuss the matrices in their groups until they reached consensus on the allocation of the valued outcomes. They then recorded

their allocation decisions individually, and completed the social identification and hostile norm scales independently. The content of the group interaction was not recorded, but the experimenter and her assistants remained in the room and unobtrusively monitored the discussion during the study to ensure groups kept to the intended topic. Participants were then debriefed.

Analytic Strategy

Condition effects

Considering that the responses to the measures at the consensus phase were shaped by the intra-group interaction, this data is interdependent. To address this issue, mixed design hierarchical linear modelling (HLM) analyses were performed in HLM for Windows 6.03 (Raudenbush & Bryk, 2002) on the data. Mixed design HLM analyses are appropriate for repeated measures data (Level 1) in which one level of analysis (individuals, level 2) is nested within another (groups, level 3). This procedure examined the behaviour of the level 2 outcome (discrimination and the hostile norm) at both the pre-consensus and consensus phase as a function of both level 1, level 2 and level 3 predictors. Dummy variables were created to represent the independent variables (i.e., a multilevel equivalent of the General Linear Model approach). The estimated models were:

$$\text{Level-1 Model: } Y = \Pi_0 + \Pi_1 (\text{phase}) + E$$

$$\text{Level-2 Model: } \beta_0 + R$$

$$\text{Level-3 Model: } \beta_0 = \gamma_{00} + \gamma_{01} * (\text{obstruction}) + \gamma_{02} * (\text{valence of monetary allocation}) + \gamma_{03} * (\text{2-way interaction}) + U_0$$

In these models, Y represents the dependent variable of interest (either discrimination or the hostile norm), E, R and U the errors at the intra-individual, individual and group levels respectively, and Π , β and γ are regression coefficients at the intra-individual, individual and group levels respectively.

Tests of Specific Hypotheses

Tests of specific hypotheses for the allocations within each condition were carried out by means of contrasts (Judd, McClelland, & Culhane, 1995; Rosenthal & Rosnow, 1985; Wilkinson et al., 1999), separately in the pre-consensus and consensus phases. Recommendations were followed for contrast analysis in the testing of regression models by Cohen, Cohen, West and Aiken (2003, pp. 332-341). Table 2.1 provides the coding of the contrast variables to aid interpretation. Four hypotheses were tested: (a) That there would be more discrimination when participants were obstructed relative to when they were not obstructed; (b) That allocating rewards would increase discrimination relative to allocating fines (the PNAE) (c) That interaction would transform the discriminatory behaviour exhibited, and (d) That these effects would be mediated by the group norm. Level 1 contains the individual level variables and the group-level contrast variables were entered as uncentered variables at level 2. The estimated models were:

$$\text{Level-1 Model: } Y = \beta_0 + R$$

$$\text{Level-2 Model: } \beta_0 = \gamma_{00} + \gamma_{01} * (\text{contrast1}) + \gamma_{02} * (\text{contrast2}) + \gamma_{03} * (\text{contrast3}) + U_0$$

In these models, Y represents the dependent variable, R and U the errors at the individual and group level respectively, and β and γ are regression coefficients at the individual and group levels respectively. In the results section, *t* and *p* values are reported as a direct test of the relation between the independent (via dummy and contrast variables) and dependent variables. Means and standard deviations are reported in Table 2.2.

Table 2.1

Coding of contrast variables for HLM, Study 2.2

		Contrast A: Obstructed vs. Not- obstructed conditions			Contrast B: Reward vs. Punishment Conditions		
		Contrast variables			Contrast variables		
Condition		C1	C2	C3	C1	C2	C3
Reward	Not	1/2	1/2	0	1/2	1/2	0
	Obstructed						
	Obstructed	-1/2	0	1/2	1/2	-1/2	0
Punishment	Not	1/2	-1/2	0	-1/2	0	1/2
	Obstructed						
	Obstructed	-1/2	0	-1/2	-1/2	0	-1/2

Results

*Discrimination**Condition Effects*

A three-way mixed HLM analysis (obstruction of in-group advancement x valence of monetary allocation x phase) examined discrimination scores. There was a main effect for the valence of monetary allocation $t(28)=-3.36, p<.01$, with significantly more discrimination on rewards than fines. There was also a phase main effect, with more discrimination prior to discussion ($M=.56, SD=.59$) than after discussion ($M=.45, SD=.55$) $t(194)=-5.20, p<.01$. There was a significant interaction between the valence of monetary allocation and phase: prior to discussion, there was more discrimination on rewards than fines. After discussion, there was much more discrimination on the fines (Table 2.2), $t(194)=3.41, p<.01$. However, these effects were qualified by a significant three-way interaction, $t(194)=-1.96, p=.05$. Examination of means suggests that, prior to interaction, all cells show relatively equal levels of discrimination, except for participants in the non-obstructed condition who allocated (negative) fines. After interaction however, the least discrimination is displayed by participants in the non-obstructed condition who allocated (positive) rewards (Figure 2.1).

Specific Hypotheses

Pre-consensus phase. At pre-consensus, there was a significant main effect for obstruction, $t(98)=-2.54, p=.01$, with more discrimination when participants were obstructed ($M=.67, SD=.56$) than not obstructed ($M=.43, SD=.60$). The main effect for valence of monetary allocation was also significant, $t(98)=2.25, p=.03$, with more discrimination overall on the rewards ($M=.66, SD=.62$) than fines ($M=.43, SD=.51$). This finding confirms that overall, we found evidence for the PNAE, as predicted by hypothesis (b).

The PNAE was then explored in the non-obstructed and obstructed conditions separately, using contrast tests. In the non-obstructed condition there was a PNAE, with significantly more discrimination when money was allocated as rewards ($M=.56, SD=.62$) than as fines ($M=.16, SD=.46$) $t(98)=2.22, p=.03$. In the obstructed condition however, the PNAE was eliminated. Discrimination here was equal in the reward ($M=.76, SD=.62$) and punishment ($M=.57, SD=.48$) conditions, $t(98)=1.04, p=.30$, supporting hypothesis (a).

Consensus phase. After group consensus had been reached, there was a significant main effect of obstruction on discrimination, $t(98)=4.48, p=.001$, with greater discrimination when participants were obstructed ($M=.62, SD=.58$) than when they were not obstructed ($M=.25, SD=.44$). There was also a significant main effect of valence of the monetary allocation on discrimination, $t(98)=2.26, p=.03$, with more discrimination when allocating fines ($M=.53, SD=.45$) than rewards ($M=.40, SD=.61$). This shows that consensual discussion resulted, overall, in a significant *reversal* of the PNAE, which provided evidence for hypothesis (c).

These effects were further explored in the non-obstructed and obstructed conditions separately, using the contrasts. In the non-obstructed condition the reverse PNAE was significant: there was more discrimination on fines ($M=.52, SD=.57$) than

on the rewards ($M=.12$, $SD=.29$), $t(98)=-2.26$, $p=.03$. In the obstructed condition there was no significant difference between reward and punishment conditions, $t(98)=1.10$, $p=.27$. Therefore, perceived obstruction eliminated the PNAE also after consensus, which was further support for hypothesis (b). Positive-negative asymmetries were found in the non-obstructed conditions *only*: means were in the direction of the traditionally reported PNAE at pre-consensus, but the effect *reversed* as a result of group discussion.

Table 2.2

Mean scores at pre-consensus and consensus, Study 2.2

		Rewards		Punishments	
		Not obstructed	Obstructed	Not obstructed	Obstructed
Discrimination, pre-consensus	<i>M</i>	.56	.76	.16	.57
	<i>SD</i>	.62	.62	.46	.48
Discrimination, consensus	<i>M</i>	.12	.70	.52	.53
	<i>SD</i>	.29	.72	.57	.39
Hostile norm, pre-consensus	<i>M</i>	3.74	3.86	3.54	4.46
	<i>SD</i>	1.30	1.20	1.21	0.73
Hostile norm, consensus	<i>M</i>	3.37	3.90	4.37	4.43
	<i>SD</i>	.92	1.19	1.63	.58
Overall social identification	<i>M</i>	4.99	4.65	5.01	4.71
	<i>SD</i>	1.02	1.19	.80	.78
Social self-categorisation	<i>M</i>	4.29	4.10	4.22	3.83
	<i>SD</i>	1.35	1.28	.53	1.15
Group commitment	<i>M</i>	5.54	4.77	5.33	4.94
	<i>SD</i>	1.06	1.50	1.10	1.05
Group self-esteem	<i>M</i>	5.10	4.96	5.35	5.19
	<i>SD</i>	1.27	1.22	1.14	1.03

Notes. The total sum of money allocated to the in-group at each phase (where zero is equivalent to no discrimination, or equal distribution of money between groups), was standardised. For the standardised scales, a higher number indicates a higher propensity on that measure.

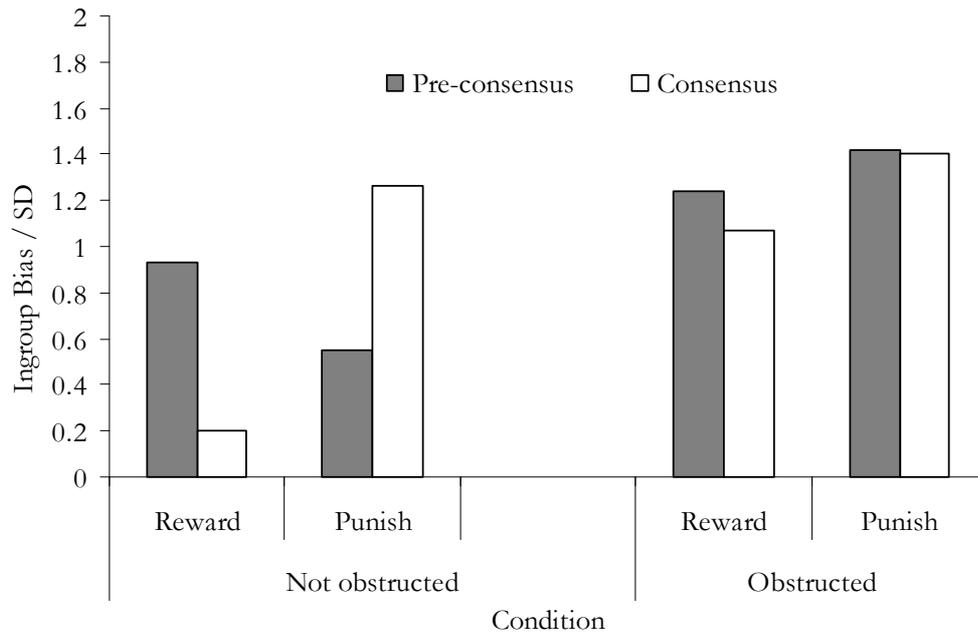


Figure 2.1. Mean discrimination scores, Study 2.2

Hostile Norm

Condition Effects

As before, a three-way HLM (obstruction of in-group advancement x valence of monetary allocation x phase) examined effects on the hostile norm. There was no main effect for the obstruction of in-group advancement, $t(26)=-1.08$, $p=.29$, although the means suggest that the norm legitimised somewhat more hostility when participants felt obstructed by the out-group than when they did not feel obstructed (Table 2.2). There was a main effect for the valence of monetary allocation, $t(26)=-3.32$, $p<.01$, with a more hostile norm when allocating fines than rewards (Table 2.2). Furthermore, there was an interaction between the valence of monetary allocation and phase, $t(178)=2.88$, $p<.01$. There was a more hostile norm on the fines both prior to and after discussion, but the norm became more hostile in the punishment condition overall, and less hostile in the reward condition, which may partly explain the reversal of discrimination behaviour. However, these results were qualified by a significant three-way interaction,

$t(178)=-2.83, p<.01$ (Figure 2.2). Examination of means in the obstructed condition suggests that there was a more hostile norm in the punishment condition compared to the reward condition, and interaction had little effect on this norm. In contrast, in the not obstructed condition, the norm was similar for rewards and fines *prior* to interaction. After interaction however, the pattern changes, and the norm is much more hostile in the punishment conditions than in the reward conditions.

Specific Hypotheses

Pre-consensus phase. A 2 (obstruction of in-group advancement: obstructed vs. not obstructed) x 2 (valence of monetary allocation: reward vs. fine) HLM analysis was conducted on the hostile norm scores in the pre-consensus phase. The main effect for valence of monetary allocation was not significant, $t(98)=0.12, p=.91$, neither was the obstruction main effect, $t(98)=0.44, p=.67$, nor the 2-way interaction, $t(98)=1.04, p=.30$. However, contrast tests performed within the obstructed condition showed that there was a significantly more hostile norm when money was allocated as a fine ($M=4.46, SD=.73$) than when it was allocated as a reward ($M=3.86, SD=1.20$) $t(90)=-2.05, p=.05$. There was no valence of monetary allocation main effect in the non-obstructed condition $t(90)=.49, p=.62$.

Consensus phase. Hostile norm scores from the consensus phase were entered as the dependent variable at level one and the dummy variables were entered at level two. There was no main effect of obstruction on the hostile norm, $t(28)=1.31, p=.20$. There was however, a main effect of valence of the monetary allocation on the hostile norm, $t(28)=2.06, p=.05$, with more hostility when allocating fines ($M=4.41, SD=1.01$) than rewards ($M=3.62, SD=1.08$).

The norm was then explored in the non-obstructed and obstructed conditions separately, using the contrast variables. There was more hostility in the unobstructed conditions when money was allocated as fines ($M=4.37, SD=1.63$) than as rewards

($M=3.37$, $SD=.92$) $t(90)=-2.54$, $p=.01$. There was no effect of valence of the monetary allocation on the hostile norm in the obstructed conditions $t(90)=-1.50$, $p=.14$.

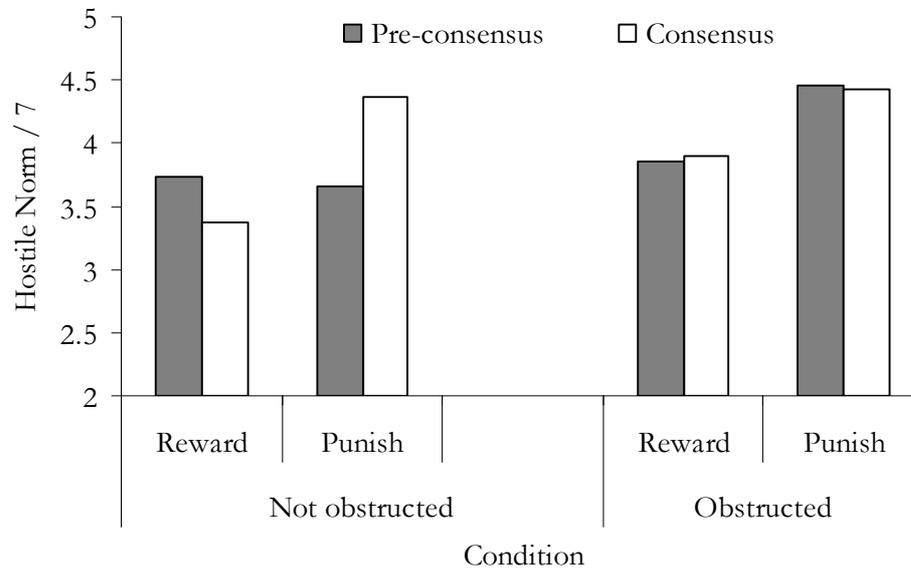


Figure 2.2. Mean hostile norm scores, Study 2.2

Mediation

A three-way mixed HLM analysis (obstruction of in-group advancement x monetary allocation x phase) was conducted to examine hypothesis (d): that changes in the group norm could account for the discrimination effects. This was done by adding the group norm as a level-1 predictor to the model examining condition effects (i.e., as if it were a covariate). In this mediation analysis, the norm was indeed a very significant predictor of the degree to which discrimination occurred, $t(191)=6.60$, $p<.01$. Looking at the condition effects, the main effect for the valence of monetary allocation remained $t(28)=-2.92$, $p=.01$; as did the phase main effect $t(194)=-4.24$, $p<.01$ and the interaction between the valence of monetary allocation and phase $t(191)=2.72$, $p=.01$. However, the three-way interaction which previously qualified these effects became non-significant, $t(191)=-1.47$, $p=.14$. The fact that the Sobel test also showed that the reduction in effects was significant, $Z=-2.62$, $p=.01$, demonstrates that the group norm fully mediated this effect, providing evidence for hypothesis (d).

Social Identification

Overall social identification scores from the consensus phase were entered as the dependent variable at level one and the dummy variables were entered at level two. There was neither a main effect of obstruction on social identification, $t(28)=-0.81$, $p=.43$, a main effect for valence of monetary allocation, $t(28)=-0.19$, $p=.86$, nor a 2-way interaction $t(28)=0.24$, $p=.81$. Social identification also did not significantly predict discrimination scores $t(100)=.12$, $p=.91$ when added as a predictor at level 1, with discrimination as the dependent variable. This suggests that changes in identification cannot explain the condition effects reported above for discrimination and the hostile norm.

Each of the identification sub-scales were then analysed for condition effects. There was no main effect of obstruction on social self-categorisation, $t(28)=-0.39$, $p=.70$, nor was there a main effect for valence of monetary allocation, $t(28)=-0.20$, $p=.85$, nor a 2-way interaction $t(28)=0.10$, $p=.92$. For group self-esteem there was also no main effect of obstruction, $t(28)=-0.29$, $p=.77$, or valence of monetary allocation, $t(28)=0.37$, $p=.71$, nor a 2-way interaction $t(28)=0.16$, $p=.87$. Please refer to Table 2.2 for means.

For group commitment, there was a main effect of obstruction, $t(28)=-2.12$, $p=.04$, with greater commitment when participants felt unobstructed, ($M=5.47$, $SD=1.06$) than when they felt obstructed ($M=4.86$, $SD=1.29$). There was no valence of monetary allocation main effect, $t(28)=-0.93$, $p=.36$, however, nor a 2-way interaction $t(28)=0.83$, $p=.42$. Neither self-categorisation $t(100)=-0.73$, $p=.47$, group self-esteem $t(100)=.36$, $p=.72$, nor group commitment $t(100)=0.17$, $p=.87$ significantly predicted discrimination.

Proportion of Respondents Behaving Fairly

The finding that group discussion produced a complete *reversal* of the PNAE in the non-obstructed condition was unexpected. In order to explore why consensualisation could have produced this particular effect, we examined the distribution of pre-discussion responses. It has been found in the minimal group paradigm that a substantial proportion of respondents choose to distribute resources equally (e.g., Bornstein, Crum, Wittenbraker, Harring, Insko, & Thibaut, 1983). A typical finding in the small group decision making literature is that if a majority of people support a particular viewpoint in a group discussion, this majority viewpoint tends to be adopted by the whole group (Hastie & Kameda, 2005). We reasoned that the “majority wins” rule could have something to do with this reversal. We calculated how many participants had allocated money equally between the in-group and out-group. An analysis of the number of participants distributing equally versus those displaying discrimination showed, somewhat unexpectedly, that a larger proportion of participants distributed money equally in the positive reward condition ($N=22$, or 37%) than in the punishment condition ($N=6$, or 14%; $\chi^2=45.87$, $p<.01$, Table 2.3).

Table 2.3

N participants choosing to distribute resources equally between groups

	Rewards	Punishments
Equal distribution	22	6
Unequal distribution	37	37

This, in conjunction with the finding that group norms mediated the discrimination effects, suggests that a different intra-group dynamic may have occurred in this condition, with majority opinion leading to the inference of a norm for equality, driving the group averages down.

General Discussion

Previous research has shown that whilst in-group members are happy to display in-group favouritism whilst allocating rewards between groups, they are generally reluctant to punish out-group members by allocating them higher fines than they allocate to the in-group. The pilot study conducted here suggests that this may be because negative allocations are considered to have a negative impact on recipients and this is perceived to be illegitimate behaviour. The purpose of the main study was to examine the impact of intra-group discussion and the obstruction of in-group advancement on this phenomenon. First, it was hypothesised (a) that perceived obstruction would reduce or remove the illegitimacy of punishing the out-group (and therefore, the PNAE). Results supported this hypothesis. There was more discrimination when participants felt obstructed than when they did not feel obstructed. This effect was in evidence before consensus and after group discussion. Moreover, at the pre-consensus phase the traditional PNAE was replicated in the non-obstructed conditions, with more discrimination displayed when participants allocated rewards than punishments. This provides evidence to support our second hypothesis (b), that the PNAE would be replicated prior to interaction. Yet in the obstructed conditions, the PNAE was eliminated. This finding is consistent with the more general finding that socio-structural conditions can directly impact upon inter-group behaviour (Ellemers, 1993, Sherif, 1967; Wright, Taylor, & Moghaddam, 1990).

Indeed, this is not the first time that research demonstrates that there are conditions under which the PNAE is attenuated (e.g., Sassenberg et al., 2003). However, the present results are nonetheless of interest. In particular, unlike previous research the present findings suggest that the PNAE may be particular to experimentation on “minimal groups” which, by definition, do not include socio-structural conditions. The conceptualisation of the PNAE within a richer than minimal

context, i.e., against a backdrop of a competitive socio-structural context; is a small step towards increasing the ecological validity of this line of research.

The third prediction (c) was that intra-group interaction should have a transformative effect on discrimination. There was no evidence for this prediction in the condition where participants were obstructed by the out-group: Here, the in-group displayed high levels of discrimination across the board. But in the condition where there was no such obstruction, intra-group interaction did indeed *reverse* the traditional PNAE. Thus, after group discussion and reaching consensus, there was less discrimination when rewards were allocated, and significantly more discrimination when fines (punishments) were being handed out.

Finally, it was predicted (d) that the reason for this transformative effect of intra-group interaction would be related to the formation of in-group norms that justified hostile treatment of the out-group. Thus, the group norm should mediate the effect of intra-group interaction on discrimination. Indeed, the pattern of effects for the group norm of hostility largely mirrored those of discrimination. In the *unobstructed* condition, there was no difference in the norm for allocating rewards and for punishments prior to discussion, (although the pattern appeared to echo that of the discrimination results). After interaction, the pattern reverses in the same way as the discrimination scores do: the norm is more hostile in the *punishment* condition than in the reward condition. There was one difference between the discrimination and norm results: In the obstructed conditions, the norm was more hostile in the punishment condition than in reward condition. And yet, discrimination was equally strong in both conditions. This deviation between norms and discrimination is interesting in light of the mediation analysis.

In line with the prediction, the changes in discrimination exhibited in the non-obstructed conditions were fully mediated by the in-group norm. Indeed, in the

consensus phase, when participants had had the opportunity to interact, the norm becomes more hostile on punishments than rewards, following the same pattern as discrimination scores. This suggests that the norm explains behaviour in the non-obstructed conditions. But evidence for mediation was less obvious in the conditions where the group was obstructed (indeed the obstruction and obstruction by phase interaction effects did not show evidence of full mediation). A sense of obstruction appears to have (at least partially) overridden normative influence on behaviour. Retrospectively, this is perhaps not surprising: in these conditions, the obstruction itself may have carried sufficient weight to render the process of consensualisation and norm formation of secondary importance. Here, the nature of inter-group relations, in themselves, would appear to have been sufficient to legitimate punishments. Nevertheless, the in-group norm did fully mediate the remaining condition effects.

These findings confirm the transforming effect that group interaction can have on perceptions of legitimacy of in- and out-group actions (see also Stott & Drury, 2004). It appears that consensualisation can work both for and against the legitimisation of hostile norms, contingent upon the generic norm in place at the beginning of the interaction and the nature of the social context. This finding is consistent with Mummendey and Otten's (1998) suggestion that norms are an important factor in the occurrence of the PNAE. But the nature of these norms is in no sense fixed and generic: intra-group interaction can override and thus *qualitatively transform* any generic norms for equality and against administering harm to an out-group. The meaning of what is legitimate, or *normative*, inter-group behaviour can easily change as a function of intra-group debate (see also Amiot & Bourhis, 2003, 2005). In this way, consensualisation processes informed the in-group's identity by shaping its normative content.

Whilst the results suggested that neither group discussion nor valence of allocation appeared to change in-group identification *per se* (there being no change in identification across conditions), it is possible that with the normative shift, the meaning of the in-group (as well as the out-group) changed in the non-obstructed conditions. In other words, the meaning of being an in-group member may have changed according to the valence of the allocation and before and after discussion because of the shift in the in-group norm. This is consistent with previous research on identity content and inter-group relations (Haslam, Turner, Oakes, Reynolds et al., 1998; Haslam, et al., 2002; Reicher, 1996; Stott & Drury, 2000; 2004).

Although the complete reversal of the PNAE (as opposed to its mere elimination) as a result of intra-group interaction was somewhat unexpected, this does confirm the assertion that intra-group interactions are not neutral (or merely amplificatory, cf. Allport, 1924; Moscovici & Zavalloni, 1969) with respect to inter-group behaviour. Results confirm that intra-group dynamics have the capacity to transform inter-group behaviour. This reversal may be explained by the resource distribution choices of participants before interaction. Paradoxically, while the level of *discrimination* was on average higher when the monetary allocation was framed as a reward, there were also more individual participants who chose to divide resources *equally* in the same condition. The traditional normative account of the PNAE (i.e., in terms of general norms against negative discrimination; Mummendey & Otten, 1998) would be more consistent with the finding that more participants make equal distributions when allocating punishments. A further inspection of the distribution of allocation responses showed that when rewards were allocated there was the bimodal distribution that is typically found in studies of discrimination, with a large group of participants choosing equality, and another (somewhat smaller) group displaying a

substantial amount of discrimination in favour of the in-group. In the case of allocating punishments, the range of allocations was greater.

These differences in the distributions of allocations potentially explain the reversal of the PNAE in the non-obstructed groups. To the extent that groups followed the commonly applied “majority rule” in their decision making (Hastie & Kameda, 2005), groups in the positive reward conditions should have shifted more toward equality norms. After all, here the majority unequivocally chose to allocate equally. In the punishment condition, there was no majority for any course of action, and hence the majority rule could not be applied. In a context where fining was the only form of inter-group differentiation, the groups converged on the consensus that hostility would be legitimate. This is a well-known finding: group norm formation tends to accentuate inter-group differences (see also Postmes et al., 2005). Hence intra-group interaction led to the endorsement of the hostile norm and a shift toward “punishment”. The data supports this post-hoc analysis.

Conclusion. Results suggest that there are two qualifications to the PNAE. First, when in-group advancement is obstructed, no PNAE was found: groups felt sufficiently justified to punish the out-group. Second, when in-group advancement was not obstructed, the PNAE reversed after group discussion, such that more hostility occurred when participants administered punishments than when they awarded prizes. Because this reversal was mediated by processes of consensualisation and norm formation, the proposal that intra-group interactive processes can qualitatively transform the nature of inter-group behaviour was supported.

In sum, the PNAE appears to be restricted to particular conditions of making allocations to in- and out-groups. The PNAE would appear to be contingent upon both the intra-group and inter-group dynamic, which in turn affect each other. With the inclusion of consensualisation processes within a study on inter-group behaviour,

we hoped to make a small step towards redressing the lack of ecological resonance between research on in-group bias and inter-group discrimination and the prejudices that are prevalent in contemporary society. Moreover, the perception of inter-group relations and the (lack of) correspondence of group trajectories (i.e., the anticipation of future conflict or the experience of current obstruction) are important factors in the occurrence of discrimination. Perceived inter-group relations and intra-group dynamics are both important elements of how insidious cycles of wrongdoing come into being and why they continue.

CHAPTER 3

The Power of Talk: Legitimising Hostile Social Action

“Let me say it forcefully: From our perspective—that of modernity—a lot of what Islam and her culture bring is either objectionable or downright backward. [...] We rule our part of the world and all those who are foreign and above all alien, must adjust [...] to what we see as our core norms and values. [...] Incidentally, not with fists and bombs, but *by using our words as weapons.*”
(Fortuyn, 2002, pp. 154-155, authors’ translation, italics added)

The influence of intra-group dynamics on the individual has long been acknowledged (e.g., Lewin, 1953). However, the study of how these intra-group processes affect relations between groups is a more recent concern. There are many examples that speak to the power of intra-group talk. As reflected in the quote above, the Dutch politician Pim Fortuyn set the agenda for a national debate about immigration. The watershed of anti-immigrant vitriol that was unleashed shocked many, especially internationally. Within a relatively short time, it became acceptable mainstream behaviour to vent prejudiced views of immigrants, and their second generation descendants. This was remarkable, given the pervasive belief that Dutch national identity was defined by tolerance. Within the Netherlands, the prevailing atmosphere was one of liberation: It was good that these issues could finally be discussed openly, and Fortuyn (assassinated in 2002) was credited with lifting the taboos.

Paradoxically, public opinion research shows that this overt discussion of anti-immigrant attitudes was neither preceded, nor accompanied, nor followed by a sharp rise in anti-immigrant attitudes: Xenophobia among the Dutch was as high as in other

EU countries before the national debate changed, and stayed more or less at the same level until 2007 (e.g., Dagevos & Gijsberts, 2007). At the same time, the changing national debate did encourage political and behavioural change (European Commission against Racism and Intolerance, 2008). Verbal and (in some cases) physical aggression against “aliens” increased and a “tough” line on immigration was widely advocated. Dutch immigration policy is now one of the strictest in the EU.

Although it seems plausible that intra-group debate may have contributed to this increase in inter-group hostility, there is little research on the issue. Yet it is clear that the in-group plays a vital role in shaping inter-group behaviour, functioning as a social resource (Correll & Park, 2005) by validating and supporting the psychological reality of individual in-group members. Thus, including intra-group interaction in the study of inter-group relations adds a key ingredient: the power of talk.

The Power of Talk

During the Second World War, Lewin (1953) famously attempted to persuade women to feed offal to their families. His study compared the persuasive effects of an informative lecture versus group discussion on women’s decision to feed. Group discussion was far more persuasive. Only 3% of those who heard a lecture served offal, whereas after group discussion, 32% served it (cf. Coch & French, 1948, for a similar but more controlled study of worker productivity). These studies demonstrate that group discussion can have profound behavioural effects. Indeed, later research suggested that group discussion can lead individuals to accept higher levels of risk (Fraser, Gouge, & Billig, 1971; Stoner, 1968), and that it can polarise their opinions and attitudes (Lamm & Myers, 1978; Moscovici & Zavalloni, 1969; Turner, 1991).

More recent research has begun to examine the consequences of intra-group discussion for stereotype consensus (e.g., Thompson, Judd, & Park, 2000). This

research not only shows that intra-group communication has a significant impact on stereotype content (e.g., Kwok, Wright, & Kashima, 2007), but also that consensus is enhanced by social identity salience (Haslam, 1997; Haslam, Oakes, Reynolds, & Turner, 1998; Haslam, Turner, Oakes, Reynolds, & Doosje, 2002). It can be concluded from this literature that intra-group discussion can have a significant influence on social stereotypes. The present research integrates these different literatures, and demonstrates that intra-group discussions have the power to change inter-group behaviour.

Normative and Informational Explanations

In line with Deutsch and Gerard's (1955) classic work on social influence, two different classes of explanations for the effects of group discussion have been suggested: informational and normative. "Informational influence" may be defined as the pressure to accept information obtained from another as evidence about reality. To the extent that this information is processed systematically (Eagly & Chaiken, 1993) or elaborated cognitively (Petty & Cacioppo, 1986), attitudes should change. A key characteristic of this explanation for social influence within groups is that it does not depend on group processes *per se*, merely on the processing of information.

There is considerable evidence for the operation of these informational processes in group discussion. For example, Vinokur and Burnstein (1974) demonstrated that polarisation occurs when group members are exposed to novel and persuasive arguments during group discussion. However, information processing is not necessarily the only or even primary factor in group polarisation (De Dreu, Nijstad, & Van Knippenberg, 2008). There are numerous studies (a.o. Lewin's, 1953, seminal study of persuasion) which demonstrate that interaction can produce strong social effects even when the information provided to participants is held constant across

experimental conditions (Turner, 1991). In other words, group discussion produces psychological change over and above informational influence.

One such change that occurs is the development or inference of group norms that accompanies social interactions within groups. Groups are often forums for debate about the appropriate interpretations of events, or the valuation of information. Even if such values are not explicitly addressed, group members readily detect social conventions non-consciously, and adapt their behaviours to them (Postmes, Spears, Sakhel, & De Groot, 2001). In this vein, emergent norm theory (Turner & Killian, 1972) posited that norms develop organically through group interaction, because of a drive to make sense of an event. These normative inferences that readily accompany the exchange of information can have a significant influence on group behaviour that is independent of actual attitude change (Terry & Hogg, 1996).

The term “normative influence” is often used to describe the social influence that is exerted by social expectations or group standards. This may be the result of direct pressure exerted by the group (i.e., coercion). More commonly, social influence occurs because in-group members can function as a source of validation (or “confirmation”) for individuals’ beliefs and social perceptions (Festinger, 1954). From this perspective, intra-group discussion does not function merely as a source of information: It presents an opportunity to assess whether one’s individual beliefs are supported. To the extent that there is consensus, this process of validation enables a group to establish a shared sense of social reality (Festinger, 1950; Hardin & Higgins, 1996; Turner, 1991).

The process of social validation is particularly relevant for inter-group relations. After all, our perceptions of out-groups are dependent on social consensus and have limited basis in physical reality (cf. Searle, 1995). Thus, we depend on other in-group

members for the validation of our perceptions of out-groups. Importantly, these socially shared perceptions have consequences for inter-group relations: To the extent that stereotypes are socially validated, they also become normative (see also Sherif, 1935). Stereotypes thus carry with them implicit ideologies for inter-group behaviour (e.g., Jost & Banaji, 1994). As a result, stereotypes and their associated prejudices may legitimise particular treatments of the out-group. Furthermore, a norm serves as an anchor for what is perceived as valid, as well as socially supported and desirable, and all this bolsters group members' confidence in their actions towards an out-group (Van Zomeren, Postmes, & Spears, 2008).

In sum, two distinct forms of influence may be exerted within group interaction. According to informational influence principles, social information exchange produces individual attitude change, which can lead to behavioural change. According to normative influence explanations, information exchange may validate individual beliefs, which then become firm foundations for action. In practice, these two processes are likely to be related to each other (e.g., Wetherell, 1987). However, a key difference between the two is that social validation of existing social stereotypes may occur in the absence of attitude change.

Stereotype Consensualisation and the Normalisation of Implicit Ideologies

The literature above suggests that particularly in the domain of inter-group relations, individuals engage in reality testing through intra-group interaction. Indeed, links between consensus-based validation, confidence and judgment intensification have been documented in prior research (Baron et al., 1996; Luus & Wells, 1994). The present research investigates whether comparable processes affect out-group stereotyping, and examines the behavioural consequences for hostile action.

Although the inter-group relations literature has largely ignored the roles of validation and intra-group interaction, it nevertheless suggests several reasons why stereotype consensus would have consequences for hostility. We know that stereotypes are powerful, functional, dynamic and value-laden constructs, reflecting ideologies and a particular world view. As such, they can be used to achieve certain social outcomes (e.g., Haslam, 1997; Jost & Banaji, 1994; Reicher, Hopkins, & Condor, 1997; Reynolds, Oakes et al., 2000). It has been argued that when there is a perception of consensus on social stereotypes, they become subjectively valid beliefs (Haslam, Turner, Oakes, McGarty et al., 1998) and this can influence attitudes and behaviour (Sechrist & Stangor, 2001). It has also been suggested that inter-group behaviour is contingent upon the validity of inter-group beliefs (Reicher et al., 1997). Furthermore, stereotypes may include implicit prescriptions for appropriate collective behaviour toward the stereotyped group (Haslam, 1997; Reicher et al., 1997). It follows that stereotype content can propel social protest (Stott & Drury, 2004) and that it can challenge (or reinforce) social hierarchies (Reynolds, Oakes et al., 2000).

In line with these ideas, the current research defines consensualisation as the process by which, through discussion, group members negotiate and create shared understandings (for example, of the out-group stereotype). We assume that consensualisation validates social stereotypes held by individuals, and thereby defines what constitutes legitimate, normative social behaviour towards the out-group. In other words, consensualisation is an intra-group process which anchors individual beliefs to those of the in-group, thereby making those beliefs subjectively valid and normative. As a consequence, individuals may be bolstered to act upon stereotypical beliefs.

There is some indirect support for this theory. Stott and Drury (2004) found that intra-group interaction increased stereotype consensus, and argued that this increased the likelihood of collective protest against unfair treatment by the experimenter. In Chapter 2, we showed that intra-group interaction can lead to marked changes in inter-group behaviour, and that interaction fosters the development of norms which may legitimise acts of inter-group hostility. However, this prior research was somewhat limited by the artificial context it created: consensualisation may have added a sense of reality to the experimenters' feedback rather than the supposed inter-group dynamic. Moreover, neither provided evidence for the entire causal chain investigated in the present research.

Present Research Overview

In order to investigate the process of intra-group consensualisation and its effect on hostile social action, a series of studies were conducted within a real inter-group dynamic: The British people and immigration. This topic was at the centre of an ongoing debate in Britain. Three studies examined consequences of intra-group debate among 16-18 year old school pupils on hostile (anti-immigrant) intentions and actions. The use of minors as participants raised ethical concerns (which were taken into account throughout), but also adds to the ecological validity and importance of the results. The overall aim of the research was twofold: to demonstrate that intra-group talk has the power to produce inter-group hostility, and to examine the process by which this occurs.

STUDY 3.1

Using white British citizens as the target in-group, the first study simply compared the products of an intra-group discussion about the out-group (immigrants) with a no discussion control condition in which individuals reflected about the out-

group. Although group discussion has been shown to encourage shifts in attitude of either valence, because of the prevalent negative bias against immigration at the time, it was hypothesised that groups would consensualise on a hostile stereotype of the out-group. As a consequence they should advocate more hostile action towards the out-group after group discussion than when they reflected on the topic alone. We also hypothesised that a norm that supported inter-group hostility would be positively related to hostile action, and would help explain the difference between the two experimental conditions.

Method

Participants and Design

Participants were 39 white British⁴ sixth form students (Mean age=17.0, 56% female) from a state comprehensive school in the UK, recruited during an educational visit to their establishment. Participants were randomly allocated to $n=13$ groups of three. All were unpaid volunteers and naïve about the purpose of the study. Parental/guardian (opt-out) consent was requested in advance. The study had a 2-condition (Type of reflection: group discussion vs. individual reflection) between-subjects design.

Procedure

For 5 minutes, participants either engaged in a group discussion or individually reflected on the out-group (immigrants). The instructions were to “Formulate an impression of immigrants. What do you think immigrants are like as a group of people? What do they do? Who are they? Where do they come from? What is their relation to British people?” In order to keep participants focused on the same objective, they were

⁴ For each study, participants were asked to state whether or not they were British, although the length of their British citizenship was not recorded.

asked to write down the key points that arose and then asked to complete a questionnaire individually and in silence. The questionnaire consisted of standardised scales. After completing the measures, participants were debriefed in person by the experimenter and provided with a full written debriefing in order to neutralise potentially negative effects of the intervention. The debriefing included references to information about immigration issues and support help-lines.

*Dependent Measures*⁵

The written description of the out-group was content analysed. The main dependent measure was a 3-item social policy support scale ($\alpha=.67$), which asked, “How much social security money should immigrants get?” “How much free healthcare should be provided for immigrants?” and “How much should immigrants get in benefits?” Items were answered on a seven point Likert-type scale (1=Much more than British people, 4=Equal, 7=Much less than British people). This scale was designed to measure advocacy of hostile social policy towards the out-group.

All other scales consisted of statements with which participants indicated agreement on a 7-point Likert-type scale (1=“strongly disagree,” 7=“strongly agree”). A 4-item scale measured perceptions of the in-group norm ($\alpha=.72$): “We believe that keeping Britain British is the right thing to do”; “For students like us it is normal to favour our own kind”; “Students like us believe that it's right to treat immigrants better than British people” (reverse coded); and “The average sixth form student believes that

⁵ All three studies included measures of efficacy and anger. Both variables have been related to inter-group action in past research (e.g., van Zomeren, Postmes, & Spears, 2008). There were no significant condition effects or significant mediation for either variable, with one exception: In Study 3.1, anger mediated the condition effect together with group norm. As this effect could not be replicated in subsequent research, we assume that it was not robust and may have been due to Type II error.

immigration enriches British society” (reverse coded). A factor analysis on the norm and social policy items using principle axis factoring and oblique rotation extracted the two factors respectively. This suggests that the hypothesised outcome (hostile policy support) and process (norm) variables measured distinct constructs. The social policy support factor explained 38% of the variance, $\lambda=2.71$; the norm explained a further 17% variance, $\lambda=1.19$; and the inter-factor correlation was low ($r=.11$).

Attitudes towards immigrants were measured through an adaptation of a generalised prejudice scale as used in previous research into attitudes towards immigrants (Dambrun & Guimond, 2001; Guimond and Dambrun, 2002; Postmes & Smith, in press; 10-items, $\alpha=.83$). Items included, “The entry of foreign families into Britain should be more stringent” and “Immigrants should not be given responsibility or positions of authority over the British”.

A 7-item adaptation of Ellemers, Kortekaas and Ouwerkerk’s (1999) measure of identification was included ($\alpha=.51$), with sample items “I identify with the other students” and “I feel good about these students”. Identification was included to examine whether group discussion *per se* may galvanise social identity. Inter-scale correlations (Table 3.1) suggest that the discriminant validity of the scales was acceptable.

Analytic Strategy

A content analysis was performed on the written descriptions of the out-group provided by participants. The primary coder followed the iterative inductive and deductive coding procedures recommended by Miles and Huberman (1994). In accordance with these procedures the primary coder was not blind to the hypotheses. A second coder, who was blind, independently coded the data. Inter-rater agreement was 95%, and inter-rater reliability was excellent: Cohen’s (1960) kappa ranged from

Table 3.1

Mean individual-level scores and scale intercorrelations, Study 3.1

		Group	Individual	1	2	3	4
		Discussion	Reflection				
1. Social action	<i>M</i>	4.97 _a	4.47 _b	(.67)	.45*	.67**	.00
	<i>SD</i>	0.77	0.58				
2. Hostile Norm	<i>M</i>	4.35 _a	3.00 _b	.51**	(.72)	-.05	.10
	<i>SD</i>	1.68	1.77				
3. Prejudice	<i>M</i>	4.34 _a	3.87 _a	.64**	.54**	(.83)	.07
	<i>SD</i>	0.92	1.15				
4. Identification	<i>M</i>	4.47 _a	4.23 _a	.09	.10	-.03	(.51)
	<i>SD</i>	0.85	1.41				

Notes. A higher number indicates a higher propensity on each measure. Means in the same row with different subscripts differ significantly at $p < .05$. Cronbach's alphas are reported in parentheses. ** $p < .01$. Full scale intercorrelations are reported below the diagonal, and partial correlations (controlling for the effects of the manipulation) are reported above the diagonal.

$\kappa=.71$ to $\kappa=1.00$ across codes. The content of the stereotypes was assessed and compared across conditions.

Analyses were conducted with hierarchical linear modelling in HLM 6.03 (Raudenbush & Bryk, 2002) on the quantitative data, as is appropriate when individuals are nested within groups. We examined the level 1 outcomes (e.g., hostility) as a function of a level 2 predictor (as well as level 1 mediators). A-priori predictions were tested by means of contrasts (e.g., Wilkinson et al., 1999), following the procedure outlined by e.g., Cohen, Cohen, West and Aiken (2003, pp. 332-341). Two hypotheses were tested: (a) group discussion increases hostile social policy support relative to individual reflection, and (b) the group norm mediates this condition effect.

Results

Discussion Content

On average, participants listed 4.36 ideas (Figure 3.1; see Appendix 3.1 for a full explanation of the codes and example quotes). There were considerable differences in out-group stereotyping across conditions, both in descriptive content and affective tone. The stereotype was noticeably more hostile in the group discussion condition. In the discussion condition, immigrants were described as poor (62% vs. 30% in the individual reflection condition), and as taking advantage of the welfare system (87.5% vs. 30%), for example, “Some immigrants are merely here to deal drugs and rake money off the state”.

In the individual reflection condition, participants were more resistant to stereotyping immigrants: “They come from all walks of life: every country, every race and every ethnic background”. Participants in this condition expressed more sympathy towards immigrants (60% vs. 25%) and emphasised that many are refugees (65% compared to 37.5%), who simply seek a better life in the UK (60% compared to

37.5%), e.g.: “They must be scared to come over from a foreign country, to be so scared to have to leave home... just ordinary people like you and me who are desperate”.

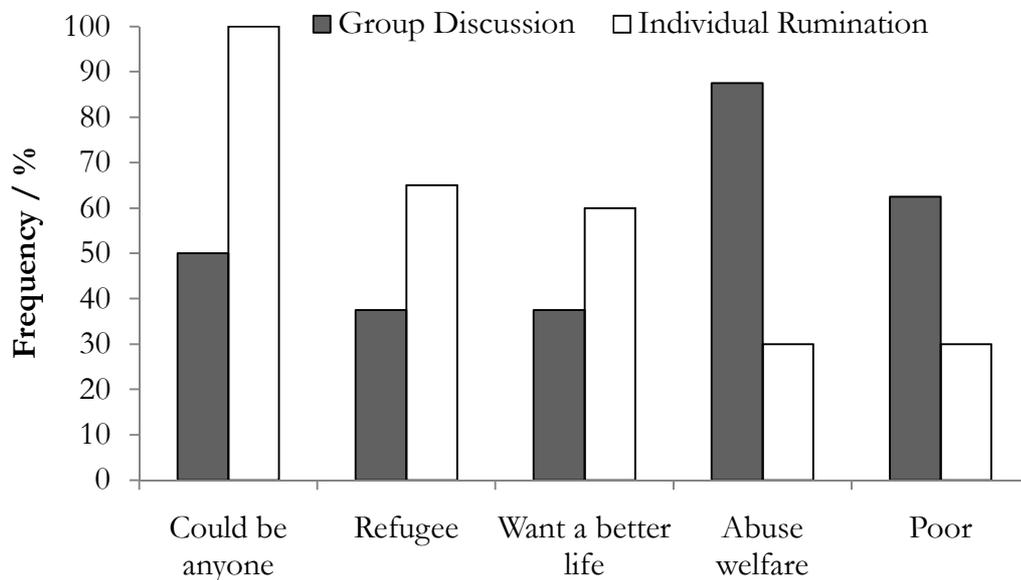


Figure 3.1. The percentage of participants who mentioned each key code, Study 3.1

HLM Analyses

In the initial model, action was the Level-1 dependent variable and the contrast the level-2 predictor. As summarised in Table 3.1, there was significantly more endorsement of hostile social policy in the experimental condition, $\gamma=.50$, $p=.04$. There was also a significant effect of condition on the group norm, $\gamma=1.35$, $p=.03$, with a greater perceived normalisation of hostile social policy in the discussion condition than the control condition. Condition had no significant effect on generalised prejudice ($p>.20$). There was also no difference in overall social identification or on the subscales between the two conditions.

Mediation. There were significant condition effects on the norm, but not on any of the other potential process variables. Therefore, we tested for mediation by entering

the norm into the multi-level model as a level-1 grand centered effect, predicting hostile policy endorsement at Level 1, with the contrast at Level 2. The slope for the group norm was significant, $\gamma=.20, p<.01$. On entering this mediator into the model, the difference between the experimental condition and the control condition became non-significant, $\gamma=.23, p=.33$ (Figure 3.2), suggesting that the norm mediated the condition effect, as confirmed by a significant Sobel test, $Z=2.01, p=.04$.⁶

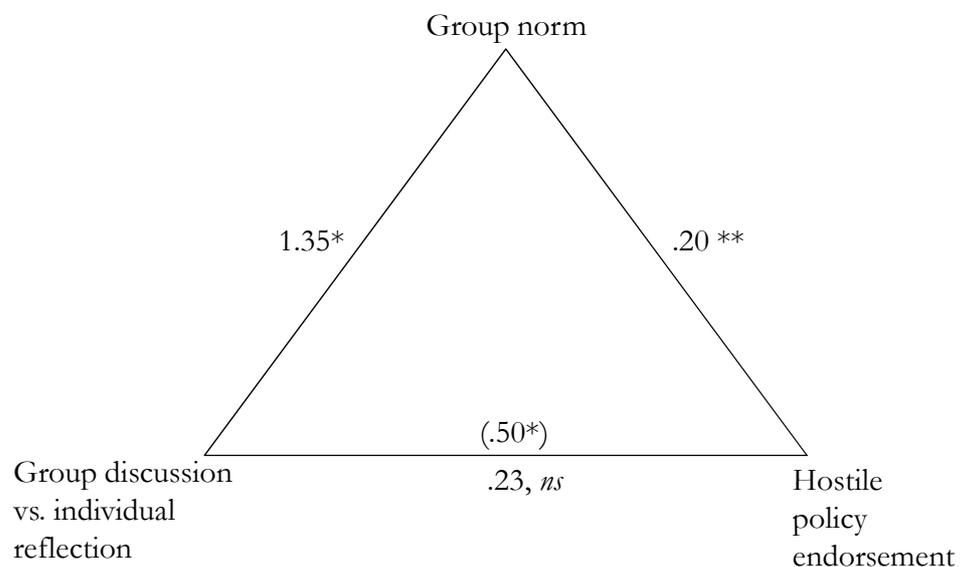


Figure 3.2. Gamma coefficients for the mediation model, Study 3.1 (* $p<.05$, ** $p<.01$)

Discussion

Study 3.1 examined the effects of engaging in a group discussion on social action tendencies. Results showed that group discussion of the out-group stereotype

⁶ We used the multilevel first-order Taylor series approximation to estimate the standard error of the mediated effect as is recommended for multilevel mediation with a Level-2 predictor, Level-1 mediator, and Level-1 outcome (see Krull & MacKinnon, 1999).

led to more support for hostile policies than did individual reflection. This effect was mediated by the in-group norm. Could group polarisation account for this result? It has often been demonstrated that group discussion can polarise individual attitudes that are already tending towards a particular extreme (e.g., Moscovici & Zavalloni, 1969). In the present study, there was no significant difference in post-intervention prejudice, and hence there was no evidence that polarisation occurred, nor that it mediated the effect of condition on policy support. Indeed, given that the average prejudice was close to the midpoint of the scale (i.e., not already tending towards an extreme), strong attitude polarisation would not normally be expected. Nevertheless, there was a descriptively small difference between conditions which will be followed up in the next studies as a potential mediator.

What was very clear was that discussion caused a normative shift, and this shift fully mediated the condition effect on policy support. This effect replicates the paradoxical public opinion data, which showed that the Dutch national debate about immigration preceded political and behavioural change, but with no accompanying increase in anti-immigrant attitudes (Dagevos & Gijsberts, 2007). In both cases discussion appears to have liberated citizens from the restraint imposed by “generic” societal norms. Indeed, the qualitative data supports the view that intra-group discussion informed the content of the out-group stereotype, such that participants became less resistant to stereotyping and less sympathetic towards the out-group. Thereby, intra-group interaction enhanced hostility, qualitatively transforming the psychological reality of the inter-group context for in-group members.

STUDY 3.2

The effects in Study 3.1 may be accounted for by two related but conceptually distinct processes; 1) norm formation may have occurred because interaction bolstered

a sense of shared in-group identity (cf. Postmes, Spears, Lee, & Novak, 2005), and 2) norm formation may have occurred because inter-group stereotypes were validated. These stereotypes may have been thus transformed from tentative beliefs, into social perceptions that *ought* to be true. A further process that may have played a role is that potential policies may have been explicitly discussed in groups: because these discussions were more outspoken and extreme than individual reflection, they would lend themselves better to the inference (and discussion) of specific actions.

These possibilities informed design choices of the follow-up study. In comparison with Study 3.1, a different control condition was used. In order to assess whether group discussion per se may bolster a sense of shared identity which leads to the inference of shared norms and solidarity, the control condition had groups also interact but on an unrelated topic. In order to assess whether stereotype validation could account for effects, we adjusted the norm measure. In order to assess whether the discussion of actions could account for the effects, we added a condition in which participants were explicitly asked to do this. Finally, the study used a more concrete action intention measure.

Four hypotheses were tested: (a) after discussing an out-group stereotype, action intentions towards that group would be more hostile than after an irrelevant discussion, (b) after discussing the stereotype plus what actions to take towards the out-group, action intentions would be more hostile than after an irrelevant discussion, and, (c) action intentions after discussing the out-group action would be more hostile than after discussing the stereotype only. Furthermore, (d) the effects of condition on hostile intent would be mediated by a sense of social validity.

Method

Participants and Design

$N=75$ White British sixth form students volunteered from a different state comprehensive school to Study 3.1, recruited during an educational visit. The mean age was 17.0, 31 participants were female. Participants were randomly allocated to 25 groups of three. All were unpaid and naïve as to the purpose of the study. Consent and debriefing arrangements were identical to Study 3.1. The design was a 3-condition (Topic of interaction: out-group stereotype vs. out-group stereotype plus a plan for social action vs. irrelevant control) between-subjects design.

Procedure

Participants engaged in a 5-minute discussion. Instructions varied by condition. In the *Out-group stereotype* condition, instructions were identical to the discussion condition of Study 3.1, e.g., “Formulate an impression of immigrants...” In the *Stereotype plus plan for action* condition, instructions were identical for the first part: “Formulate an impression of immigrants and a plan to combat any problems immigration causes”, followed by a similar set of questions. In the *control* condition participants discussed an irrelevant topic: “Please discuss in your group, whether or not you think that the British monarchy is outdated.” Unlike Study 3.1, participants were not instructed to write down responses, and there was no qualitative data to analyse.

Dependent Measures

The measure of hostile action intentions was a 4-item scale designed to measure whether or not participants would be prepared to engage in concrete and direct hostile action against the out-group ($\alpha=.76$). In order to ensure that it was completely unambiguous whether they supported action or not, participants responded 1=“No”, or 2=“Yes” to the questions, “I agree to participate in a demonstration about immigration policy in Britain,” “I agree to have my name added to a petition to make the government listen to my group’s views,” “I would do something with fellow group

members to address the problems that immigration brings to Britain,” and “I would participate in raising our collective voice to address the problems that immigration brings to Britain.”

For all subsequent items, participants indicated agreement on 7-point Likert-type scales (1=“strongly disagree,” 7=“strongly agree”). Social validation was measured to tap into the process by which group norms exert their influence, i.e., through the legitimisation and validation of views about immigration. A further reason to change this measure was an in-group norm measure would not have been appropriate for use in the control condition of the present study. The social validation scale consisted of 3-items ($\alpha=.79$): “The discussion made me feel that my views are more legitimate”; “My beliefs about immigration are justified” and “I feel that my views on immigration are well-founded”.⁷ A factor analysis with oblique rotation on the items in the validation and hostile action scales extracted two distinct factors, suggesting that the outcome and process measures are statistically distinct constructs. The action intentions factor explained 45% of the variance, $\lambda=3.13$; social validation explained a further 24% variance, $\lambda=1.67$. The inter-factor correlation was relatively low ($r=.35$).

To assess effects of the condition in which stereotypes plus action were discussed, a single item measured participants’ certainty about future courses of action “I have a clear idea about where things need to go with immigration in this country”. The scale of generalised prejudice was the same as in Study 3.1 (10-items, $\alpha=.88$). A single item measured social identification: “I identify with the other members of my

⁷ Reviewers expressed concern about mixing items referring to the discussion with items which did not refer to the discussion but did refer to immigration. We verified that there were no major between-condition differences in scaling, and that alpha coefficients were acceptable to good in each individual condition.

group”. Inter-scale correlations (Table 3.2) suggest that the discriminant validity of scales was acceptable.

Analytic Strategy

HLM analyses were performed as before. Hypotheses were tested via comparisons made using sets of orthogonal dummy and contrast variables, created from the three conditions as recommended by Cohen et al., (2003). The dummies were coded so that the control condition was the referent group. Therefore, the dummies for the stereotype, action and control condition respectively were: [0,1,0] and [0,0,1]. The orthogonal contrasts ([1/3, 1/3, -2/3] and [1/2, -1/2, 0]) were coded to make the remaining comparison between the stereotype and action condition.

Results

In the initial model, action intention was entered at level one and a set of orthogonal dummies at level two. This initial model was significantly different to the null model, $\chi^2(24)=49.71, p<.01$, indicating significant between-condition differences overall. Please refer to Table 3.2 for means. Action intentions were significantly lower in the irrelevant discussion condition compared to the out-group stereotype condition, $\gamma=.93, p<.01$. Intentions in the irrelevant discussion condition were not significantly different to the action condition $\gamma=.52, p=.37$. Action intentions were also significantly higher after discussion of the stereotype only, than after discussion of the stereotype plus courses of action, $\gamma=1.13, p=.01$. Paradoxically, this demonstrates that discussing the out-group stereotype enhanced action intentions more successfully than discussing a specific “plan of attack”.

The effects of condition on social validation were then explored, $\chi^2(24)=37.93, p=.03$. Participants felt more validated in the stereotype condition than the irrelevant condition, $\gamma=.51, p<.01$, and in the action condition than the irrelevant condition

$\gamma=.57, p=.04$. Although the mean was higher in the stereotype condition than the action condition, the difference was not quite significant, $\gamma=.47, p=.11$. These results suggest that the topic of interaction also affected the extent to which group members felt validated in their views of immigrants.

A potential explanation for the unexpected effects in the action condition was suggested by participants' responses to the item measuring certainty about actions to deal with immigration. On this item, there was somewhat more certainty in the stereotype only condition than the irrelevant condition, although not significantly so, $\gamma=.38, p=.10$. Notably, there was significantly *less* certainty in the action condition than in the irrelevant condition, $\gamma=.57, p=.04$. There was also less certainty in the action condition than when groups discussed the stereotype only, $\gamma=.95, p=.03$. As in Study 3.1, there were no condition effects for identification or generalised prejudice (all p 's > .08).

Mediation. Validation was predicted to mediate. Preconditions for this were met: validation showed the same pattern of between-condition differences as action intentions, and the relationship between validation and action intentions was marginally significant, $\gamma=.29, p=.06$. Thus, validation was entered as a grand centered level 1 predictor, with action intentions as the outcome variable. The dummy variables were entered separately at level 2. The slope for validation was significant, $\gamma=.36, p<.01$. On entering this mediator into the model, the difference between the stereotype and the irrelevant discussion conditions reduced, from $\gamma=.93, p<.01$ to $\gamma=.77, p=.01$. Thus, validation partially mediated the effect of experimental condition on action intentions; Sobel $Z = 2.24, p = .03$ (see Figure 3.3).

Table 3.2

Mean individual-level scores and scale intercorrelations, Study 3.2

		Stereotype only	Stereotype + Action	Irrelevant	1	2	3	4	5
1. Social Action	<i>M</i>	6.96 _a	6.11 _b	5.27 _b	(.76)	.17	-.11	.18	.16
	<i>SD</i>	1.23	1.57	1.49					
2. Social Validation	<i>M</i>	4.71 _a	4.26 _a	3.69 _b	.26**	(.79)	.06	.34**	.53**
	<i>SD</i>	1.09	1.35	.95					
3. Prejudice	<i>M</i>	4.40 _a	5.10 _a	4.79 _a	-.02	-.04	(.88)	-.03	.24*
	<i>SD</i>	1.36	1.23	0.81					
4. Identification	<i>M</i>	5.03 _a	4.94 _a	5.64 _a	.07	.20*	-.05		.43**
	<i>SD</i>	1.09	1.52	1.39					
5. Certainty	<i>M</i>	4.79 _a	3.85 _b	4.60 _a	.21	.52**	.16	.43**	
	<i>SD</i>	1.50	1.68	1.40					

Notes. A higher number indicates a higher propensity on each measure. Means on the same row with different subscripts differ significantly at $p < .05$. With the exception of single items, Cronbach's alphas are in parentheses. * $p < .05$, ** $p < .01$. Full scale intercorrelations are reported below the diagonal, and partial correlations (controlling for the effects of the manipulation) are reported above the diagonal.

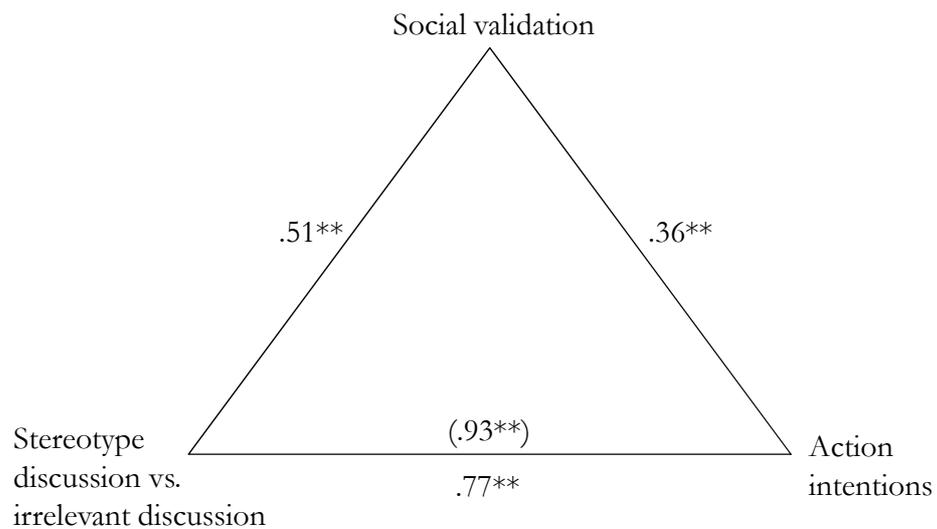


Figure 3.3. Gamma coefficients for the mediation model, Study 3.2 (** $p < .01$)

Discussion

In Study 3.2, the topic of interaction was varied to distinguish the effects of interaction *per se* from effects of discussing out-group stereotypes only, or discussing out-group stereotypes as well as potential courses of future action towards that group. First, it was hypothesised that there would be a greater propensity to endorse social action when participants discussed the out-group than when they discussed an irrelevant topic. Indeed, the lowest action intentions scores occurred in the irrelevant (control) condition compared with the experimental conditions. Participants in the control condition also scored lower on validation. It can be concluded that interaction with in-group members *per se* does not lead to greater preparedness for action.

It was also hypothesised that participants who discussed actions could be more prepared for them. Contrary to this prediction, discussing actions resulted in less hostile action intentions, and lower levels of social validation and collective efficacy.

Retrospectively, there appears to be one obvious explanation for this finding. It would seem to be easier to reach in-group consensus on an out-group stereotype than on a concrete plan of action. Indeed, discussions about courses of action *reduced* certainty about “what needed to be done about immigration,” compared with discussion of only the stereotype. It may be speculated that the discussion of acts ordinarily considered illegitimate or impractical ended up disempowering the group, therefore undermining any capacity for collective action. The implicit action intentions contained in stereotypes (e.g., in the sense that negative stereotypes are associated with prejudicial behaviour, Reicher et al., 1997) might be more effective in suggesting courses of action that, when actively voiced, could be the subject of intra-group disagreement.

Social validation appeared to be an important psychological product of intra-group interaction and mediated the effects. Thus discussions validated participants' stereotypes, thereby providing a firmer foundation for their intentions to engage in hostile social action. Their social perceptions were, in a sense, transformed into a social reality upon which they could legitimately act. This result points to the importance of consensus as a source of validation and action confidence (see also Baron et al., 1996; Luus & Wells, 1994; Petty, Briñol, & Tormala, 2002).

In conclusion, these results provide further support for the idea that intra-group interaction can transform moderately negative individual attitudes into intentions for hostile collective behaviour. The results also support the argument that social validation of out-group stereotypes plays a key role in the process, effectively justifying hostile action intent. The next study looked at actual behaviour.

STUDY 3.3

This study combined the designs of the previous two studies, by independently manipulating consensualisation (group discussion vs. individual reflection) and the

topic of discussion (out-group stereotype only vs. out-group stereotype plus action). The aim was to replicate the previous findings with a behavioural measure of hostile social action. As previous research has shown (e.g., Heckhausen & Kuhl, 1985), it cannot be taken for granted that self-reported action intentions will always translate into actual behaviour.

It was hypothesised that (a) there would be more hostile action after group discussion than after individual reflection, and (b) discussion of the out-group stereotype would produce more hostility than discussion of the stereotype plus specific actions.

Method

Participants and Design

78 White British sixth form students (Mean age=16.7; 77% female), recruited during an educational visit to a different state school to Studies 1 and 2, were randomly-allocated to 26 groups of three. Consent and debriefing arrangements were the same as for the previous studies. The study had a 2 (Type of reflection: Group discussion vs. individual reflection) x 2 (Topic of reflection: Out-group stereotype vs. out-group action) between-subjects design.

Procedure

Participants were asked to either undertake a 5 minute discussion in small groups or reflect on the specified topic alone for 5 minutes, and write down the key points that arose, using the same feedback for each condition as Study 3.1-3.2. They then completed the dependent measures individually and in silence.

Dependent Measures

The main dependent variable was a behavioural measure. Participants were asked to vote in a fictitious "UK Youth Parliament Election." They placed a vote for a

candidate of one of the following political parties: Labour, Conservative, the Liberal Democrats, or the UK Independence Party (UKIP). Fictional male candidates were presented as candidates for each party. A paragraph described each party's immigration policy. The policies described were genuine party statements taken from the respective parties' programs, and represented a continuum of political ideologies, from left- to right-wing. The Liberal Democrats were relatively pro-immigration. Labour and Conservative policies were roughly equivalent centrist ones, stating the need for balancing the benefit of immigration with control of the country's borders. The UKIP policy was explicitly anti-immigration. We emphasise that these are actual policies of the main political parties in Britain—the only aspects of the feedback that were fictional were the names of the candidates and the election they featured in. Votes were subsequently classified as hostile towards the out-group (UKIP vote: coded "1") or not hostile towards the out-group (not a UKIP vote, coded "0").⁸

The social validation scale used in Study 3.2 was expanded to 6-items ($\alpha=.88$), "I feel that my opinions on immigrants are valid"; "I feel that my views on immigrants are well-founded"; "My beliefs about immigrants are justified"; "I am certain that my views on immigration are right"; "I feel that my opinions about immigration are shared by many"; and "I feel that I have a justified opinion on immigration". The same measure of identification was included as in Study 3.1 (7 items, $\alpha=.77$). Also included was the 10-item generalised prejudice scale ($\alpha=.92$). Inter-scale correlations suggested that scales had adequate discriminant validity.

Results

Discussion Content

⁸The full range of voting behaviours were measured, but were not considered to add anything of value in terms of the analyses reported here. The results are available on request.

A content analysis was performed, following the same procedure as Study 3.1. Inter-rater agreement was 96%, Cohen's (1960) kappa coefficient ranged from $\kappa=.78$ to $\kappa=1.00$. As expected, representations of immigrants differed considerably across conditions. Participants who engaged in a discussion about the out-group stereotype generated a relatively consensual stereotype: 71% believed immigrants entered the UK to find work, claimed on the welfare system and argued that they were not well integrated in society, "Sometimes immigrants group together and form ghettos ". 43% believed immigrants clashed with British culture: "They don't respect our rules and they are trying to make everyone go under their rules in our country", and "Usually intimidating to true citizens and have too much freedom to exploit their own beliefs. If they go to a different country, they should live by the ways of that country". 14% linked immigration to terrorism, and 14% mentioned hostile actions that could be taken against the out-group, despite not being asked to discuss action, e.g.,: "We should limit the numbers let in to the country, tighter controls".

Participants who also discussed action had a similarly consensual stereotype, with a majority mentioning immigrants coming to the UK for work or welfare (57%). However, there was evidence of disagreement in plans for action towards immigration: although 29% of participants in this condition generated a plan for hostile action against the out-group, 43% also formulated a benevolent plan, for example, "Encourage multi-faith and multi-racial activities to bring people together".

Participants who reflected individually had a less homogenous and less consensual view of immigrants. Participants were more likely to characterise immigrants as refugees (40%). Furthermore, fewer participants in this condition (45%) displayed hostility towards immigrants relative to other conditions (57-70%).

HLM Analyses

Dummy variables were created to test main effects and the interaction orthogonally. Following this, tests of specific hypotheses were carried out through using two sets of 3 orthogonal contrasts, separately in the pre-consensus and consensus phases. Table 3.3 provides the coding of the contrast and dummy variables to aid interpretation.

Hypotheses tested with contrasts were that: (a) discussion conditions would increase UKIP votes relative to the individual reflection conditions, and (b) the out-group stereotype discussion would increase the number of UKIP votes relative to the stereotype plus social action discussion. Dummy and contrast variables were entered as uncentered variables at level 2. Identification, generalised prejudice and social validation scores were then analyzed using the HLM strategy described for Studies 3.1 and 3.2.

A Bernoulli sampling HLM analysis using a logit link function was conducted to assess the magnitude of variation in UKIP votes (a binary measure) between conditions. This procedure examined voting behaviour as a function of condition. This analysis yields the probability that participants cast an anti-immigrant vote (i.e. the odds of voting UKIP rather than another party). Overall, the analysis showed significant between-condition differences, $\chi^2(25)=36.84, p=.05$. There was a main effect for topic of reflection, with more UKIP votes when the stereotype was reflected upon than when the stereotype plus actions were considered, $\gamma=1.18$, log-odds=3.27, $p=.05$. The main effect for type of reflection was not significant, $\gamma=.11$, log-odds=1.12, $p=.85$, contrary to hypothesis (a). The 2-way interaction was also not significant $\gamma=.94$, log-odds=2.55, $p=.12$.

Table 3.3

Coding of dummy and contrast variables for Bernoulli HLM, Study 3.3

		Main effects and interaction			Contrast A: Stereotype only vs. action conditions			Contrast B: Discussion versus individual reflection conditions		
		Dummy variables			Contrast variables			Contrast variables		
Condition		D1: Main effect 1	D2: Main effect 2	D3: Interaction	C1	C2	C3	C1	C2	C3
Discussion	Stereotype	1/2	1/2	1/2	1/2	1/2	0	1/2	1/2	0
	Stereotype + Action	1/2	-1/2	-1/2	1/2	-1/2	0	-1/2	0	1/2
Individual reflection	Stereotype	-1/2	1/2	-1/2	-1/2	0	1/2	1/2	-1/2	0
	Stereotype + Action	-1/2	-1/2	1/2	-1/2	0	-1/2	-1/2	0	-1/2

Table 3.4

Frequency of UKIP votes, mean individual-level scores and scale intercorrelations, Study 3.3

	Group Discussion		Individual Reflection		1	2	3	4
	Stereotype	Stereotype + Action	Stereotype	Stereotype + Action				
1. UKIP votes (%)	54.5 _a	10.5 _b	35.0 _{ab}	26.3 _{ab}		.17	.29*	.42**
2. Identification	<i>M</i> 5.25 _a	5.32 _a	5.72 _a	5.15 _a	.18	(.77)	.28*	.09
	<i>SD</i> 1.10	0.91	0.78	0.81				
3. Social Validation	<i>M</i> 4.59 _a	4.61 _a	4.83 _a	4.58 _a	.28*	.30**	(.88)	.18
	<i>SD</i> 1.16	1.04	1.50	1.32				
4. Prejudice	<i>M</i> 4.59 _a	4.42 _a	4.73 _a	4.81 _a	.39**	.11	.18	(.92)
	<i>SD</i> 1.24	1.21	1.26	1.38				

Notes. A higher number indicates a higher propensity on each measure. Cronbach's alphas are in parentheses. * $p < .05$, ** $p < .01$. Means on the same row with different subscripts differ significantly at $p < .05$. Full scale intercorrelations are reported below the diagonal, and partial correlations (controlling for effects of the manipulation) are reported above the diagonal.

Contrasts then tested the second hypothesis and other between-condition effects. The topic of reflection main effect appeared to be driven by a highly significant increase in probability of voting UKIP after the out-group stereotype discussion, which had the most UKIP votes overall (54.5%) compared to the action discussion, which had the least (10.5%; Table 3.3), $\gamma=2.12$, $\log\text{-odds}=8.33$, $p=.01$, confirming the prediction and replicating the results of Study 3.2. There was no difference between the two individual reflection conditions $\gamma=.25$, $\log\text{-odds}=1.28$, $p=.78$. There were no simple main effects for type of reflection in the stereotype only $\gamma=1.05$, $\log\text{-odds}=2.84$, $p=.18$, or action conditions $\gamma=-.83$, $\log\text{-odds}=.44$, $p=.35$.

Neither main effects nor 2-way interactions were significant for identification, social validation and generalised prejudice (all p 's $>.20$), although both prejudiced attitudes and validation were significantly positively correlated with voting UKIP. Please refer to Table 3.4 for means and correlations.

Discussion

The purpose of Study 3.3 was to replicate the results of the previous two studies with a behavioural measure. Overall, the pattern of results was as predicted. The first hypothesis, that there would be more hostile social action when participants engaged in a group discussion than when they reflected individually, was not confirmed, although the difference was in the predicted direction. As in Study 3.2, discussion of the stereotype as well as specific courses of action did not produce an increase in anti-immigrant voting behaviour.

The content analysis showed that after discussing actions, a substantial percentage of participants consensualised on benevolent plans for action. This suggests that discussing specific courses of action leads to less consensus and thereby undermines the general hostility that results from discussing only the stereotype. It

may be postulated that this is because out-group stereotypes (unlike specific actions) are to some extent societally shared, and hence easier to reach consensus on.

The second prediction, that there would be a difference in voting behaviour between the discussion conditions, received strong support. More specifically, discussion of the out-group stereotype led to more hostile behaviour than discussion of the stereotype plus actions. Neither validation, identification, nor generalised prejudice varied significantly across conditions. Therefore, changes in these factors cannot account for the patterns in voting behaviour displayed in this study. Nevertheless, it was apparent that the group discussions differentially affected the content of the out-group stereotype, providing group members with information about normative behaviour.

General Discussion

Three studies investigated the effect of intra-group interaction on hostile social action. It was predicted that intra-group consensualisation would increase in-group members' preparedness for inter-group action. The process by which this takes place, we proposed, is the validation of an internalised in-group norm for what are considered "appropriate" measures to address immigration issues.

The results of three studies tended to confirm these predictions. The main conclusions are that (a) Intra-group interaction informed the out-group stereotype, making it more concrete and more hostile, (b) More hostile action was directed towards the out-group when in-group members discussed only the out-group stereotype, because when they also consider specific courses of action, the group consensus breaks down, and (c) the reason why interaction has these profound effects is that it validates hostile stereotypical perceptions of immigrants, contributing to a normative climate in which hostile actions can be seen to be both supported and legitimate.

The Power of Talk

The existing literature on group discussion tends to focus on the polarising effects of group discussion. In the present research however, the overall effect of group discussion on attitude change was descriptively small: A meta-analysis across the three studies comparing the effects of stereotype discussion with the control condition shows the average weighted Cohen's D is essentially zero, $d = -.02$, Stoffer's $Z = 0.12$, *ns*. Therefore, it seems that informational social influence may not be able to account for these effects. Indeed, the results attest to the capacity of social interaction to validate the out-group stereotype, thereby establishing what is normative and what is not.

Past research has demonstrated that group discussion can inform stereotypes (e.g., Haslam, 1997; Stangor, Sechrist, & Jost, 2001; Thompson, Judd, & Park, 2000), and researchers have suggested that such consensualised stereotypes provide a basis for subsequent inter-group behaviour (e.g., Haslam et al., 2002; Stott & Drury, 2004). The present data confirms this, and demonstrates that consensualisation about stereotypes does indeed mobilise social action. Moreover, it provides some initial evidence for the process by which this could happen: interaction may propel a set of (potentially implicit) shared beliefs about an out-group into a shared ideology for inter-group behaviour.

The central point of these studies is that they highlight the important role of intra-group consensualisation in the development of inter-group behaviour. The results suggest that interaction with in-group members can qualitatively transform the psychology of those individuals, and lead to the normalisation of hostile inter-group behaviours. Therefore, when investigating hostile inter-group behaviour, not only must the inter-group context be taken into account, but also how and why social reality may

be changed through intra-group dynamics (see also Postmes, Haslam, & Swaab, 2005; Chapter 2).

Limitations and Future Research

The study of intra-group interaction effects on inter-group behaviour is relatively embryonic. Perhaps inevitably at such a stage, the research raises many questions. One objective of follow-up research would be to introduce greater experimental control over the content of discussions. Even though our instructions sought to control the content of interaction by focusing them on specific target questions to be addressed, we did not standardise the content of group discussions as much as would have been possible with confederates, for example. Also, we did not have the technological means to record the discussions within so many parallel groups—in future, a content analysis of the actual discussion content would further corroborate inferences about the processes at work. Even so, the content analysis of what groups wrote down about the out-group afterwards (Study 3.1, Study 3.3) confirms that consensualisation took place, as predicted on the basis of past research.

A further challenge for future research is to determine whether the group dynamics documented here can also contribute to *improvements* in inter-group relations. In theory, there is no reason why the same principles operating to increase inter-group hostility could not also be mobilised to reduce it. Thus, discussions among members of deprived groups could equip them with the social solidarity and cognitive equipment to better resist stereotyping and prejudice (see also Stott & Drury, 2004). Furthermore, discussion within higher status groups can be equally productive, provided that a significant minority within that group has the courage to voice dissent and undermine perceptions of social consensus.

Conclusion

The present research examines the consequences of intra-group interaction for inter-group behaviour. Intra-group consensualisation leads to a sense of social validation, and this contributes to hostile intentions and actions. This implies that the in-group provides a forum within which the perceived legitimacy of hostile inter-group behaviour is negotiated. However, although the findings demonstrate that such debates may have problematic consequences, they simultaneously show that debates will only have such consequences to the extent that the discussion is allowed to focus on the narrow range of xenophobic prejudices upon which a majority can agree. The moment that more nuanced and complex issues of inter-group relations are discussed, intra-group debate ceases to be such a problem. Indeed, future research should explore under what conditions such nuanced discussions can provide integrative solutions for inter-group relations.

CHAPTER 4

The Role of In-group Consensualisation in Identity (In)Formation

“The perceived, expected or believed agreement of similar others in the same situation implies that our behaviour is a function of the *objective world* rather than our personal biases, prejudices and idiosyncrasies,” (Turner, 1991; p. 161, italics added).

“In order for perceivers to have confidence in the correctness of their beliefs – and therefore use them as a basis for coordinated and subjectively meaningful social behaviour – those beliefs need to be validated by other people [...] Consensus is an *outcome and expression of an identification process* that is capable of mobilising and motivating group members in a way that independent activity and an individual product never could,” (Haslam, Turner, Oakes, McGarty et al., 1998; p. 232; italics added).

As expressed in the opening quotes, social consensus about reality is often assumed to provide individuals with a basis for inferring the validity of their beliefs (Festinger, 1950; Turner, 1991). So far, the studies presented in this thesis have shown that when group members are asked to consensualise (that is, reach consensus through discussion) on the out-group and inter-group action, the social validation this provides transforms their perceptions of the group norm. The purpose of the research in this chapter is to demonstrate that it is the in-group *consensus* achieved (and its implications for the validation of reality) that is critical in the (in)formation of in-group identity content.

The Crucial Role of Consensus

The literature reviewed in Chapter 1 leads to the suggestion that individuals

engage in reality testing through processes of collective discussions and social interpretation. If an individual's perception of reality is validated socially, this social reality may be objectified (Turner, 1991). This knowledge that others share our cognitive perceptions of inter-group relations may increase our confidence in our beliefs and even accentuate them—for example, by polarising our social stereotypes of an out-group (e.g., Haslam, Oakes, Reynolds, & Turner, 1999). However, perceived objectivity, or the (tacit) belief in a social reality is *necessary* if we want to act on these beliefs. For example, members of low status groups need in-group consensus that there is pervasive discrimination against them before they can resist the discrimination collectively. Similarly, high status group members can not engage in hostile action towards another *group* on the basis of idiosyncratically held individual beliefs: If they seek to collectively stigmatise or disadvantage some out-group, they must do so on the basis of some shared understanding among in-group members. In other words, consensus is a valuable commodity for understanding the world, and it is invaluable for acting upon it. The process of reaching consensus on reality within one's group is the basis for social action. Isolated individual cognition can never provide this foundation (Haslam, Turner, Oakes, McGarty et al., 1998; p. 232).

Indeed, group members tend to have fairly consensual views about inter-group relations. More specifically, members of a given in-group may, in a context in which comparisons with other groups are invoked, have a relatively consensual idea of in-group norms, ideology, stereotypes (both of in-group and out-groups) and perceptions of the socio-structural context more generally. Indeed, several authors have noted the close correspondence between stereotypes, particular ideologies and certain normative implications that these may have for behaviour (Haslam, 1997; Jost & Banaji, 1994; Reicher, Hopkins, & Condor, 1997; Reynolds, Oakes, Haslam, Nolan, & Dolnik, 2000;

Tajfel, 1981; see Chapter 1). Against this backdrop, it is perhaps not surprising that the mere discussion of out-group stereotypes may be sufficient for the justification and legitimisation of inter-group violence (Haslam, Turner, Oakes, McGarty et al., 1998; Tajfel, 1981). Indeed, there is evidence that socially shared stereotypes can be used as a mobilisation tool. As the previous chapter has shown, increasing consensus through social interaction on an out-group stereotype may increase hostility and collective action intention towards the out-group (see also Stott & Drury, 2004).

However, the extent to which interaction was successful in increasing stereotype consensus was dependent upon social identity salience (Haslam, 1997; Haslam et al., 1999; Haslam, Turner, Oakes, McGarty et al., 1998; Haslam, Turner, Oakes, Reynolds et al., 1998). Haslam and colleagues argued that, in line with self-categorisation theory (SCT; Turner, 1985; Turner et al., 1987), this was because social identity salience increased in-group members' influence on each other by promoting perceptions of in-group homogeneity (via self-stereotyping) and thus increasing the group's capacity for mutual validation (Haslam, Turner, Oakes, McGarty et al., 1998).

Although this research on consensualisation represents a significant advance in our understanding of the importance of communication to social identity processes, there is scope for further elaboration of this perspective. The traditional approaches to studying social identity effects in groups assume that the influence of the group on individual cognitions can be explained by increases in the salience of a particular social identity (vs. personal or alternative social identities). For example, the opening quote by Haslam, Turner, Oakes, McGarty et al., (1998) conceptualises consensus as an "outcome and expression of an identification process". Therefore, in this view, consensus tends to be a dependent variable, rather than a *process* leading to identification (see Chapter 1). By limiting the conceptualisation of consensualisation (and for that

matter, communication) to a dependent variable, researchers limit the way in which it is studied. This circumvents the possibility of discovering that the process of consensualisation may allow the formation of novel identity content and identification with a new or transformed in-group. Throughout this thesis, we have aimed to study the effects of interaction and consensualisation on identity content formation, rather than the other way around. This allowed us to study the effects of the consensualisation process on perceived social realities.

Alternative Explanations and Empirical Challenges

The research described thus far has not sufficiently addressed the key underlying assumption that would make intra-group interaction such a powerful process: that the achievement of intra-group *consensus* transforms an agglomerate into a group entity. Alternative explanations do exist. For example, taking an attitude change perspective, one could suggest that listening to other group members' arguments changes the individual's perspective on inter-group relations (e.g., group polarisation), and it is this shift in individual beliefs (rather than their validation through consensus) which explains why people become more hostile.

In contrast to prior research, therefore, one would have to demonstrate that *achieving consensus* is an essential step towards producing the effects. In order to establish this, the level of consensus achieved during discussion would need to be controlled. Therefore, instead of engaging in direct interaction in the following studies, participants were shown films of a group of other in-group members having a discussion to consensus. The present research used this methodology to address several different concerns. First, Study 4.1 addressed (1) the question of whether *consensual* discussion of future courses of action promotes action intentions (in contrast to chapter 3), (2) the question of whether consensual discussion of even very extreme courses of action

(which are not ordinarily endorsed) promote action intentions, and (3) the question of whether these effects occur in a context which is lacking an inter-group dimension that is clearly defined *a priori*. Next, Study 4.2 then addressed the further question (4) of whether the process of reaching consensus is essential or not when persuasive arguments are exchanged.

(1) *The consensual discussion of actions.* One of the surprising findings of Study 3.2 was that intra-group interaction did *not* increase action intention when the group discussed potential courses of action relative to when they discussed the out-group stereotype. The post-hoc explanation of this finding was that although groups had no difficulty reaching consensus about the stereotype and its implied prejudices, they failed to reach consensus about potential courses of action. In order to investigate this explanation further, Study 4.1 controlled the level of consensus reached across conditions, and varied the content of discussion. The group either discussed the stereotype, or the stereotype plus specific actions.

Hypothesis 1: The consensual discussion of a course of action should promote action intentions in line with that course that are as strong, or stronger, than when the group had merely discussed an out-group stereotype.

(2) *Discussion of extreme actions.* We suggest that interaction can change perceptions of action legitimacy. Therefore, it is possible that interaction may increase perceptions of how acceptable “extreme” actions are. In order to investigate whether the legitimacy and validity of “extremism” can be negotiated, Study 4.1 included a condition in which discussants consensualised over actions which are ordinarily considered extreme. This condition was compared to conditions in which mainstream actions were discussed, as well as the conditions in which action was not explicitly mentioned.

Hypothesis 2a: Overall, it was hypothesised that the strongest intention for collective action would be when the consensual recommendations for action were mainstream (legal, and what we considered, reasonable ways to protest) rather than extreme (illegal, unreasonable). Crucially however, collective action intentions would vary significantly with how much the action discussed was perceived to be mainstream and legitimate (Hypothesis 2b), and with how much participants felt validated on this dimension by the in-group (Hypothesis 2c).

(3) *The prior inter-group context.* The research presented in Chapter 2 showed that in conditions of negative inter-group interdependence (obstruction), intra-group consensualisation had little effect on the discriminatory behaviours of the in-group. However, when no negative interdependence existed, intra-group interaction led to a marked increase in the willingness to punish the out-group. This finding was worthwhile of follow-up, since it addressed one of the core issues in current debates surrounding social identity theory (SIT; Tajfel & Turner, 1979; 1986). SIT is founded partly on the demonstration that, even in minimal group situations without negative inter-group interdependence or realistic group conflict, people are willing to discriminate in favour of the in-group. In light of findings that this is merely the result of “in-group love” (favouritism of in-group others) and not of “out-group hate” (see Chapter 2), questions have been raised over the ability of SIT to explain hostile inter-group behaviour. The suggestion of our previous results is, however, that intra-group debate may be sufficient to justify hostile inter-group behaviour (in the absence of prerequisite antecedents); but the question is why and how.

There is a further reason to study the effect of consensualisation on action in the absence of a pre-defined inter-group context. The target out-group in the studies in Chapter 3 was immigrants: a traditionally lower status out-group in the UK. This status

differential may explain why participants had difficulty reaching agreement over particular courses of action. Taking action against a lower status out-group raises all manner of concerns, not least because open hostility is potentially detrimental to the legitimacy of the high status groups' position (Postmes & Smith, in press; Turner, 2005). Therefore, the design of Study 4.1 aimed to remove this status differential by removing the presence of a pre-defined out-group and inter-group context.

Hypothesis 3: Intra-group interaction would increase action intent in the absence of a pre-defined inter-group context.

(4) *The process of consensualisation.* The arguments put forward suggest that in-group consensus is a necessary condition for the formation of norms for social action. An alternative approach may suggest that the exchange of persuasive arguments provides the necessary information for in-group members to form norms for social action. The purpose of Study 4.2 was to investigate whether the process of reaching in-group consensus is essential for norm formation or not, when (the same) persuasive arguments are exchanged. This should disentangle the role of persuasive argumentation from the role of consensualisation.

Hypothesis 4: The understanding that the in-group has reached consensus through discussion will lead to the formation of identity content, including norms for social action. Conversely, the undermining of consensus with intra-group dissent will make norm formation impossible.

Overview. Two studies investigated the role of intra-group consensus in normalising inter-group action from within the in-group. Study 4.1 was conducted with a non-specific out-group to investigate whether high intra-group consensus on action can increase the perceived legitimacy of collective action. Study 4.2 manipulated the process of consensualisation by showing a film of a discussion which began with intra-

group dissent and ended in consensus, or vice versa.

STUDY 4.1

Study 4.1 was conducted partly to investigate the paradoxical finding of Chapter 3, (Studies 3.2 and 3.3), that when groups discussed simply the out-group stereotype, they were more likely to act than when they discussed potential courses of action against an out-group. In a context within which there was no existing conflict, this study examined the role of high consensus in the way groups create and legitimise a hostile ideology. Through manipulating the content of a message to include validation of the in-group's opinion and endorsements of social actions which varied in legitimacy, the degree to which participants were willing to infer norms for social action (which may negatively affect the out-group) from the message was measured.

Method

Participants and Design

Participants were 89 university students recruited via email, randomly allocated to one of 4 experimental conditions (Content of filmed discussion: irrelevant control vs. stereotype only vs. stereotype plus mainstream action vs. stereotype plus extreme action). The mean age was 20.56 years old; 28 participants were male and 61 were female. All participants were unpaid and naive as to the purposes of the study.

Stimuli and Procedure

Participants were emailed a link to an online survey. After informed consent was obtained, this link randomly directed them to watch one of four films. The films lasted approximately 5 minutes and ostensibly showed a group of 5 fellow university students (one male and four female) sitting in a semi-circle, having an informal discussion. The content of the discussion was either irrelevant (whether or not the British monarchy is out-dated); a stereotyped discussion of fast food chains as a

societal problem; the stereotype plus ideas for mainstream action against fast food chains; or the stereotype plus ideas for extreme action. The discussion content was highly consensual and included affirming body language.

The videos containing stereotyped information focused on several well-known reasons why fast food chains allegedly cause societal problems. The reasons mentioned during discussion were: the health cost of obesity caused by the food they serve, cruelty against animals, the marketing of unhealthy food to children, the global monoculture they promote, and the unfair competitive advantage they gain through corporate legislation which favours them over alternative, smaller, food retailers.

The video with consensus on mainstream actions mentioned only legal actions: petitioning the government for enforced transparency of ingredients, personally boycotting fast food chains and encouraging others to join this boycott. The video in which extreme actions were advocated mentioned only illegal actions: Vandalising local fast food restaurants, personally harassing staff and their family, barricading the entrance to the restaurant, and setting battery hens free.

After watching the film, participants were asked, “Imagine you are involved in the discussion with the students on the film you just watched. What would you say when they asked your opinion?” They were also asked to outline the most compelling argument on the film. This was designed in order to increase participants’ perception of involvement in the discussion. Participants then completed the standardised scales.

Dependent Measures

Standardised scales were used (Table 4.1 for inter-scale correlations). Participants were presented with a series of statements, and asked to indicate on a 7-point Likert-type scale to what extent they agreed or disagreed (1=“not at all,” 7=“very much”). The primary dependent measure was a 9-item collective action intentions scale

($\alpha=.99$) with the items, “I would do something with fellow students to address the problems that fast food brings”; “I would personally boycott fast food”; “I would encourage others to boycott fast food”; “I agree to have my name added to a petition to raise awareness of the problems that fast food brings”; “I would participate in raising our collective voice to address the problems that fast food brings”; “Fast food restaurants should be allowed to continue as they are now” (reverse-coded); “I would participate in a demonstration against fast food corporations”; “I want to help to wipe out fast food outlets”; and, “I don’t want to do anything about fast food” (reverse-coded).

Next, a 6-item social validation scale ($\alpha=.81$) was included (adapted from Chapter 3). Items included, “I feel that my opinions on fast food are valid”; “I feel that my views on fast food are well-founded”; “My beliefs about fast food are justified”; “I feel that I have a legitimate attitude on fast food”; “I am certain that my views on fast food are right”; “I feel that my opinions about fast food are shared by many”; and, “I feel that I have a justified opinion on fast food”.

In order to assess whether the content of discussions impacted upon the extent to which actions were perceived as legitimate, we included a 5-item perceived legitimacy of collective action scale ($\alpha=.99$). Items included, “I believe that it would be legitimate to take action against fast food outlets”; “It is not fair to cause any problems for fast food restaurants” (reverse-coded); “It is OK to take action against fast food”; “It is justified to let fast food chains continue serving food as they are now” (reverse-coded); and “We would be justified in taking action against fast food chains”.

A 4-item scale was designed to measure the extent to which the views expressed on the film seemed normal ($\alpha=.86$). This was included to investigate whether perceptions of the in-group norm altered with discussion content. Items were: “The

people on the film seem to hold normal views”; “The views expressed in the film seemed excessive” (reverse-coded); “The people on the film hold the views of someone like me”, and, “The views expressed in the film seemed extreme” (reverse-coded).

Furthermore, a 3-item scale was included to measure the extent to which participants endorsed the message in the film ($\alpha=.90$): “I support the ideas expressed in the film”; “I tend to think that the message in the film is right”; and, “I would join the people in the film in the course of action they recommend”.

Inter-group anger, social support and collective efficacy are common factors often forwarded in the literature as being important to undertaking collective action (e.g., Van Zomeren, Spears, Leach, & Fischer, 2004). In order to test their relative contribution to action intentions, a 4-item inter-group anger scale ($\alpha=.99$) was included, for example: “The problems fast food brings to the country make me furious”; “I feel irritated about the problems fast food cause”; “I would tend to feel angry about the problems fast food cause”; and, “I tend to feel incensed by fast food culture”.

Additionally, a 5-item social support scale ($\alpha=.99$) included the items, “I feel supported in my beliefs about fast food”; “Other students back me up in my beliefs about fast food”; “Others support me in my views about fast food”; “Others would support my ideas about what to do about fast food”; and, “My own views on fast food are supported by many”.

Finally, a 6-item adaptation of Ellemers, Kortekaas and Ouwerkerk’s (1999) measure of identification was included ($\alpha=.88$). Items were, “I am like the people on the film”; “The people on the film are a reflection of who I am”; “I think the people on the film have little to be proud of” (reverse-coded); “I feel good about the people on the film”; “I have little respect for the people on the film” (reverse-coded); “I would

rather not belong to the same group as the people on the film” (reverse-coded). Data were analysed using univariate analyses of variance (ANOVA). Simple cell comparisons were made using contrast tests, as recommended by Rosenthal and Rosnow (1985).

Results

A one-way analysis of variance (ANOVA) comparing 4 conditions (Content of discussion: irrelevant control vs. stereotype only vs. stereotype plus mainstream action vs. stereotype plus extreme action) was conducted on collective action intention scores. There was a significant omnibus effect of condition, $F(3, 84)=3.08, p=.03, \eta_p^2=.10$. As can be seen in Figure 4.1, the means were in the predicted direction (also Table 4.1), with the control condition containing the least action intention ($M=3.29, SD=1.57$), strongest advocacy of social action in the mainstream condition ($M=4.44, SD=1.50$), which was closely followed by the extreme condition ($M=4.41, SD=1.62$), then the stereotype only condition ($M=3.88, SD=1.20$). Although the mainstream and extreme conditions were not significantly different from each other $F(1, 84)=.01, p=.95, \eta_p^2 < .01$, they were together significantly different from the control condition, $F(1, 84)=9.11, p<.01, \eta_p^2=.10$, but not from the stereotype only condition $F(1, 84)=1.91, p=.17, \eta_p^2=.02$. The stereotype condition was not significantly different from the control condition $F(1, 84)=1.76, p=.19, \eta_p^2=.02$, which suggests that listening to a discussion with high consensus about the stereotype plus action was more predictive of action intention than listening to a discussion about only the out-group stereotype.

There was also a significant condition effect for the perceived legitimacy of action, $F(3, 84)=5.46, p<.01, \eta_p^2=.16$. The means (Table 4.1) followed the same pattern as scores for collective action intentions, with which they were highly correlated $r=.72, p<.01$. Collective action was perceived as most legitimate in the mainstream

condition ($M=4.90$, $SD=1.05$), followed by the extreme condition ($M=4.70$, $SD=1.24$), then the stereotype only condition ($M=4.66$, $SD=1.18$) with the control condition yielding the lowest scores ($M=3.59$, $SD=1.41$). The mainstream and extreme conditions were not significantly different from each other $F(1, 84)=.28$, $p=.60$, $\eta^2<.01$, they were together significantly different from the control condition, $F(1, 84)=15.13$, $p<.01$, $\eta^2=.15$, but not from the stereotype only condition $F(1, 84)=.20$, $p=.66$, $\eta^2<.01$. The stereotype only condition was also significantly different from the control condition $F(1, 84)=8.50$, $p<.01$, $\eta^2=.09$, which suggests that listening to a discussion with high consensus about the stereotype alone or the stereotype plus action can increase action intention.

There was a similar condition effect on social validation $F(3, 84)=7.42$, $p<.01$, $\eta^2=.21$. Again, the means followed the same pattern as scores for collective action intentions, with which they were significantly correlated $r=.52$, $p<.01$. The highest social validation was reported in the mainstream condition ($M=5.87$, $SD=.77$), followed by the extreme condition ($M=5.63$, $SD=.86$), then the stereotype only condition ($M=5.56$, $SD=.74$) with the control condition yielding the lowest scores ($M= 4.70$, $SD=1.17$). Interestingly, the mainstream and extreme conditions were considered equally validating $F(1, 84)=.75$, $p=.39$, $\eta^2<.01$. The two action conditions were together significantly different from the control condition, $F(1, 84)=20.67$, $p<.01$, $\eta^2=.20$, but not from the stereotype only condition $F(1, 84)=.56$, $p=.46$, $\eta^2<.01$. The stereotype only condition was also significantly different to the control condition $F(1, 84)=10.25$, $p<.01$, $\eta^2=.11$, which suggests that listening to a discussion with high consensus about the stereotype alone or the stereotype plus action can be similarly validating.

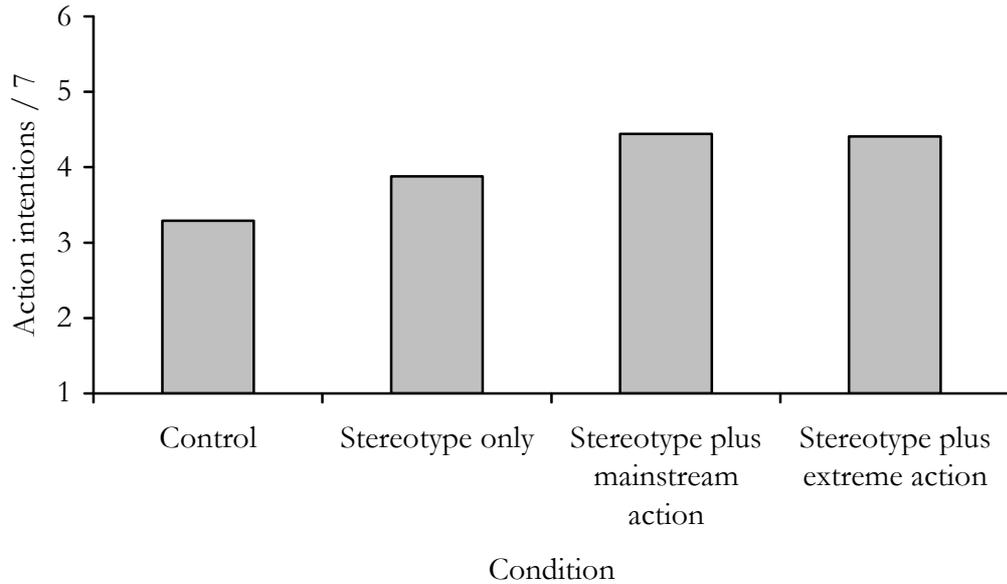


Figure 4.1. Mean collective action intention scores, Study 4.1.

In addition, there was an omnibus condition effect for endorsement of the message in the film $F(3, 84)=8.11, p<.01, \eta^2=.23$. The message was endorsed most in the mainstream condition ($M=5.52, SD=1.21$), followed by the stereotype only condition ($M=5.35, SD=1.20$), then the control condition ($M=4.09, SD=1.34$) with the extreme condition yielding the lowest scores ($M=3.94, SD=1.59$). The message in the mainstream condition was significantly more endorsed than that in the extreme conditions $F(1, 85)=14.62, p<.01, \eta^2=.15$ and the control condition $F(1, 85)=13.20, p<.01, \eta^2=.13$. The mainstream condition was not significantly more endorsed than the stereotype only condition $F(1, 85)=.18, p=.68, \eta^2 < .01$, however. The stereotype only condition was endorsed significantly more than the extreme condition $F(1, 85)=11.58, p<.01, \eta^2=.12$ and the control condition $F(1, 85)=10.18, p<.01, \eta^2=.11$. The extent to which the extreme and control conditions were endorsed was not significantly different $F(1, 85)=.15, p=.70, \eta^2<.01$.

Table 4.1.

Mean scores and scale intercorrelations, Study 4.1

		Control	Stereotype only	Stereotype plus mainstream action	Stereotype plus extreme action	1	2	3	4	5	6	7	8
1. Collective action intentions	<i>M</i>	3.29 _b	3.88 _{ab}	4.44 _a	4.41 _a	(.94)							
	<i>SD</i>	1.57	1.20	1.50	1.62								
2. Social validation	<i>M</i>	4.70 _b	5.56 _a	5.87 _a	5.63 _a	.52**	(.81)						
	<i>SD</i>	1.17	0.74	0.77	0.86								
3. Legitimacy of action	<i>M</i>	3.59 _b	4.66 _a	4.90 _a	4.70 _a	.72**	.56**	(.99)					
	<i>SD</i>	1.41	1.18	1.05	1.24								
4. Social support	<i>M</i>	4.42 _a	4.88 _a	5.20 _a	5.00 _a	.44**	.61**	.48**	(.99)				
	<i>SD</i>	1.45	0.74	0.79	0.81								
5. Inter-group anger	<i>M</i>	3.55 _a	3.98 _a	4.42 _a	4.46 _a	.83**	.55**	.65**	.43**	(.99)			
	<i>SD</i>	1.67	1.40	1.53	1.68								
6. Social identification	<i>M</i>	4.43 _a	4.71 _a	4.90 _a	4.22 _a	.49**	.29**	.38**	.24*	.52**	(.88)		
	<i>SD</i>	0.95	1.00	1.24	1.46								
7. Endorsement of message	<i>M</i>	4.09 _b	5.35 _a	5.52 _a	3.94 _b	.46**	.27**	.54**	.30**	.41**	.64**	(.90)	
	<i>SD</i>	1.36	1.20	1.21	1.59								
8. Characterisation of message as normal	<i>M</i>	5.60 _{ab}	4.63 _{ab}	5.24 _b	3.96 _a	.33**	.01	.22*	-.02	.41**	.54**	.55**	(.86)
	<i>SD</i>	0.87	1.68	1.35	1.62								

Notes. A higher score indicates a higher propensity on that measure. * $p < .05$, ** $p < .01$. Cronbach's alphas are in parentheses. Means in rows with different subscripts differ at $p < .05$.

Finally, there was also an omnibus condition effect for characterisation of the message as “normal”, $F(3, 84)=5.95, p<.01, \eta_p^2=.18$. The message was characterised as most normal in the mainstream condition ($M=5.24, SD=1.35$), followed by the stereotype only condition ($M=4.63, SD=1.68$), with lowest scores in the extreme condition ($M=3.96, SD=1.62$). The message in the mainstream condition was perceived as significantly more normal than in the extreme condition, $F(1, 84)=8.81, p<.01, \eta_p^2=.10$. The other simple cell comparisons were not significant (all p 's $>.10$). On all other dependent variables, there were no significant condition effects, neither for inter-group anger, $F(3, 84)=1.72, p=.17, \eta_p^2=.06$; social support, $F(3, 84)=2.44, p=.07, \eta_p^2=.08$; or social identification, $F(3, 84)=1.41, p=.25, \eta_p^2=.05$.

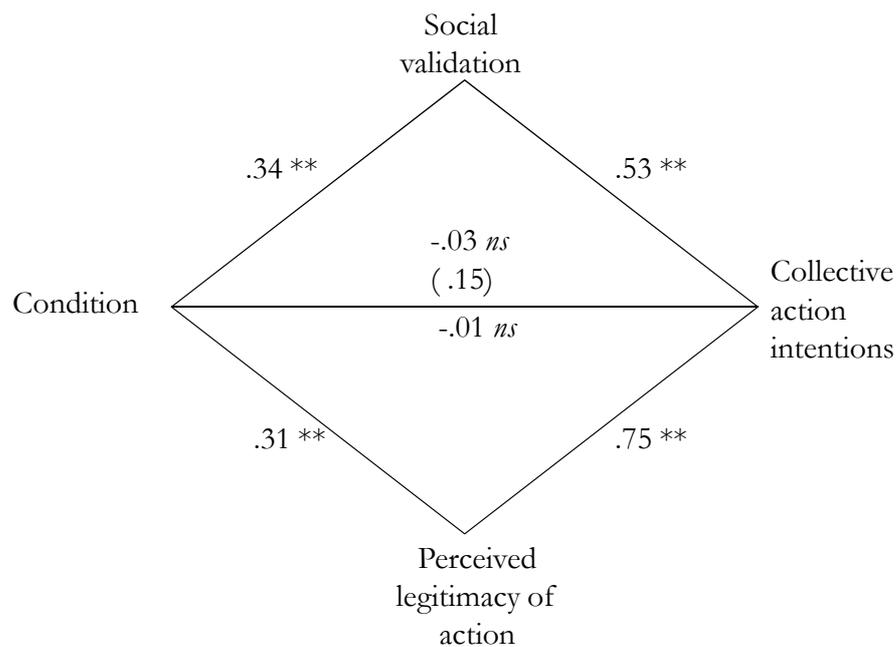


Figure 4.2. Beta coefficients for the mediation model, Study 4.1 (** $p<.01$)

Mediation. Legitimacy and social validation scores showed the same pattern of results across conditions as action intentions, with which they were significantly positively correlated. Both variables are likely candidates for mediation. In order to

examine this, we conducted mediation analyses (Figure 4.2).⁹ On entering legitimacy into the model, the omnibus effect of condition on hostile action intentions disappeared, $F(1, 83)=.79, p=.50, \eta^2=.03$. Legitimacy was a significant covariate, $F(1, 83)=77.59, p<.01, \eta^2=.48$, and the mediation effect was significant too, Sobel $Z=2.47, p=.01$, bootstrap confidence interval .11 to .55. This suggests that perceived legitimacy of action mediated collective action intentions.

Upon entering social validation as a covariate, again the difference between conditions on collective action intentions disappeared, $F(1, 83)=.62, p=.60, \eta^2=.02$. Validation was also a significant covariate, $F(1, 83)=21.02, p<.01$, and the mediation effect was significant, $\eta^2=.20$ Sobel $Z=2.58, p<.01$, bootstrap confidence interval .11 to .45, suggesting that social validation also mediated collective action intentions. If both mediators were entered into the model at the same time, all paths were significant and the direct effect of condition on action intentions was reduced significantly, however validation (bootstrap confidence interval .01 to .22) appeared to influence this reduction slightly more than legitimacy (bootstrap confidence interval .08 to .51).

Discussion

The purpose this study was three-fold. First, it investigated the effect of the consensual discussion of actions on action intention, relative to the discussion of only the out-group stereotype. In contrast to the findings of Chapter 3, in-group interaction increased action intention when the group discussed potential courses of action and when they discussed only the out-group stereotype. It seems, therefore, that when participants perceived explicit in-group consensus about potential courses of action, they were able to infer norms for action. In other words, *consensus mattered*.

⁹ Testing mediation using a one-way analysis of variance has the advantage of being able to test for the mediator's ability to account for the entire pattern of mediation, in a way that is impossible using a regression approach with dummy coded variables.

The second purpose of this study was to investigate the effect of the discussion of extreme versus mainstream actions. As hypothesised, collective action intentions were greatest when the consensual recommendations for action were mainstream rather than extreme, and collective action intentions varied significantly with how much this action was perceived to be legitimate and was validated by the in-group. This suggests that the consensus information provided by the discussion increased participants' perceptions that social action was legitimate and valid, relative to the control condition. Furthermore, this research shows that high consensus surrounding even extreme inter-group acts can lead to the legitimisation of social action. In the extreme condition (which was seen as the least normative), social action was legitimised, validated and advocated *despite the fact* that the discussion was seen as extreme.

The third and final purpose of this study was to disentangle the influence of any prior inter-group context from the effect of group discussion, on hostile social action. The suggestion of our previous results (in particular Study 2.2) is that intra-group debate may be sufficient to justify hostile inter-group behaviour (in the absence of prerequisite antecedents). Indeed, in this study intra-group interaction increased action intent in the absence of a pre-defined inter-group context.

Although they are considered important predictors of collective action in the literature (see Van Zomeren, Spears, Leach & Fischer, 2004), there were no significant condition effects for inter-group anger, social support, or social identification. This replicates the results of Chapter 3. Although these are null effects, and therefore conclusions cannot be drawn, it is interesting that we repeatedly find that the effects of these factors are descriptively small. It seems that (in these experimental contexts), validation, legitimacy and the normalisation of action were more important considerations. It may be speculated that whichever mediators (or moderators) of

action are most significant is dictated by the specific political and/or socio-structural context in which the action is undertaken.

In sum, the research shows that the legitimacy of hostile behaviour is *negotiable*. Social action can be legitimised even in the *absence* of an existing antagonistic inter-group dynamic, and even when the actions discussed are considered extreme and are illegitimate (in the sense of being illegal). Although this study answers some of the alternative explanations mentioned in the introduction, it fails to address one further alternative explanation. In none of the studies conducted so far could we control the content of communications in such a way that a consensus experimental condition could be compared with a no-consensus control condition, whilst keeping the content of the information conveyed constant across conditions. Of course, the content of information is a key factor to control, e.g., because groups could polarise merely on the basis of the content of the information conveyed, irrespective of the group dynamic involved. Study 4.2 was designed to investigate this possibility.

STUDY 4.2

This study examines whether or not the process of consensualisation is critical in the legitimisation of a hostile ideology. The objective was to keep the information conveyed constant, whilst manipulating whether consensus would be achieved or not. To our knowledge, this is the first time that the *process* of consensualisation has been manipulated. Such a design allows one to test whether the effect of social interaction on action intention and identity content is driven by exposure to persuasive arguments alone, or by the perception of the process of reaching an in-group consensus on the out-group stereotype. As before, in this study participants watched a filmed discussion with controlled content. There were two versions of this discussion, both with the same arguments. However, the order in which arguments were presented was varied, so

that in one condition the discussion appeared to converge, whilst in the other the discussion diverged.

Method

Participants and Design

N=63 participants were recruited via university email lists and a social networking website. Mean age was 26.79 years, and 47 participants were females. 90.5% stated their nationality as British¹⁰. Participants were unpaid and naïve as to the purposes of the study, and were randomly assigned to one of three conditions of a one-factor (Direction of consensus: Consensus to dissent vs. Dissent to consensus vs. collective action base-rate), between subjects design.

Stimuli and Procedure

The study was conducted online. In the control condition, participants were informed that they were taking part in a study on opinions towards immigration and simply answered the questionnaire (which included only the relevant scales). This enabled us to collect base-line action intention scores. In the experimental conditions, participants were asked to watch a short film clip then answer the questionnaire. After informed consent was obtained, participants were asked to state whether or not they were British, in order to make national identity salient. In the experimental conditions, participants were asked first to watch a 5 minute film of a group of British people (i.e., in-group members) ostensibly discussing their opinions on immigration.

The film stimulus consisted of 6 British actors (3 male, 3 female) who held a scripted discussion about immigration in the UK. The statements in this script were collected from previous research conducted in a sample with a similar demographic to

¹⁰ All participants were included in the analyses regardless of their nationality, as analyses suggested that removing non-British participants did not make a qualitative difference to the results.

the participants (the results reported in Chapter 3). Each explicit statement about immigration was volunteered in previous research. Thus, participants were not exposed to any arguments that they could not potentially encounter in their everyday lives, and the experimenter's only involvement was to select a coherent and representative set of statements about immigrants, and order them so that they appeared to progress logically. Although both films contained the same set of statements, the order in which they were presented was systematically varied so as to show a discussion either from consensus to dissent, or from dissent to consensus. There were two parts to the discussion. In one part (X), the group was in dissent: two actors forwarded pro-immigration arguments (e.g., economic stability, asylum, diversity) whilst the other four actors disagreed with them. In the other part (Y), there was apparent consensus: four actors forwarded anti-immigration arguments (e.g., welfare fraud, illegal immigration, cultural clashes) whilst the other two actors remained unresponsive. In the dissent to consensus condition, the film was edited to show (X) then (Y). In the consensus to dissent condition, the film showed (Y) then (X). This gave the impression that the group either consensualised or did not consensualise, respectively.

In order to involve participants actively in the debate, they were asked, "Imagine you are discussing immigration with the people on the film you just watched. What would you say about immigration?" Participants then completed the questionnaire, which consisted of standardised scales.

Dependent Measures

Collective action intentions were measured with eight items ($\alpha=.86$), "I would do something to address the problems that immigration brings"; "I would encourage others to do something about immigration"; "I agree to have my name added to a petition to raise awareness of the problems that immigration brings"; "I would

participate in raising our collective voice to address the problems that immigration brings”; “Immigration policy should be allowed to continue as it is currently” (reverse-coded); “I would participate in a demonstration against immigration”; and reverse-coded: “I want to help to decrease immigration”; “I don’t want to do anything about immigration”.

A 3-item scale measured how much participants endorsed the opinions expressed in the film ($\alpha=.91$): “I support the ideas expressed in the film”; “I tend to think that the message in the film is right”; and, “I tend to agree with the ideas brought up in the filmed discussion”. A 4-item scale measured to what extent the views expressed in the film were normative ($\alpha=.83$). Items were, “The people on the film seem to hold normal views”; “The views expressed in the film seemed excessive” (reverse-coded); “The people on the film hold the views of someone like me”; and, “The views expressed in the film seemed extreme” (reverse-coded).

The social identification scale from Study 4.1 was expanded to include the item, “I identify with the people on the film” (7-items, $\alpha=.90$). The social validation scale was adapted from Study 4.1 to refer to immigration rather than fast food (6 items, $\alpha=.90$), e.g., “I feel that my opinions on immigrants are valid”. All items were answered on a Likert-type scale, where 1=“Strongly disagree” to 7=“Strongly agree”. As in Study 4.1, data were analysed using univariate analyses of variance (ANOVA). Simple cell comparisons were made using contrast tests, as recommended by Rosenthal and Rosnow (1985).

Results and Discussion

A one-way analysis of variance (direction of consensus: dissent to consensus vs. collective action base-rate vs. consensus to dissent) was conducted on collective action intention scores. The condition effect was significant, $F(2, 60)=3.62, p=.03, \eta_p^2=.11$,

forming a linear trend (Figure 4.3). Exploration of the means with contrasts showed that there was significantly more action intent after watching the discussion which ended in consensus ($M=4.25$, $SD=1.45$) than the discussion which ended in dissent ($M=3.21$, $SD=1.07$) $F(1, 60)=7.24$, $p<.01$, $\eta_p^2=.11$, as predicted. The base-rate level of action intentions ($M=3.73$, $SD=1.19$) fell in the middle of the means in the two experimental conditions, and it was not significantly different to either (both p 's $>.16$).

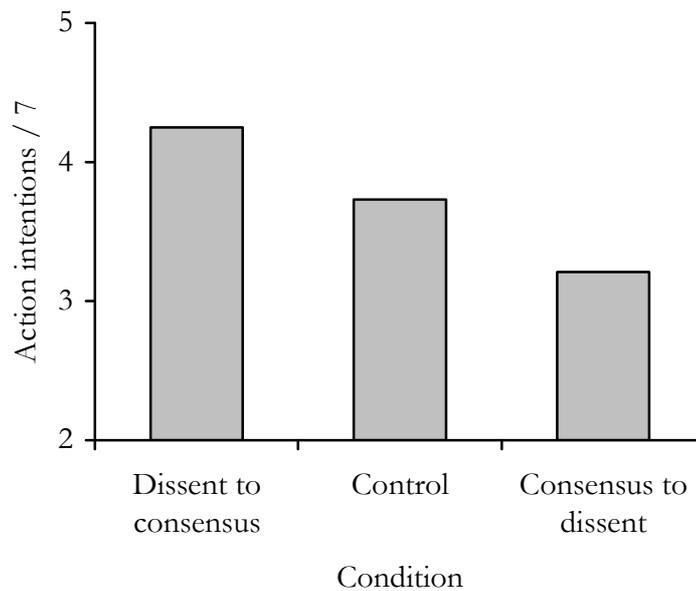


Figure 4.3. Mean collective action intention scores, Study 4.2

A one-way ANOVA comparing 2 conditions (direction of consensus: consensus to dissent vs. dissent to consensus) was conducted on scores from each of the additional standardised scales. There was a significant condition effect on the norm scale, $F(1, 39)=6.00$, $p=.02$, $\eta_p^2=.13$. Participants considered the views expressed in the discussion to be more normative when the group reached consensus ($M=4.74$, $SD=1.22$) than when the discussion ended in dissent ($M=3.79$, $SD=1.26$). The condition effect on social validation was not significant, however $F(1, 60)=.43$, $p=.52$, $\eta_p^2<.01$, (Table 4.2 for means).

There was a main effect of experimental condition on identification with the group members in the discussion, $F(1, 39)=6.04, p=.02, \eta_p^2=.13$, with significantly more social identification after watching the discussion to consensus ($M=4.21, SD=1.26$) than to dissent ($M=3.33, SD=1.10$). Therefore, as hypothesised, the process of consensualisation appeared to increase participants' sense of shared identity with the discussants.

There was more endorsement of the views expressed in the film when the discussion ended in consensus ($M=3.64, SD=1.26$), compared to when it ended in dissent ($M=2.94, SD=1.15$), although the main effect was marginal $F(1, 39)=3.44, p=.07, \eta_p^2=.08$. These results together, suggested that consensualisation informed identification *and* identity content.

Table 4.2.

Mean scores and scale intercorrelations, Study 4.2

		Consensus to dissent	Dissent to consensus	1	2	3	4	5
1. Collective action intentions	<i>M</i>	3.21 _a	4.25 _b	(.86)				
	<i>SD</i>	1.07	1.45					
2. Social validation	<i>M</i>	4.78 _a	5.02 _a	.32*	(.90)			
	<i>SD</i>	1.11	1.43					
3. Norm	<i>M</i>	3.79 _a	4.74 _b	.40**	.11	(.83)		
	<i>SD</i>	1.26	1.22					
4. Endorsement of message	<i>M</i>	2.94 _a	3.64 _b	.52**	.16	.70**	(.91)	
	<i>SD</i>	1.15	1.32					
5. Identification	<i>M</i>	3.33 _a	4.21 _b	.40**	.02	.79**	.80**	(.90)
	<i>SD</i>	1.10	1.26					

Notes. A higher score indicates a higher propensity on that measure. * $p<.05$, ** $p<.01$. Cronbach's alphas in parentheses. Means in rows with different subscripts differ at $p<.05$.

Mediation. There were positive correlations between the results of the collective action intention and norm scales $r = .40, p < .01$ and collective action intention and identification scales $r = .40, p < .01$. In order to examine whether these effects could account for the effect of condition on action intention, we conducted mediation analyses (Figure 4.4). The norm was a marginally significant covariate, $F(1, 38) = 3.84, p = .06, \eta^2 = .09$, reducing the difference between the experimental conditions from $F(1, 39) = 6.96, p = .01, \eta^2 = .15$ to $F(1, 38) = 3.34, p = .08, \eta^2 = .08$, Sobel $Z = 1.53, p = .13$, bootstrap confidence interval .01 to .90.

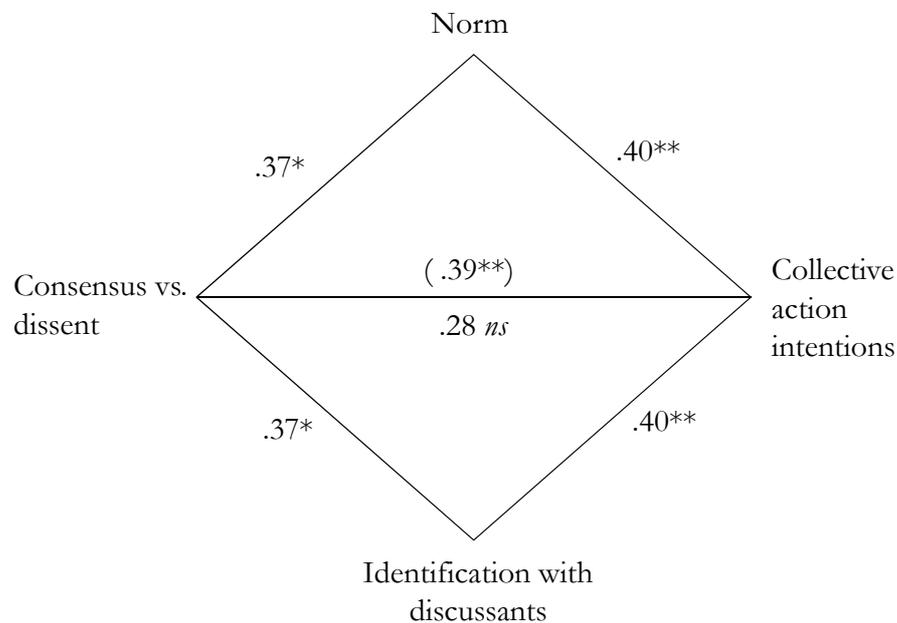


Figure 4.4. Beta coefficients for the mediation model, Study 4.2 ($p < .05, **p < .01$)

Identification was also a marginally significant covariate, $F(1, 38) = 3.91, p = .06, \eta^2 = .09$, reducing the difference between the experimental conditions to $F(1, 38) = 3.32, p = .08, \eta^2 = .08$, Sobel $Z = 1.54, p = .12$, bootstrap confidence interval .03 to .77. In view of the fact that the condition effect became non-significant in both mediation analyses, we can conclude that mediation was in evidence. If both mediators were entered into

the model at the same time, only the norm remained a significant mediator (bootstrap confidence interval for norm -.26 to .80; and identification -.20 to 1.04). However, the fact that identification becomes non-significant may be due to the relatively high ($r=.79$) between identification and the norm.

The results of Study 4.2 show that, as hypothesised, (a) more collective action intent was exhibited after witnessing the discussion to consensus than the discussion to dissent, and (b) this was mediated by the extent to which participants considered the views espoused in the discussion to be normative. Therefore, the results of Study 4.2 further suggest that it is not the content of group discussion *per se* that drives the effect of interaction, but that reaching an intra-group consensus is an important springboard for action and norm formation. Furthermore, participants identified more with the discussants if they reached consensus relative to when the ended in dissent. It appears, therefore, that a sense of shared identity with the discussants emerged (or was undermined in the dissenting condition) as well as group norms. It is interesting that identification mediated action intent in this study, but in none of the other studies in this thesis. It may be speculated that this is because the comparison condition in this study is the only one in which we specifically manipulated in-group dissent. Therefore, the difference in identification between conditions was markedly significant in this study.

General Discussion

Two studies investigated the role of intra-group consensus in normalising inter-group action from within the in-group. Study 4.1 was conducted with a non-specific out-group to investigate whether high intra-group consensus on specific actions can alone increase the perceived legitimacy of collective action. Study 4.2 manipulated the process of consensualisation by showing a film of a discussion which began with intra-

group dissent and ended in consensus, or vice versa. The purpose of these studies was fourfold: to investigate (1) whether *consensual* discussion of future courses of action promotes action intentions, (2) whether consensual discussion of even extreme courses of action can promote action intentions, (3) whether the legitimisation of inter-group action requires a pre-defined inter-group context, and finally, (4) whether the perception of in-group consensus is necessary in informing identity content or not.

First, (1) it was hypothesised that the consensual discussion of a course of action would promote action intentions that are as strong, or stronger, than when the group had merely discussed an out-group stereotype. Indeed, Study 4.1 demonstrated that when there was high perceived in-group consensus over actions, discussing them was more likely to increase action intentions than discussing the stereotype alone.

Next, it was hypothesised that, (2) collective action intentions would be strongest when the consensual recommendations for action were mainstream rather than extreme. In fact, there was little difference between action intentions when the actions espoused in the discussion were mainstream or extreme. Instead, and in line with the subsequent predictions, action intentions varied significantly with how much the action was perceived to be legitimate, and how much participants felt validated on this dimension by the in-group. Third, (3) it was hypothesised that intra-group interaction would increase action intent even in the absence of a pre-defined inter-group context. Indeed, the results of Study 4.1 suggested that pre-existing inter-group tensions were not necessary for the development of norms for action. Finally, (4) it was hypothesised that only when consensus was reached would the discussion increase action intention and norm formation. The results of Study 4.2 are consistent with this hypothesis, showing that when the persuasive argumentation is kept constant, consensus encourages action while dissent discourages it in comparison.

The present research is consistent with the theoretical assertions made in the literature that the achievement of in-group consensus can validate social reality (Festinger, 1950; Turner, 1991) and provide a psychological foundation for action (e. g., Haslam, Turner, Oakes, McGarty, et al., 1998). However, the research presented in this chapter also makes a unique contribution to this literature, in that it manipulated the *process* of consensualisation. This is important, as it demonstrates that the perception of reaching in-group consensus is not merely a product of shared identity but a *process* and catalyst for identity (in)formation, i.e., it is impacted by both increasing identification and perceptions of the (valid and legitimate) in-group norm for action.

Notably, it appeared that the discussion of stereotypes was a major step toward increasing action intentions. Consistent with Chapter 3, the additional discussion of group norms for action added little to this. As the level of consensus was controlled in the studies presented in this chapter (unlike those in Chapter 3), we conclude that this result was found because stereotypes are powerful as *socially shared* carriers of implicit norms and ideologies (e. g., Haslam, 1997; Jost & Banaji, 1994; Reicher, Hopkins, & Condor, 1997; Reynolds, Oakes, Haslam, Nolan, & Dolnik, 2000; Tajfel, 1981; see Chapter 1).

The common way of theorising about the processes behind social action is to look for external antecedents and moderating inter-group factors. However, the present research demonstrates that there may be internal (in-group) factors which have a strong influence on why and how these factors become a *shared, validated* and *legitimised* basis for collective action.

Implications and Future Directions

The present results show that social action intention can be increased by collectively consensualising on an action plan and/or a stereotype. The results diverge

into two directions: first, achieving consensus is crucial for identity (in)formation, and can have powerful repercussions for inter-group action. This supports the argument forwarded in the inductive model of identity formation (IMIF; Postmes, Haslam et al., 2005; Postmes, Spears et al., 2005; see Chapter 1) that social identities are socially shared and *socially constructed* entities. It also supports Haslam and colleagues' research on stereotype consensualisation. This research has been used to support the assumption that when there is a perception of consensus on social stereotypes, they are transformed into *subjectively valid* beliefs and can be used as a basis for social behaviour (e.g., Haslam, Turner, Oakes, McGarty, & Reynolds, 1998). However, importantly we also show the opposite: that undermining the development (or illusion) of consensus can also have powerful implications: the explicit lack of consensus prevents inter-group action from occurring.

The results of Studies 4.1 and 4.2 also speak to the power of intra-group interaction to legitimise collective action within an in-group not already premised upon a shared opinion or stereotype (cf. Bliuc, McGarty, Reynolds, & Muntele, 2007; Thomas & McGarty, in press). The implications of findings such as these are that interaction not only amplifies existing psychological factors (such as identification or attitudes), but also aids the *formation* of shared identity content.

There is potential, therefore, for the ideas articulated in the current research to be utilised to develop applied positive interventions. The findings presented in this chapter imply that interaction not only amplifies existing identity salience or attitudes, but aids their (in)formation. If indeed, social interaction processes have this effect because in-group members can validate their reality, social consensus could therefore provide a solid basis to *change* norms (for the better) by a simple re-framing of the discussion topic. Furthermore, the undermining of consensus through nuanced debate

could be utilised to prevent social action. It is therefore important that future research investigates further the process through which intra-group interaction can achieve identity content change.

One potential limitation of these studies is that they show that consensualisation produces powerful effects even if participants themselves are not actively engaged in social interaction. This raises the question of whether group interaction is a necessary condition for social validation. Previous research (Chapters 2 and 3) has shown that group discussion serves to establish social validation and consequent confidence in one's beliefs regarding out-groups and (implicit) appropriate actions. Here, we show that providing the participants identify with the discussants and see themselves as potential in-group members, the mere observation of a discussion can have a similar effect. This raises the question of whether group discussion is but one means of elevating confidence or providing validation for one's beliefs on a variety of topics (cf. Petty, Briñol & Tormala, 2002; see also Chapter 3). The issue of whether group interaction is a particularly effective (or superior) means of manipulating belief confidence may be considered in future research.

The research has not been designed to show that active participation is necessary, and therefore does not test this prediction. Even though participants were not themselves in the discussion they witnessed, psychologically they were clearly involved (not least because the instructions invented them to be). This is also evident from their endorsement of the message, and their identification with the actors they viewed. It is also clear that in the absence of such identification, action intentions are formed much less strongly. All of this provides some evidence for the social underpinnings of these effects.

The elementary nature of interactions with others does not stem, therefore,

from the active participation in discussions, or from the ability to discuss things with other in-groupers *face-to-face*. Indeed, mass communication alone should be able to produce many of the same effects as those observed here. What is critical, we suggest, is that people have the opportunity to gather information about what others like them believe and feel about the issues that concern them. This is the only way to develop a valid image of the social world, after all. In this sense, interaction with others is a *logical* necessity. The present research shows that interaction is the process variable which explains why people may resort to hostile or violent collective action. That we may design interaction out of our experiments by substituting direct interaction with a different form of communication (e.g., with the experimenter themselves) does not negate the underlying premise that the sharing of cognitions with others through interaction is an essential prerequisite to them being enacted socially.

Conclusion

In this research, the consensual discussion of a course of action promoted action intentions which were stronger than when the in-group had discussed the out-group stereotype only. This effect did not differ with whether the consensual recommendations for action were mainstream or extreme. Rather, the crucial factor was the extent to which this action was perceived to be legitimate, and how much participants felt action was validated by the group discussion. These results were demonstrated in the absence of a pre-defined inter-group context, suggesting that salience of the out-group is not a necessary pre-condition for social action or norm formation. Instead, these results showed that the perception of in-group consensus was a necessary condition for the formation of norms for social action.

In conclusion, the studies reported here show that if people perceive that the in-group has gone through a process of consensualisation in conversation, they are able to

infer norms for social action. This does not require the pre-existence of inter-group conflict or even a salient out-group. Social consensus, therefore, can provide a powerful springboard for social action. Furthermore, in-group consensualisation appears to offer great potential not only to crystallise in-group norms for social behaviour, but also to *change* them: a proposal investigated in the final empirical chapter.

CHAPTER 5

Changing Social Identities: Social Consensus as a Foundation of Stereotype Threat

“Stereotype threat” is the phenomenon that negative stereotypes can adversely affect the psychology and behaviour of members of the stereotyped group within a stereotype-relevant domain (e.g., Steele, 1997; Steele & Aronson, 1995). Stereotyping is assumed to operate (at least in part) as an automatic, implicit process (Devine, 1989). Therefore, where social stigmas are pervasive, stereotype threat can be a substantial problem for stigmatised group members. This chapter shows that a brief opportunity for intra-group interaction offers a way to *eliminate* the stereotype threat effect.

The Negative Effects of Stereotype Threat

Many studies have shown that when an individual becomes aware of their membership in a stigmatised group, their performance will be negatively affected by this stereotype if measured in a relevant domain (e.g., Aronson et al., 1999; Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995). Factors which have been argued to mediate this effect include anxiety and evaluation apprehension (Spencer et al., 1999), but consensus on the critical mediators is yet to be reached (Jamieson & Harkins, 2007). Furthermore, there are several known variables which moderate stereotype threat, such as domain identification (Aronson et al., 1999; Steele, 1997), public visibility of performance in the stereotyped domain, and salience of/identification with the negatively stereotyped in-group identity (e.g., Shapiro & Neuberg, 2007; for a review). Therefore, it is now acknowledged that there are various types of, sources of, and responses to stereotype threat (Shapiro & Neuberg, 2007), and along with each of

these is a selection of bespoke moderators and mediators.

Stereotype lift. The consequences of stereotype threat are not restricted to members of stigmatised groups. Evidence has emerged that the negative stereotyping of an out-group can be utilised to *enhance* in-group performance (*stereotype lift*; Walton & Cohen, 2003; Marx & Stapel, 2006). There are benefits to performance, therefore (as well as self-esteem), associated with belonging to a group with a positive self-stereotype.

Combating Stereotype Threat

Evidence suggests that it is possible to overcome the “deficiencies” in performance associated with stereotype threat through self-confirmation (Martens, Johns, Greenberg, & Schimel, 2006). Individuals may engage in cognitive self-esteem enhancing strategies to buffer the stigma attached to their group. However, an individual-level intervention would not change the negative content of the in-group stereotype *per se*, it merely allows one to re-categorise at the individual level, and bypass the group-level threat for the moment. A longer-term strategy would be to directly invalidate or challenge the negative in-group stereotype (Haslam, Salvatore, Kessler, & Reicher, 2008). However, such strategies are likely to be of limited success if undertaken alone (i.e., individually): Stereotypes are by definition *socially shared* cognitions, therefore the validity of these beliefs must be contingent on a certain degree of (implicit) consensus among members of the in-group. It follows logically that the best strategy for stereotype change is through reaching a different consensus about those stereotypes in interaction with other in-group members (i.e., collectively).

The Power of Talk: Transforming Social Identities.

Intra-group communication has been shown to enhance consensus about *out-groups* (e.g., Thompson, Judd, & Park, 2000). Reaching consensus on stereotypes through group discussion is defined here as *consensualisation*. Many studies have shown

effects of consensualisation on stereotype content (Haslam, 1997; Haslam et al., 1999; Haslam, Turner, Oakes, McGarty et al., 1998; Haslam, Turner, Oakes, Reynolds et al., 1998; Chapter 3). Further research has detailed the underlying process (Kashima, 2000; Lyons & Kashima, 2003). Recent research has suggested that consensualisation may aid the development of in-group norms for social behaviour (Postmes, Haslam, & Swaab, 2005; Swaab, Postmes, van Beest, & Spears, 2007; Chapter 2; Chapter 3). It has also been suggested that consensualisation about out-group stereotypes may partly occur because in-group members are a source of validation (“confirmation”) for individuals’ belief systems and ideologies (Correll & Park, 2005; Festinger, 1954; Chapter 3).

The parallels between out-group stereotypes and in-group norms (see also Sherif, 1935) are, to some extent, entirely logical. After all, if “we” discuss what “they” are like, then it is highly likely that we also draw inferences about in-group stereotypes (e.g., “we” are not like “them”, see Turner, 1991). These emergent in-group stereotypes are likely to go beyond the descriptive aggregation of in-group views, in the sense that they are accompanied by inferences about what is “appropriate”, and therefore normative, social behaviour. Thereby, discussions about the out-group stereotype are not merely informative about the out-group stereotype, but they implicitly shape the content of the *in-group* identity and (importantly) the in-group’s self-stereotype (see also Postmes, Spears, et al., 2005). Importantly, however, research on intra-group discussion, communication and consensualisation has thus far largely ignored their implications for self-stereotyping.

One exception to this is the collective action literature, in which recent attempts have been made to chart the implications of intra-group interaction for collective

action. In particular, research has shown that discussing the out-group stereotype may motivate the in-group to *mobilise* (Stott & Drury, 2004; see also Chapter 2; Chapter 3).

The Present Research

The current studies seek to integrate these literatures. One purpose of the present research is to demonstrate that consensualisation through communication in small groups does not merely affect out-group stereotypes, but can also be used to directly change in-group stereotypes. A key element of this is the demonstration that consensualisation in a small group interaction can *discredit and overcome* a negative in-group stereotype, through the endorsement of an alternative (positive) in-group stereotype. To the extent that in-group stereotypes can indeed be redefined in this way, intra-group interaction should also have the capacity to overcome the stereotype threat effect. After all, if the self-stereotype becomes more positive, its threatening impact to the self should be neutralised.

The present research thus examines (a) the potential for intra-group discussion to *change* existing self-stereotypes (i.e., to confirm or to challenge them), and (b) the implications this has for the stereotype threat effect. Two studies were conducted to investigate whether challenging a negative in-group stereotype through intra-group discussion could enable in-group members to resist the effects of stereotype threat on performance. If the perception and consequences of stigma are indeed malleable, then the practical and theoretical implications of these methods are considerable. They offer a potential way for marginalised and oppressed groups to overcome the psychological restraints of their position.

STUDY 5.1

The first study sought to establish the effects of challenging vs. confirming a negative in-group stereotype through individual thought, or through group discussion.

Performance on a stereotype-relevant task was then examined.

Hypotheses. It was hypothesised that (a) group discussion would qualitatively transform the consensual content of the in-group and out-group stereotype. Moreover, it was predicted that (b) participants who confirmed the negative stereotype would have lower maths scores than participants who challenged the stereotype, and (c) as it relies upon in-group consensus, this effect would only be achieved through group discussion, and not through individual cognition.

Method

Participants and Design

Participants were 77 female first year Psychology undergraduates from Exeter University, recruited during a practical class. The mean age was 20.04 years old. All participants were unpaid and naive as to the purposes of the study. Maths ability was controlled by only including participants who did not study maths at a degree level. The design was a 2 (Topic: Confirm stereotype vs. challenge stereotype) x 2 (Reflection: Group discussion vs. Individual cognition) between-subjects factorial design. Participants were randomly allocated to the conditions, and within those to 3-person groups ($N=26$).

Materials and Procedure

Consenting participants were informed that they were taking part in a study on mathematical ability, which involved a modified brainstorming task, a maths test and answering a few questions about their views on numerical ability. The experimenter then split participants into groups. Participants were informed that first of all, they would be taking part in an exercise to help them generate their opinions. They were then asked to discuss in their groups or think about alone, either “Why it is true that men are better than women at maths”, thus confirming the stereotype; or “Why it is

not true men are better than women at maths”, thus challenging the stereotype. To ensure that participants in all conditions worked to a similar goal, they were all asked to write down their top 5 ideas. After 10 minutes, participants were asked to stop discussing or thinking about this. They then individually completed a maths test, which began with the following feedback: “There is a stereotype in society that women are worse at maths than men. Gender differences have been found in the past on this maths test. This test is *diagnostic of your numerical ability*, so please do your best to answer the following questions”. The test contained 7 questions of approximately GCSE standard, which was to be completed in silence, individually and without the use of a calculator. Although there was no time limit on completing the test, the amount of time taken was recorded in order that its effect could be controlled. Participants then filled out a questionnaire comprised of standardised scales, and were debriefed.

Dependent Measures

The performance measure was the total score on maths test, which contained 7 items: “ $3/5 - 1/4 = ?$ ”; “If $b = -2$ and $c = 6.1$, what is the value of $2b^2 + 3c^2$?”; “Make a the subject of this formula: $r = \sqrt{4a + b}$ ”; “Write down the gradient and the intercept of the straight line graph: $2y = 4x + 8$ ”; “ $2/3 + 2/4 = ?$ ”; “A formula is given by $y = mx + c$. Find the value of x when $y = 13$, $m = 5$ and $c = 3$ ”; and “What is 20% of £25?”. Answers were marked either correct (1) or incorrect (0), and correct answers were totalled to give a score out of 7.

Psychological outcomes were measured via standardised scales (Table 5.3 for intercorrelations). A 4-item measure of evaluation apprehension ($\alpha = .93$) was adapted from Spencer, Steele, and Quinn (1999), who found it to be a significant mediator of stereotype threat: “If I do poorly on the maths test, people will look down on me”; “People will think I have less ability if I do not do well on the maths test”; “If I don't

do well on the maths test, others may question my ability”; and “People will look down on me if I do not perform well on the maths test”. Anger has been shown to impact upon social behaviour in past research (e.g., Van Zomeren, Spears, Leach, & Fischer, 2004). Therefore, in order to assess the psychological impact of the challenge conditions, a 4-item anger scale was also included ($\alpha=.90$), “I would tend to feel angry about the stereotype about gender differences in ability”; “I feel irritated about the stereotype about gender differences in ability”; and “The problems that the stereotype about gender differences in ability brings makes me furious”.

Furthermore, in order to measure the impact of the manipulations on gender identification, a 17-item scale was included. This scale (Ryan, Iyer, Hersby, & Kulich, in prep, adapted from Cameron, 2004) measured three factors of gender identification - in-group ties, affect, and centrality. Within this overall gender identification scale ($\alpha=.86$), 6-items measured in-group ties ($\alpha=.87$), for example: “I feel strong ties to other women”, “I have a lot in common with other women”, and “In a group of women, I really feel that I belong”; 5-items measured affect ($\alpha=.86$), for example: “Generally, I feel good when I think about myself as a woman”, “I often regret that I am a woman” (reverse-coded), and “Just thinking about the fact that I am a woman sometimes gives me bad feelings” (reverse-coded); and 6-items measured centrality ($\alpha=.77$), for example: “Being a woman has very little to do with how I feel about myself” (reverse-coded), “In general, being a woman is an important part of my self-image”, and “In my everyday life, I often think about what it means to be a woman”. Participants indicated their degree of agreement with each statement (1=“strongly disagree,” to 7=“strongly agree”).

Analytic Strategy

Qualitative Data

A content analysis was performed on the written description of the in-group and out-group stereotype provided by participants across conditions, following iterative inductive and deductive coding procedures recommended by Miles and Huberman (1994). Please refer to Appendices 5.1 and 5.2 for definitions of the codes. The purpose of this analysis was to explore the content of the stereotype formed through group discussion or individual reflection for any descriptive and/or affective differences. Inter-rater agreement was 99% for the confirm condition and 95% for the challenge condition, and reliability was good, Cohen's (1960) $\kappa=.68$ to $\kappa=1.00$.

Measure of consensus. The theory presented in this chapter hinges on the definition of stereotypes as *socially shared* cognitions, i.e., as based on consensus among in-group members. Therefore, for the confirm/challenge conditions, scores were calculated to represent the amount of consensus surrounding the content of the stereotypes. We computed the average proportion that category codes were mentioned (Tables 5.1 and 5.2). Then, we computed how frequently each group of $N=3$ participants mentioned each code (where 1=all mentioned, 0=none mentioned). Next, for each group, we computed a relative score of how frequently they mentioned each code, minus the average proportion that the categories were mentioned in the condition. We then normalised the scores for each group, so that scores range from zero to one (where 1=complete consensus that a code either applied or did not apply)¹¹.

The scores for each group were then aggregated across codes, to create one number per group, ranging from zero (no deviation from average proportions) to one (consistent consensus). In order to compare consensus on the stereotype across

¹¹ In order for a score of "1" to represent consensus, if the category was mentioned by all members of a group, scores were divided by 1-condition average. If the category was *not* mentioned by all members of the group, scores were divided by the condition average itself.

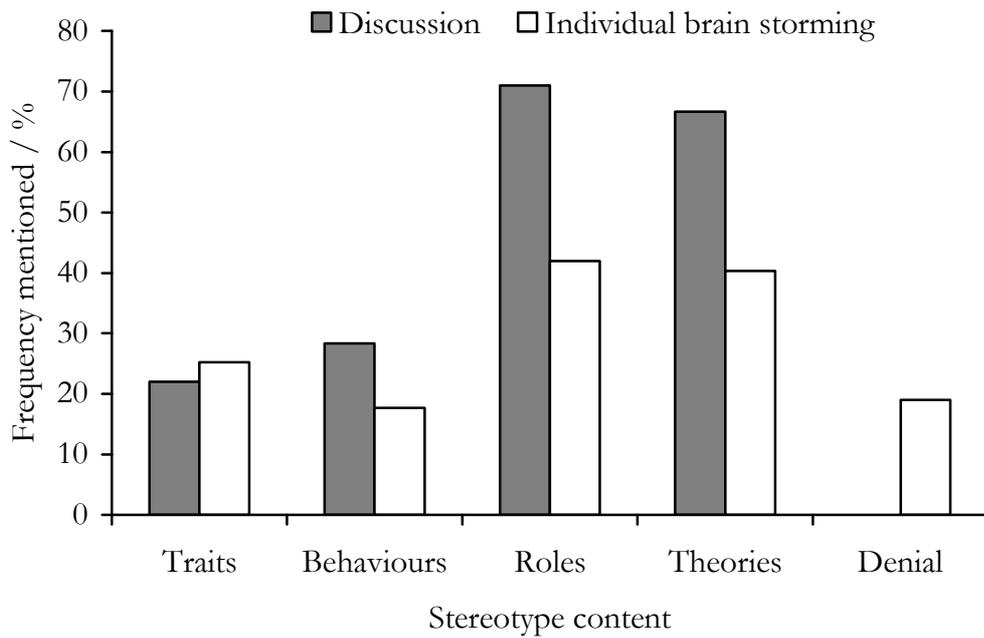
conditions, these consensus scores were analyzed via analysis of variance (ANOVA) at the group level.

Results

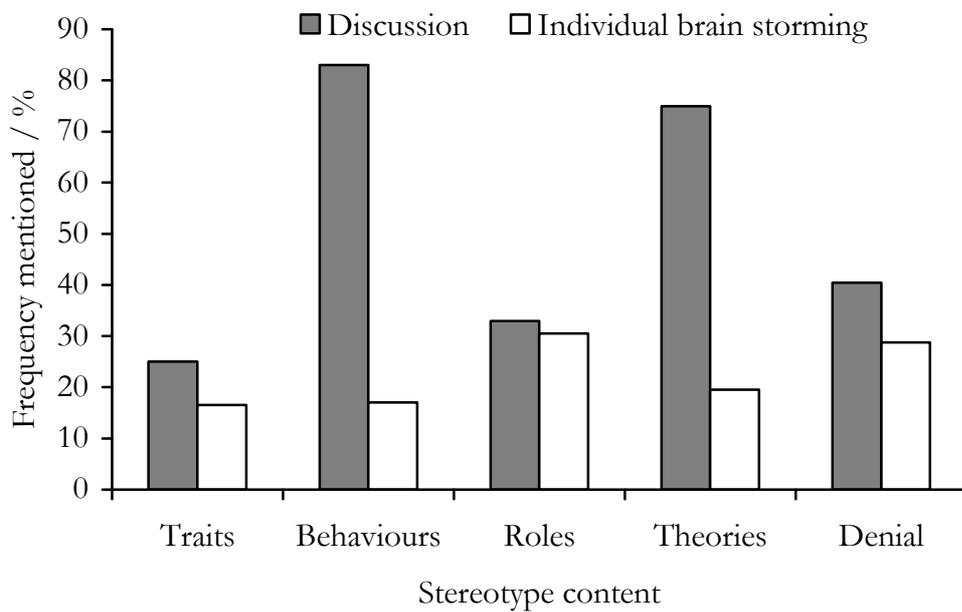
Stereotype Consensus

A content analysis was performed on the written descriptions of the gender stereotypes. A 2 (Topic: Confirm stereotype vs. challenge stereotype) x 2 (Reflection: Group discussion vs. individual thought) between-subjects ANOVA was conducted on the group-level consensus scores generated from the content analysis, revealing considerable effects of reflection condition. Analyses revealed a main effect for reflection condition: there was greater consensus after discussion ($M=.91, SD=.10$) compared to individual thought ($M=.77, SD=.04$), $F(1,21)=24.05, p<.01, \eta^2=.53$. There was also significantly more consensus when participants confirmed the stereotype ($M=.87, SD=.10$) than challenged it ($M=.81, SD=.10$) $F(1,21)=4.33, p=.05, \eta^2=.17$. There was no difference in consensus when confirming or challenging the stereotype through individual thought $F(1,21)=.33, p=.57, \eta^2=.02$. However, there was significantly more consensus when confirming the stereotype through discussion than challenging it through discussion $F(1,21)=5.78, p=.03, \eta^2=.22$.

In the confirm stereotype condition, participants listed an average of 4.65 ideas to answer the question, “Why is it *true* that men are better than women at maths?” There were 17 distinct ideas in total (see Table 5.1 and Appendix 5.1). These were collapsed into 5 categories: traits, behaviours, roles, theories and denial (Figure 5.1). Most strikingly, after group discussion (compared with individual reflection) there was a greater degree of consensus surrounding *theories* confirming the purported gender difference, and gender roles (Figure 5.1). Simple cell analyses of the consensus score revealed significantly more consensus about the stereotype after confirming it through



A: Why are men better at maths?



B: Why are men *not* better at maths?

Figure 5.1. Stereotype content used to explain why men are better (A) or not better (B) than women at maths, Study 5.1.

discussion ($M=.95$, $SD=.05$) compared to individual thought ($M=.78$, $SD=.05$)

$F(1,21)=19.71$, $p<.01$, $\eta^2=.48$.

A large proportion of participants in the group discussion condition theorised that the “gender gap” is caused by physical differences (86%) and/or because of social reasons (62%). 43% also believed that it could have been caused by evolution, but only 16% in the individual condition mentioned this. Thus, the stereotype appeared to be more predetermined, irrefutable and concrete in this condition. Furthermore, only in the individual condition did participants on occasion deny the stereotype (19%). Therefore, socially confirming the stereotype appeared to disable participants’ willingness to contest it.

Responses by participants who reflected on the question alone were more varied and less consensual. Overall, participants in the individual reflection condition focused more broadly on both genders than participants in the group discussion condition (Table 5.1). Participants in the individual condition appeared to focus slightly more on stereotypical traits, and less on theories and gender roles, than participants in the discussion condition.

For the question, “Why is it not true that men are better than women at maths?” participants listed 4.14 ideas on average. There were 14 distinct ideas mentioned in total (Table 5.2 and Figure 5.2 for a full summary of the data). Again, there were significant effects of reflection condition. There was a greater degree of consensus in each category of codes (traits, behaviours, roles, theories, denial) after group discussion ($M=.86$, $SD=.12$) than individual reflection ($M=.76$, $SD=.02$), $F(1,21)=6.41$, $p=.02$, $\eta^2=.23$, (Figure 5.2).

Table 5.1

Percentage of participants who mentioned each code in response to the question, “Why is it true that men are better than women at maths?”

Focus	Code	Group Discussion	Individual thought
1. Traits	SPAT	38	32
	SIN	10	0
	SAT	0	16
	LOG	67	63
	NUM	10	47
	BET	5	0
	CRE	5	26
	EMO	0	11
	VERB	10	32
2. Behaviours	OTH	10	5
	ED	48	37
	COMP	10	11
3. Roles	PROF	62	42
4. Theories	SOC	62	58
	BRA	86	47
	EVO	43	16
5. Denial	EQ	0	19

Note. Percentages were computed at the individual level. Codes are described in Appendix 5.1.

The most prominent difference between the conditions surrounded behaviour and theories for the purported gender difference. Group discussion appeared to focus participants’ attention on positive in-group characteristics – reasons why women are *better* than men (i.e., not equal). 72% of women in the group discussion condition mentioned the behaviour that girls do better in exams/at school than boys, compared to only 17% in the individual condition. 83% of participants in the group discussion condition also theorised that any purported gender difference would be due to prejudices in society and that in fact men and women are born equal (compared to only 35% in the individual condition). 33% argued that environmental factors (not gender) determine numerical skill. In other words, participants in this condition were better

able to construct a positive identity, with concrete theories and evidence as to why this is the case.

Responses by participants in the individual reflection condition were less consensual and more varied. Overall, participants in the individual condition appeared to focus more on equality and individualisation, and less about why the in-group is superior to the out-group; relative to participants in the group discussion condition (Table 5.2).

Table 5.2

Percentage of participants who mentioned each code in response to the question, "Why is it not true that men are better than women at maths"

Focus	Code	Group Discussion	Individual thought
1. Traits	ATT	6	11
	INT	17	22
2. Behaviours	EX	72	17
3. Roles	FEM	22	28
	ROL	28	33
4. Theories	SOC	83	35
	NUR	33	6
5. Denial	EQ	50	56
	DED	33	33
	DBRA	22	17
	DEVO	11	11
	RST	33	50
	IND	17	17
	GEN	22	17

Note. Percentages were computed at the individual level. Codes are described in Appendix 5.2.

Performance Effects

Hierarchical linear modelling (HLM) analyses were performed in HLM 6.03 (Raudenbush & Bryk, 2002). Main effects and the 2-way interaction were tested by means of dummy variables (Judd, McClelland, & Culhane, 1995; Rosenthal & Rosnow,

1985; Wilkinson et al., 1999), and contrast variables were created to compare specific cells, following recommendations for contrast analysis in the testing of regression models by Cohen, Cohen, West, and Aiken (2003, pp. 332-341).

Maths test score. In the initial model, maths scores were entered at level one and the dummy variables were entered at level two. This initial model was significantly different to the null model, $\chi^2(22)=37.10, p=.02$, indicating significant between-condition differences overall. Analyses using the dummy variables indicated that there were significantly higher maths scores after the stereotype was challenged ($M=4.25, SD=1.71$), compared to when it was confirmed ($M=3.31, SD=1.61$), $\gamma=-.28, p=.03$, confirming hypothesis (a). There was no overall difference between group discussion ($M= 3.61, SD=1.87$) and individual reflection ($M=3.92, SD=1.55$) $\gamma=-.08, p=.53$. However, there was a significant 2-way interaction, $\gamma=-.56, p=.01$ (Table 5.3, Figure 5.2).

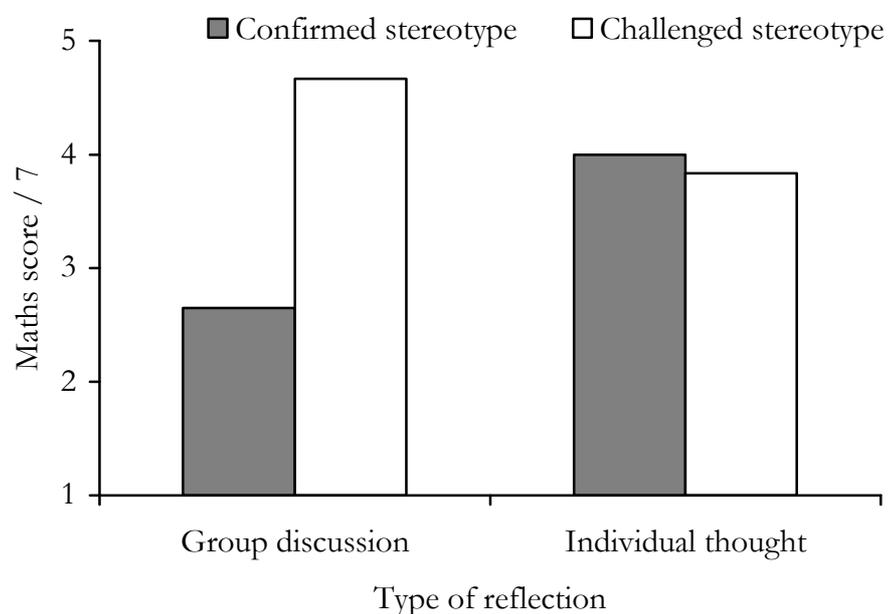


Figure 5.2. Mean maths test scores, Study 5.1

Table 5.3

Means, standard deviations, scale reliabilities, and intercorrelations of dependent variables, Study 5.1

Reflection condition	Topic	Group discussion		Individual thought		1	2	3	4	
		Confirm stereotype	Challenge stereotype	Confirm stereotype	Challenge stereotype					
1.	Consensus score	<i>M</i>	.95 _a	.86 _b	.78 _b	.76 _b				
		<i>SD</i>	.05	.12	.05	.02				
2.	Maths test score	<i>M</i>	2.65 _b	4.67 _a	4.00 _a	3.83 _{ab}				
		<i>SD</i>	1.50	1.68	1.45	1.69				
3.	Evaluation apprehension	<i>M</i>	3.05 _a	3.24 _{ab}	2.14 _b	2.85 _{ab}	-.14	(.93)		
		<i>SD</i>	1.58	1.88	1.20	1.35				
4.	Anger	<i>M</i>	3.12 _a	4.00 _b	3.08 _a	2.95 _a	.06	.27*	(.90)	
		<i>SD</i>	1.54	1.41	1.79	1.44				
4.	Gender identification	<i>M</i>	4.97 _a	4.93 _a	5.22 _a	4.97 _a	.18	-.27*	-.05	(.86)
		<i>SD</i>	1.17	1.17	.62	.68				

Notes. * $p < .05$. Cronbach's alphas in parentheses. Consensus scores were calculated at the group level, all other scores calculated at the individual level. Means in rows with different subscripts differ at $p < .05$.

In order to explore this interaction, between-cell comparisons were conducted using contrast variables. Overall, participants who confirmed the stereotype through group discussion ($M=2.65$, $SD=1.50$) performed significantly worse than participants who confirmed the stereotype individually ($M=4.00$, $SD=1.45$), $\gamma=-.28$, $p=.02$. When participants were asked to *challenge* the stereotype, participants who engaged in a group discussion ($M=4.67$, $SD=1.68$) performed somewhat better than participants who reflected on the stereotype alone ($M=3.83$, $SD=1.69$), but this was not a significant result $\gamma=.16$, $p=.19$. Among those who reflected individually, it made no difference whether they confirmed or challenged the stereotype, $\gamma=.04$, $p=.82$. Among those who discussed the stereotype in a small group, however, the difference between those who confirmed or challenged it was highly significant, $\gamma=-.42$, $p<.01$.

Psychological Effects

Evaluation apprehension. There was no overall main effect for topic $\gamma=-.15$, $p=.21$, and a marginal main effect for reflection condition, $\gamma=.15$, $p=.07$. The 2-way interaction was not significant either, $\gamma=.15$, $p=.48$. However, analyses using contrast variables showed that there were significantly higher evaluation apprehension scores for participants who confirmed the stereotype by discussion ($M=3.05$, $SD=1.58$) compared to individual reflection ($M=2.14$, $SD=1.20$) $\gamma=.22$, $p=.02$. Therefore, when confirming the stereotype, discussion appeared to increase evaluation apprehension.

Anger. There was no overall main effect for topic $\gamma=-.12$, $p=.19$, but the main effect for medium was highly significant: there was more anger when participants discussed the stereotype in the group ($M=3.55$, $SD=1.53$) than when they reflected on it alone ($M=3.02$, $SD=1.61$) $\gamma=.34$, $p<.01$. The 2-way interaction was marginally significant $\gamma=-.29$, $p=.08$ (please refer to Table 5.3 for means). Contrast analyses showed significantly more anger was experienced when challenging the stereotype

through discussion ($M=4.00$, $SD=1.40$) than when confirming it ($M=3.12$, $SD=1.54$) $\gamma=-.20$, $p=.03$. When the stereotype was challenged, group discussion generated significantly more anger ($M=4.00$, $SD=1.41$) than individual reflection did ($M=2.95$, $SD=1.44$) $\gamma=.23$, $p<.01$. Therefore, it was particularly when challenging the stereotype that group discussion somewhat increased anger.

Gender identification. There were no main effects or interactions on the overall gender identification scale or on the subscales (all p 's $>.08$). Therefore, changes in gender identification, including centrality, affect and in-group ties; could not explain the differences between conditions.

Mediation. Because there were some differences in evaluation apprehension and anger across conditions, they were each entered alone as grand centred level 1 predictors, with maths score as the outcome variable. The dummy and contrast variables were entered separately at level 2. The slope for evaluation apprehension was not significant $\gamma=-.13$, $p=.22$. Although on entering this mediator into the model, the difference between participants who confirmed the stereotype through discussion or individual reflection became non-significant $\gamma=-.26$, $p=.08$, the Sobel test suggests that conditions for statistical mediation were not met, $Z=1.47$, $p=.14$. Anger was also not a significant predictor itself, $\gamma=.07$, $p=.54$, and therefore it was not surprising that here too the Sobel test failed to reach significance, $Z=.61$, $p=.55$. Thus, there was also no evidence that anger mediated the effect of condition on maths score.

Discussion

The purpose of this study was to establish the effects of group discussion (vs. individual thought) about negative self-stereotypes on performance in a stereotype-related domain. It was hypothesised that participants who confirmed the negative stereotype that women have lower maths ability than men would subsequently score

lower on the maths test than participants who challenged the stereotype. Indeed, affirmation of the stereotype was associated with poorer performance. It was also hypothesised that this effect would occur after group discussion of the stereotype but not after individual thought, as it was argued that in-group consensus was required to drive the negative performance effects. Indeed the difference between confirming and challenging the stereotype was only significant among participants who had had a group discussion about it, not among participants who reflected on it individually. Analysis of the content of the stereotypes revealed that there was indeed greater consensus after discussion, compared to individual thought.

Evaluation apprehension and anger were included as potential mediators in the model, but neither of them convincingly demonstrated that these processes were in operation. Participants experienced more evaluation apprehension when confirming the stereotype by discussion compared to individual thought. On the other hand, when challenging the stereotype through discussion, anger appeared to be the more important factor. There was more anger when participants discussed the stereotype than when they reflected on it alone. Significantly more anger was experienced when challenging the stereotype during discussion than when confirming it. When the stereotype was challenged, group discussion generated significantly more anger than individual brain storming. There was no conclusive evidence of a single mediating process, however. There were also no differences in gender identification across conditions. Therefore, from these results the question remains, what was driving the transformative effect of interaction on performance?

Transforming Social Identities

Clues as to the processes driving the performance effects were found in the qualitative data. Participants' explanations as to *why* women were worse than men at

maths (or not) were saturated with stereotyped gender statements (cf. McGarty, Yzerbyt, & Spears, 2002). Notably, there was far more consensus on these statements after discussion compared to individual thought. Furthermore, as hypothesised, group discussion qualitatively transformed the content of in-group stereotypes. When asked to confirm the stereotype, participants in the group discussion condition readily confirmed the *positive* qualities of the *out-group* relative to women. Participants who discussed the issue were also more likely to provide (essentialising) physiological theories that would explain the negative performance of their group. This contrasts with participants in the individual thought condition, who were more likely to proffer (more malleable) social explanations. In this way, group discussion *consensually* accentuated and entrenched participants' negative views of the in-group and the commensurate positive views of the out-group. Therefore, discussion appeared to restrict participants to the mould of the negative in-group and positive out-group stereotypes when they were not given the chance to challenge them. This phenomenon resonated with participants' much lower performance in this condition and the slightly increased evaluation apprehension.

However, group discussion also bolstered successful resistance of the negative stereotype. In the condition where the stereotype was challenged, group discussion focused participants' attention on positive in-group characteristics and malleable factors which determine mathematical skills, such as social forces; rather than constraining ones, for example physical differences. In other words, participants in this condition successfully invalidated the stereotype and replaced the stigmatised identity content with a (more positive) alternative identity. In contrast, participants in the individual reflection condition appeared unable to reach consensus on theories or content which meaningfully and usefully invalidated the stereotype.

It may be speculated that reaching consensus on the in-group stereotype through group discussion, or *consensualisation* (in this case by guiding the direction of this discussion towards specific answers), affected individual performance on a subsequent unrelated maths task. Although these effects were encouraging, the evaluation of the direction and strength of this effect was hampered by, (a) the absence of a control condition in which participants did not explicitly reflect on stereotypes (i.e., the conditions under which standard stereotype threat effects occur, and therefore one other appropriate "baseline" against which to evaluate effects of group discussion), and (b) the absence of conditions with male participants who undergo the same procedure (i.e., an alternative comparison standard against which to evaluate relative performance of female participants). Therefore, the prime purpose of Study 5.2 was to replicate effects of Study 5.1 and to add these alternative control and comparison conditions.

STUDY 5.2

The second study compared the effects of intra-group consensualisation on women's and men's maths test performance. It was hypothesised that gender and discussion topic would interact: Women who confirmed the stereotype during a group discussion would show decreased performance in line with the stereotype threat effect (vs. challenged and vs. control condition). In contrast, men who confirmed the stereotype during group discussion would show improved performance (stereotype lift) relative to baseline scores. A further prediction was that when men and women were given the opportunity to challenge the stereotype through discussion, the stereotype threat effect should be eliminated, with no difference on maths performance between the gender groups.

Method

Participants and Design

Participants were 255 undergraduate volunteers (134 male, 121 female, mean age=19.7). In order to control for maths ability, all participants had achieved GCSE maths grade A*-C, but had not studied maths at a higher level. Participants were randomly assigned to one of the conditions of the 2 (Discussion topic: confirm stereotype vs. challenge stereotype) x 2 (Gender: male vs. female) between-subjects factorial design or to the no-discussion baseline condition (N participants=60). Participants were randomly allocated to same-gender 3-person discussion groups (N groups=80).

Materials and Procedure

The procedure in the experimental conditions was identical to Study 5.1, except that participants did not write down their responses, and therefore there was no qualitative data to analyse. In the baseline condition, participants were not given any feedback on the stereotype or told that the test was diagnostic, were merely asked to do their best on the test, and had no opportunity to discuss the study at any point. Therefore, we did not expect a stereotype threat effect across gender on the baseline scores.

Dependent Measures

The primary dependent measure was the total score on maths test, which was expanded from Study 5.1 to contain 11 items, to increase the sensitivity of the test. The additional items were, “ $-9+(-7)=?$ ”; “ $3/5-1/4=?$ ”; “A student got 90 out of 270 in a maths test. What percentage is this?”; and “Simplify $18^6 \div 18^2$ ”. Correct answers were totalled to give a score out of 11.

The scales used were identical to Study 5.1, with the addition of a 6-item social validation scale ($\alpha=.86$), designed to measure the extent to which participants felt that the discussion legitimised and justified their opinion on the stereotype: “I feel that my

opinions are valid”; “I feel that my views are well-founded”; “My beliefs are justified”; “I am certain that my views are right”; “I feel that my opinions are shared by many”; and “I feel that I have a justified opinion”. This scale was included as an additional potential process measure, based on previous findings on the validating aspects of consensualisation (Chapters 3 and 4).

In order to expand upon the evaluation apprehension measure used in Study 5.1, a standardised scale was also included to measure levels of anxiety experienced over the maths test (4 items, $\alpha=.81$): “Whilst doing the maths test, I felt anxious”; “It worried me when I found a question difficult on the maths test”; “Maths tests make me feel nervous”; and “I did not feel anxious when answering the questions on the maths test” (reverse coded).

Reliabilities were good for measures of evaluation apprehension (4 items, $\alpha=.92$), and anger (3 items, $\alpha = .90$). A 7-item social identification scale ($\alpha=.99$), based on Ellemers, Kortekaas and Ouwerkerk (1999) was also included in order to measure the extent to which participants identified with the small group of students with whom they engaged in discussion. Items included, “I identify with the other students”; “I am like the other students”; “The other students are a reflection of who I am”; “I think the other students have little to be proud of” (reverse-coded); “I feel good about the other students”; “I have little respect for the other students” (reverse-coded); and “I would rather not belong to the same group as the other students” (reverse-coded).

A 3-item stereotype threat scale ($\alpha=.85$) was designed to measure how threatened participants felt by the stereotype: “I feel threatened by the stereotype that men are better than women at maths”; “I felt my performance on the maths test was negatively affected by the stereotype”; and “I feel threatened by the stereotype that women are worse than men at maths”. Participants answered each item on a Likert-

type scale, (1=“strongly disagree,” to 7=“strongly agree”) to what extent they agreed or disagreed with each statement. Please refer to Table 5.4 for scale intercorrelations.

Analytic Strategy

Hierarchical linear modelling (HLM) analyses were performed on the data in the 2 (Discussion topic: confirm stereotype vs. challenge stereotype) x 2 (Gender: male vs. female) design, in HLM 6.03 (Raudenbush & Bryk, 2002). Main effects and the 2-way interaction were tested using dummy variables (Judd et al., 1995; Rosenthal & Rosnow, 1985; Wilkinson et al., 1999). First, the confirm stereotype conditions [0] were compared to the challenge stereotype conditions [+1], for men [0] and women [+1]. Second, to interpret the main effect of discussion topic on maths scores, we performed a series of contrast analyses on additional models, which included the baseline control conditions (cf. Poortvliet, Janssen, Van Yperen, & Van de Vliert, 2007).

Results

Performance Effects

Maths test score. In the initial model, maths scores were entered at level one and the dummy variables were entered at level two. This initial model was significantly different to the null model, $\chi^2(53)=72.54, p=.04$, indicating significant between-condition differences overall. There was a significant gender main effect, $\gamma=-3.42, p=.02$ and a near-significant main effect for topic of discussion $\gamma=-.2.86, p=.06$. Please refer to Table 5.4 for means. These effects were qualified however, by a significant 2-way interaction between the topic of discussion and gender $\gamma=2.05, p=.03$. Women ($M=6.65, SD=2.62$) scored more than men ($M=6.04, SD=3.01$) when participants of both genders challenged the stereotype, and men ($M=6.85, SD=2.69$) scored more highly than women ($M=5.42, SD=2.82$) when participants of both gender confirmed the stereotype (Figure 5.3).

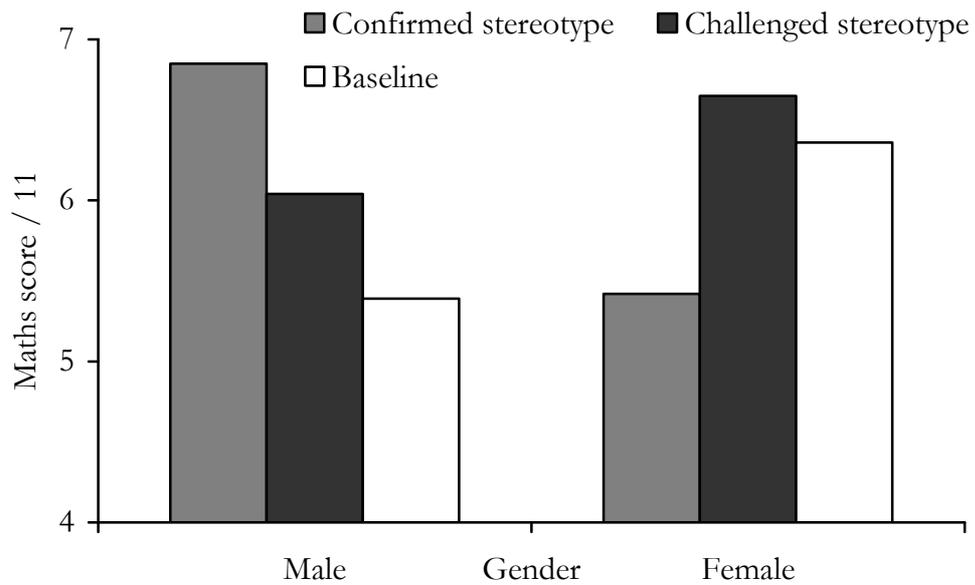


Figure 5.3. Mean maths test scores, Study 5.2.

In the baseline control condition, the difference between female and male participants was not significant $\gamma = -1.38$, $p = .11$, although there was a trend towards higher maths performance by women ($M = 6.36$, $SD = 2.77$) than men ($M = 5.39$, $SD = 2.66$).

After discussions confirming the stereotype, however, women performed significantly more poorly than did men, in line with the traditional stereotype threat effect, $\gamma = -1.25$, $p = .04$. After a discussion which challenged the stereotype, finally, women and men performed equally well, $\gamma = .81$, $p = .26$ (with women actually performing, if anything, something better), eliminating the stereotype threat effect on women's maths performance.

Examining the between-condition effects for women, a group discussion to *confirm* the stereotype produced significantly *lower* scores on the maths test ($M = 5.42$, $SD = 2.82$) compared with the baseline scores ($M = 6.36$, $SD = 2.77$), $\gamma = -1.34$, $p = .05$. When female participants challenged the stereotype, their scores were no different ($M = 6.65$, $SD = 2.62$) than scores of female participants in the baseline control condition

Table 5.4
Means, standard deviations, scale reliabilities, and intercorrelations of dependent variables, Study 5.2

Discussion condition		Male			Female			1	2	3	4	5	6	7	8
		Confirmed stereotype	Challenged stereotype	Baseline control	Confirmed stereotype	Challenged stereotype	Baseline control								
1. Maths test score	<i>M</i>	6.85 _b	6.04 _{ab}	5.39 _a	5.42 _a	6.65 _b	6.36 _b								
	<i>SD</i>	2.69	3.01	2.66	2.82	2.62	2.77								
2. Anxiety	<i>M</i>	3.78 _a	3.36 _a	3.47	4.77 _b	4.45 _b	4.49	-	(.81)						
	<i>SD</i>	1.45	1.38	1.32	1.26	1.50	1.14	.31**							
3. Evaluation apprehension	<i>M</i>	3.59 _a	3.17 _a	3.35	4.64 _b	4.26 _b	3.75	-.03	.44**	(.92)					
	<i>SD</i>	1.60	1.49	1.70	1.40	1.68	1.18								
4. Stereotype threat	<i>M</i>	1.99 _a	1.86 _a	3.66	3.17 _b	3.05 _b	3.59	.01	.16*	.19**	(.85)				
	<i>SD</i>	1.02	1.02	.64	1.49	1.46	.41								
5. Sub-group social identification	<i>M</i>	4.87 _a	5.06 _a	5.13	4.86 _a	5.05 _a	5.30	-.06	.08	-.03	.10	(.99)			
	<i>SD</i>	1.02	.96	1.02	.78	.77	.82								
6. Social validation	<i>M</i>	3.84 _{ab}	3.70 _{ab}	5.63	4.31 _b	4.79 _a	5.19	.04	-.07	-.03	-.09	.29**	(.86)		
	<i>SD</i>	1.68	1.70	1.08	1.22	1.10	.96								
7. Social support	<i>M</i>	4.88 _a	5.05 _a	5.17	5.04 _a	5.03 _a	5.05	.06	-.01	.01	-.01	.37**	.75**	(.95)	
	<i>SD</i>	.91	1.20	1.10	.84	.83	1.06								
8. Anger	<i>M</i>	2.83 _a	2.61 _a	2.84	4.01 _b	4.33 _b	3.96	.01	.23**	.19**	.24**	.04	.19**	.23**	(.90)
	<i>SD</i>	1.44	1.59	1.56	1.45	1.38	1.41								

Notes. * $p < .05$, ** $p < .01$. Cronbach's alphas in parentheses. Means in rows with different subscripts differ at $p < .05$.

($M=6.36$, $SD=2.77$), $\gamma=-.03$, $p=.97$. This result indicates that a short group discussion which challenges the conventional stereotype that women are bad at maths can eliminate the traditional stereotype threat effect on maths performance. Notably, women's scores were actually slightly *higher* when challenging the stereotype than in the baseline control condition. Finally, women who challenged the stereotype had higher maths test scores ($M=6.65$, $SD=2.62$) than women who confirmed the stereotype ($M=5.42$, $SD=2.82$) $\gamma=-1.25$, $p=.04$.

For men, a discussion that confirmed the stereotype led to higher scores on the maths test ($M=6.85$, $SD=2.69$) than the baseline ($M=5.39$, $SD=2.66$), $\gamma=1.46$, $p=.03$ providing evidence for stereotype lift. When group discussion challenged the stereotype, maths test performance ($M=6.04$, $SD=3.01$) was not significantly different from the baseline ($M=5.39$, $SD=2.66$), $\gamma=.60$, $p=.44$. It is notable, however, that even though the discussion was meant to challenge the stereotype, performance of men was somewhat *higher*. There was no significant difference between maths scores for men, $\gamma=.81$, $p=.26$, when they confirmed the stereotype ($M=6.85$, $SD=2.69$) or challenged it ($M=6.04$, $SD=3.01$), although scores in the confirm condition were somewhat elevated.

Psychological Effects

Anxiety and evaluation apprehension. Overall, women ($M=4.60$, $SD=1.33$) experienced significantly more anxiety than men ($M=3.55$, $SD=1.40$), $\gamma=.97$, $p<.01$. The topic of discussion main effect was not significant, $\gamma=.35$, $p=.17$, although there was slightly more anxiety when the stereotype was confirmed ($M=4.36$, $SD=1.51$) rather than challenged ($M=3.97$, $SD=1.54$). There was no 2-way interaction $\gamma=-.26$, $p=.61$. Similarly, there was a gender main effect for evaluation apprehension $\gamma=1.05$, $p<.01$, with women ($M=4.46$, $SD=1.53$) experiencing significantly more evaluation

apprehension than men ($M=3.36$, $SD=1.55$). However, there were no other effects ($p's > .12$). It appears therefore, that whilst consensualisation was able to eliminate the stereotype threat effect on women's performance, it did not eliminate the anxiety or evaluation apprehension experienced.

Social validation. Overall, women ($M=4.54$, $SD=1.18$) felt significantly more validated than men ($M=3.76$, $SD=1.68$) $\gamma=.79$, $p<.01$. Contrasts suggested this was due to women who challenged the stereotype ($M=4.79$, $SD=1.10$) feeling more validated than women who confirmed the stereotype ($M=4.31$, $SD=1.22$) $\gamma=-.58$, $p=.03$. There was no main effect for topic of discussion $\gamma=-.33$, $p=.26$, nor a 2-way interaction $\gamma=.52$, $p=.38$.

Anger. Overall, women ($M=4.16$, $SD=1.42$) experienced significantly more anger than men ($M=2.70$, $SD=1.51$) $\gamma=1.42$, $p<.01$. There was no main effect for the topic of discussion $\gamma=-.08$, $p=.75$ or a 2-way interaction $\gamma=.44$, $p=.39$, however.

Stereotype threat. Responses to the stereotype threat scale were then analysed for condition effects. Again, there was a main effect for gender, with women ($M=3.11$, $SD=1.47$) feeling significantly more threatened than men ($M=1.92$, $SD=1.02$) $\gamma=1.14$, $p<.01$, but neither the topic of discussion main effect, $\gamma=.19$, $p=.45$, nor the 2-way interaction were significant, $\gamma=-.12$, $p=.81$.

Identification. Social identification scores were also explored for condition differences, however, there were no main effects, nor an interaction (all $p's > .19$).

Mediation. Next, individual level anxiety, evaluation apprehension, social validation, threat and anger were entered into the model as grand centred effects at Level 1, predicting maths scores. Only the slope for anxiety $\gamma=-.63$, $p<.01$ was significant, however. As anxiety was a significant predictor of maths scores and also

differed across conditions, it was each entered alone as grand centred level 1 predictor, with maths score as the outcome variable. The dummy variables were entered separately at level 2. The slope for anxiety was significant $\gamma = -.68, p < .01$. However, on entering this mediator into the model, the 2-way interaction remained significant $\gamma = 1.90, p = .04$; suggesting anxiety did not mediate the effects of discussion condition on maths score.

Discussion

The aim of the second study was to investigate the effects of group discussion on women's and men's maths performance. First, it was predicted that gender and discussion topic would interact. Results showed that women performed significantly worse than men and than their baseline score when they had confirmed the stereotype during discussion, in line with the traditional stereotype threat effect. Male participants' maths scores, however, were lifted when they confirmed the stereotype through discussion, as hypothesised.

Second, it was predicted that when men and women were given the opportunity to challenge the stereotype through discussion, the stereotype threat effect should be eliminated, with no difference on maths performance between the gender groups. In line with this hypothesis, results showed that women did no worse than men after a group discussion which challenged the stereotype. In fact, the stereotype threat effect was slightly *reversed* when women had challenged the stereotype in their group. Therefore, whilst consensualisation processes can reinforce the negative stereotype, they also appeared to be able to equip women to *resist* its effects.

Psychological Effects

Social identification showed no effects, so the changes in performance over the

different consensualisation conditions were not due to changes in group identification. Women reported significantly more anxiety and evaluation apprehension than men. There is a plausible explanation for this, however. This finding could be a result of the “fight” or “flight” response: For women who confirmed the stereotype conditions, responses to the anxiety scale may have represented “flight”. Conversely, it may be characterised more as a “fight” response when they challenged the stereotype.

Furthermore, women were angrier about, and more threatened by the stereotype than men (when it was both confirmed and challenged). Neither anger nor perceived threat reduced for women when they challenged the stereotype, relative to when it was confirmed. Despite this emotional response, women felt significantly more validated than men, and challenging the stereotype was more validating for them than confirming it, as logic would dictate. Furthermore, anxiety was a significant predictor of maths performance. However, there was no evidence of these psychological variables mediating maths performance. It appears therefore, that whilst consensualisation was able to eliminate the stereotype threat effect on women’s performance, we cannot conclude that it significantly transformed their emotional experience.

Therefore, there was no evidence of mediation to help explain the effectiveness of our intervention. It may be speculated, therefore, that although variables such as anxiety were involved in the experience of stereotype threat, they may not have been the primary driving force of the threat effect on performance. In other words, it was not the fear of stigmatisation that was driving the stereotype threat effect. Rather, we argue that stereotype threat may have been due to a direct application of the negative stereotype to the self (i.e., self-stereotyping or the expression of identity content).

Arguably, self-stereotyping could produce these effects simply because of the application of a “valid” view of the in-group identity content to the self.

General Discussion

The purpose of the present research was twofold: First, to investigate whether challenging the negative stereotype of the in-group through intra-group discussion could enable in-group members to resist the effects of stereotype threat on performance; and second, to establish whether the former is achieved via intra-group consensualisation because it enables in-group members to collectively redefine the in-group stereotype.

The results provide support for both hypotheses. Study 5.1 demonstrated that intra-group interaction (either confirming or challenging the stereotype that women have lower mathematical ability than men) could qualitatively transform the performance of in-group members relative to individual thought about the topics. The results of Study 5.2 corroborate previous research on stereotype threat, i.e., negatively stereotyped participants’ performance was detrimentally affected by the stereotype when asked to acknowledge that the stereotype is founded in reality. The findings also replicate the more recently described phenomenon of stereotype “lift” (Marx & Stapel, 2006; Walton & Cohen, 2003), whereby the negative stereotyping of the out-group (women) was utilised to *enhance* in-group (men’s) performance. The present research makes a unique contribution to this literature, however, showing that the performance of *stigmatised* in-group members can be “lifted” after they have been given the opportunity to collectively resist the stereotype (relative to when it was confirmed). This could be more aptly named “stereotype transcendence”.

On the other hand, when trapped by the psychological shackles of the negative

in-group stereotype, consensualisation channelled and focussed in-group members' attention on the positive out-group stereotype, meaning that performance on the relevant dimension decreased for in-group members. It is argued that confirming the in-group and out-group stereotypes contextualised expectations of in-group performance within the inter-group dynamic. Analyses suggest that hearing other in-group members confirm the stereotypes solidified consensus of the out-group's positivity on the relevant dimension relative to the in-group, increasing evaluation apprehension, and decreasing performance.

These shackles of stigma were easily shed when participants were invited to challenge the negative in-group stereotype, however. Participants appeared to be able to use a brief group discussion to come up with a range of consensual explanations (or theories) which invalidated the stereotype and replaced the stigmatised identity content with a more positive alternative identity. Notably, it appeared essential to discuss this with other in-group members. Merely reflecting on this issue individually proved insufficient to invalidate the stereotype and its impact on subsequent behaviour. This is because the power of stereotypes is derived from their being *socially shared*.

The common way of theorising about the processes behind stereotype threat is to look for mediating and moderating factors. However, as stated earlier, there is little consensus in the literature about what these actually are (Jamieson & Harkins, 2007; Shapiro & Neuberg, 2007). In the present research, it is argued that variables such as anxiety, evaluation apprehension, social validation and anger could not explain the effect of stereotyping on performance: we found no evidence of mediation (although no conclusions can be drawn from such null effects). Potentially, although these variables are involved in the experience of stereotype threat, they may not be the

primary driving force of the threat effect on performance. This is a speculative argument, although we did indeed fail to find mediation across both studies. Instead, the evidence suggested that the stereotype content itself is what influenced performance. For if a group member stereotypes themselves, their performance should be implicitly guided by the norms associated with the stereotype. Crucially, what those norms for performance actually are will depend upon the consensual definition of the stereotype content. This process would not be mediated by anything except salience (which was constant across conditions in the studies presented here).

Indeed, our manipulation of identity content shifts (demonstrated by the qualitative evidence from Study 5.1) promoted performance changes. These shifts appeared to provide the in-group with a positive social identity (or in-group stereotype relative to the out-group stereotype) with associated implicit norms for behaviour on the relevant dimension. This strategy allowed in-group members to avoid the pitfalls of being “automatically” affected by a negative group stereotype. It was clear from the results, however, that these strategies cannot be effectively used through isolated individual cognitions. Instead, intra-group consensualisation enabled in-group members to exchange validating ideas about what constitute positive in-group characteristics and therefore, gain valuable, socially shared information about normative in-group attitudes and behaviour.

Implications and Future Directions

The present results show that social and behavioural change can be effected by collectively challenging a stereotype. There is potential, therefore, for the theory articulated in the current research to be utilised to develop applied interventions. The social psychological literature focuses largely on “contact”, i.e., communication with the

out-group to improve inter-group relations (e.g., Allport, 1954; Pettigrew & Tropp, 2006). However, current research points to the benefit of *intra*-group contact for social change. Further research into the contextualisation and re-formulation of the in-group stereotype within a broader inter-group framework could take a step towards redressing some of the group-based inequalities that exist in schools, colleges, universities and organisations. At the very least, the present results provide hope that small intra-group discussions can be a positive instrument for change. After all, certain stereotypes have been considered relatively stable. Considering this, it is astounding that brief, small group discussions with fellow targets of prejudice can be effective in helping them overcome the impact of negative self-stereotypes. It is therefore, imperative that future research investigates further the process through which intra-group interaction achieves this change. Intra-group interaction and interdependence can be a solution (for societal problems) rather than a problem (for methodological purity) in social psychology.

Conclusion

The studies reported here show that the consequences of negative societal stereotypes can be overcome if people are given the opportunity to critically examine those stereotypes in conversation with their in-group. Through social interaction, therefore, in-group members can overcome the strictures of social stereotypes. The perception and consequences of stigma are malleable to those who have the opportunity to reach an alternative consensus with the in-group. In the same way that social consensus forms the foundation for out-group stereotypes, it also lies at the heart of in-groups' self-stereotypes. Therefore, whilst consensualisation can maintain self-stereotypes and stigmas which are handicapping, on the positive side, it can be a powerful instrument for social change.

CHAPTER 6

General Discussion

“I think, therefore I am” (Descartes, 1656/1960)

Vs.

“We talk, therefore we are.”

I am now in a position to reflect upon the contribution of this thesis to the established literature, evaluated in Chapter 1. In the opening chapter, I described the role of communication in the provocation of the Cronulla riots in Australia, and the Rwandan genocide. I then explained how these examples rest within the theoretical agenda of the current research. To reiterate, most empirical research in social psychology to date has studied the influence of social perceptions (as cognitive input to individuals) on individuals' social behaviour (see Figure 6.1, [A]). This construes cognition, indeed the perception of all social reality, as constructed entirely within the individual mind. This metatheoretical stance is a spectre of the metaphysics of Descartes, who argued that the very existence of a thought about one's existence is proof that one exists (“I think, therefore I am”, or “cogito ergo sum”). Now, most social psychologists would agree that individual thought is interdependent (e.g., Turner & Oakes, 1997), and informed by past knowledge, including the memory of past interactions. However, the interdependent nature of Human cognition is still underplayed. For example,, an individual's cognition about a group does not necessarily mean that the individual's perception of the group is valid. The social validation of other individuals is necessary to establish that social reality. Moreover, we

are not passive observers of an objective reality that is “out there” and external to us, but we are constituents of that reality and agentic in its construction and perception (see Núñez, 1997; Varela, Thompson, & Rosch, 1991).

Yet, the question of how these shared social perceptions *develop* in the first place is rarely asked in social psychology. What generally remains implicit and empirically unexplored, is the pivotal assumption that these perceptions of social reality cannot emerge from individual cognition alone—they are, or are founded on, knowledge structures that are consensually held among members of a particular group. In other words, “We *talk*, therefore *we* are” would be a more accurate representation of the metaphysics of social reality, from this perspective.

I argued that, under certain conditions, social communication can influence in-group members’ perceptions of social reality, and transform the way they understand what constitutes legitimate, valid social behaviour. One can conceive of this as a form of identity content change, which includes the development of norms for social action. For example, the Cronulla riot protagonists used an SMS to validate and legitimise racist violence. The radio broadcast by Hutu extremists used a negative stereotype of the Tutsis to make their harm seem more valid and legitimate. It was also a signal of in-group consensus on a plan of inter-group action. This stereotypical and (implicit) ideological content was the foundation of the subsequent hostile social action by Hutu extremists.

I have provided evidence to suggest that the idea of consensualisation (or interaction), as a process through which *existing* stereotypes become more shared (or are, in other words, disambiguated), is a limited conception. Over and above this existing work on consensualisation, I wanted to show that consensualisation is the process through which social realities (not just stereotypes, but also their validity) are established and transformed. As realities, they can become the foundations for

collective action, in a way that is more difficult for privately held cognitions. In this way, this project is theoretically significant in its novel explanation of (inter-group) social action as motivated by intra-group processes, instead of exclusively through individuals' socio-structural perceptions (Figure 6.1, [A]). The new theoretical framework allows us to understand one of the reasons why people who discuss particular social ideologies are motivated to engage in social action. Through this, the research contributes to our understanding of inter-group behaviour in general.

The Purpose of this Thesis

The purpose of this thesis was twofold: (a) to examine, from a social psychological perspective, how social realities are established and transformed through interaction, and (b) to investigate why the process of intra-group interaction can spark and exacerbate social conflict. In addressing these aims, I intended to demonstrate that social interaction serves to establish norms for social behaviour by socially validating individual cognitions or observations (Figure 6.1, [B]).

From this theoretical vantage point, we can see that there are (at least) two related functions of intra-group interaction. The first function is to establish or bolster a sense of "us". The second is to inform "us" about "them", i.e., to validate out-group stereotypes. In this way, intra-group interaction may inform "us" about "who we are" and "what we should do".

Brief Summary of the Results

Nine studies were conducted in order to test this perspective (please refer to Table 6.1 for a summary of key findings). In line with the purpose of this thesis outlined above, we have demonstrated that (a) intra-group interaction transforms social stereotypes and behaviour, and (b) the processes through which this is achieved is contingent upon reaching a validating social consensus with other in-group members.

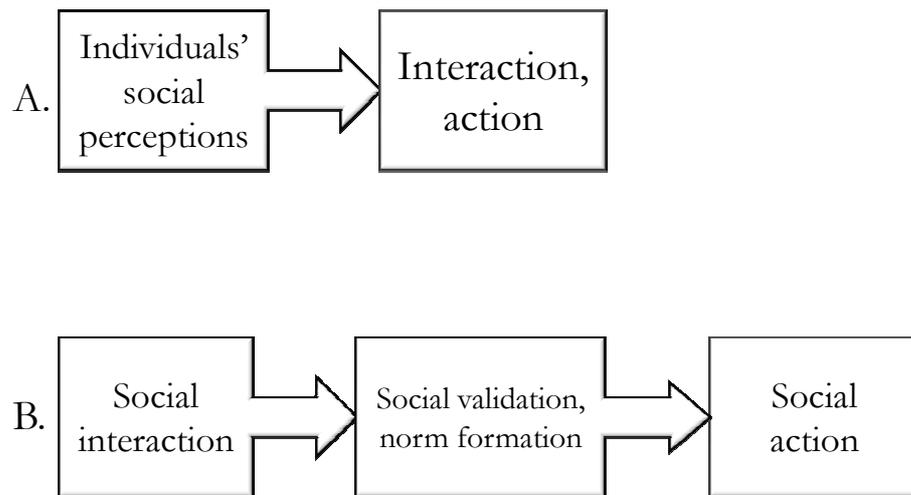


Figure 6.1. Models of social action: [A] Traditional model of social action, [B] Towards an interactive model of social action (IMSA)

In the first empirical chapter (Chapter 2), two studies investigated the positive-negative asymmetry effect (PNAE). An important point to reiterate from these results is that the novel effect – a reversal of the traditional PNAE - occurred only when there was *no* interdependence of group outcomes. Therefore, in the absence of inter-group antagonism, when given the opportunity to interact, participants looked to other in-group members to inform and validate their actions. Not only this, but interaction served to establish perceptions that an in-group norm for hostile inter-group action was legitimate.

The results of Chapter 3 provided support for the argument that social validation of out-group stereotypes which justify hostile action, plays a key role in the qualitative transformation of individual attitudes into intentions for hostile collective behaviour. It can be concluded that it was not interaction with in-group members *per se* or discussion of specific actions that led to greater preparedness for action. The implicit action intentions contained in stereotypes (e.g., in the sense that negative

stereotypes are associated with prejudicial behaviour, Reicher et al., 1997) were more effective in eliciting action in a context where the legitimacy of those actions would be contested, should they be made explicit. Thus, the discussions validated participants' stereotypes, thereby providing a firmer foundation for their intentions to engage in hostile social action. Their social perceptions were, in a sense, transformed into a social reality within which they could legitimately act.

The research reported in Chapter 4 found that crucially, action was perceived as normative only when consensualisation occurred. Under conditions of consensus, the legitimacy of hostile behaviour was *negotiable*, and social action was legitimised and validated by the interaction. Together, the results of Chapter 3 and 4 highlighted the key role of intra-group consensualisation processes in determining identity content. It is argued that these processes have a pivotal and often overlooked influence on inter-group relations.

In demonstrating that the perceived legitimacy of hostile behaviour is not fixed, this research underlines the importance of the interplay between intra- and inter-group dynamics for understanding why groups sometimes feel they can justify violent actions towards others. Therefore, research into inter-group relations which ignores the relevance of the intra-group dimension may not have the explanatory power required to adequately account for "real-world" actions, particularly if those actions would be controversial, harmful, or otherwise costly to the actor.

The empirical journey of this thesis reaches a positive endpoint. The studies reported in Chapter 5 imply that interaction can not only transform the perceived validity of out-group stereotypes, but it can similarly affect the in-group's self-stereotype. In this chapter, discussion was shown to *invalidate* a negative self-stereotype. Furthermore, and importantly, these results suggest that changing self-stereotypes can be used to promote positive behavioural change, as much as to promote hostility or (in

Table 6.1

Summary of results

Study	Topic	Design	Key Findings
2.1	PNAE	2 condition (Valence of monetary allocation: Positive vs. negative) within-subjects design.	Participants were more likely to discriminate on rewards than fines, find allocating rewards to be a more legitimate and pleasant act than allocating fines and believe that reward allocation would have a more positive impact on recipients than fine allocation. Conversely, participants thought allocating fines would have a more negative effect on recipients and felt more negative about allocating fines than rewards.
2.2	PNAE	2 (valence of monetary allocation: reward vs. fine) x 2 (group advancement: obstructed vs. not-obstructed) x 2 (phase: pre-consensus, consensus) mixed factorial design, with repeated measures on the phase factor.	When in-group advancement was obstructed, no PNAE was found: obstruction was sufficient justification for out-group punishment in its own right. When in-group advancement was <i>not obstructed</i> , the PNAE reversed after group discussion, such that more hostility occurred when participants administered fines than when they awarded rewards. This reversal was mediated by processes of norm formation.
3.1	Hostile action towards Immigrants	2 condition (Type of reflection: group discussion vs. individual reflection) between-subjects design	More hostile action intent after group discussion than after individual reflection. In-group norm supporting hostility mediated effect of condition on hostile action intentions.
3.2	Hostile action towards Immigrants	3-condition (Topic of interaction: out-group stereotype vs. out-group stereotype plus social action vs. irrelevant control) between-subjects design	The lowest advocacy of social action, validation, efficacy and anger scores occurred in the irrelevant discussion (control) condition compared with the experimental conditions. Discussing action as well as the out-group stereotype resulted in <i>less</i> hostile action intentions, social validation and collective efficacy compared to discussing only the out-group stereotype. Social validation mediated the effects.

3.3	Hostile Action towards Immigrants	2 (Type of reflection: Group discussion vs. individual reflection) x 2 (Topic of reflection: Out-group stereotype vs. out-group stereotype plus action) between-subjects design	Conceptual replication of 3.1 and 3.2 with behavioural measure. Discussion of the out-group stereotype increased hostile action relative to <i>all</i> other conditions. No significant mediation.
4.1	Collective action towards fast food outlets	4 condition (Content of discussion: irrelevant control vs. stereotype only vs. stereotype plus mainstream action vs. stereotype plus extreme action) between subjects design.	Action intentions were equally high in the two conditions in which there was high consensus over action, and lowest in the control condition. Effects of condition on action were mediated by social validation and the perceived legitimacy of action.
4.2	Collective action towards immigrants	3 condition (Direction of consensus: Consensus to dissent vs. Dissent to consensus vs. no discussion control) between subjects design	After watching a discussion during which consensualisation was achieved (rather than undermined), participants were more likely to advocate taking action. Identification and perceptions of the message as normative mediated the action intention results.
5.1	Stereotype threat	2 (Topic: Confirm stereotype vs. challenge stereotype) x 2 (Reflection: Group discussion vs. Individual reflection) between-subjects factorial design	Stereotype threat was undermined when in-group members collectively challenged the stereotype. Content analyses suggested that discussions redefined in-group and out-group stereotypes and increased consensus, providing the basis for stigma reversal or stigma confirmation.
5.2	Stereotype threat	2 (Discussion topic: confirm stereotype vs. challenge stereotype) x 2 (Gender: male vs. female) between-subjects factorial design plus a baseline control condition.	After a discussion that confirmed the stereotype, women displayed signs of stereotype threat and men's performance was "lifted". When they challenged the stereotype, performance differences between men and women were eliminated.

this case) self-handicapping behaviour. In sum, the research described in this thesis demonstrates that intra-group interaction can qualitatively transform the social behaviour of in-group members through processes of social validation, stereotype change and norm formation (see Figure 6.1, [B]).

Theoretical and Practical Implications

Towards an Interactive Model of Social Action

The findings and arguments of this thesis suggest the need to formulate a new, interactive model of social action (IMSA) which would reconceptualise the psychological changes that are assumed to occur within individuals before they act (Figure 6.1, [B]). The results reported in this thesis suggest that individuals need to form a social identity with validated and concrete content, including norms for action, before they can engage in social action. To do this, first, individuals engage in intra-group interaction. This enables the group members to develop an understanding of their common ideology, which may establish the consensual basis of their identity content. If such consensualisation occurs, this provides them with a sense that their perceptions of reality are socially valid, and gives rise to (implicit or explicit) in-group norms. This provides individuals and groups with the cognitive resources that form the foundations upon which they may proceed to act, i.e., by legitimising those courses of action.

There are significant implications of this metatheoretical shift for the ways in which both identity, and action, should be studied. As depicted in Figure 6.1 [A], social communication is often viewed as a by-product or consequence of individuals' social perceptions (e.g., Hogg, Coopershaw, & Holzworth, 1993). An IMSA would improve upon this traditional conceptualisation in an important way: it would disentangle the role of individual cognition from the input we receive from people with whom we

interact, which are often confounded in prior research. This builds on insights from the inductive model of identity formation (IMIF; see Postmes, Baray et al., 2006).

The IMIF sees polarisation not only as a consequence of self-categorisation, but as a tool for identity (or “category”) formation. If seen in this way, the process of group polarisation establishes novel group norms, or identity content. The data presented in this thesis demonstrates that it is entirely possible to alter self- and other-stereotypes, as well as social behaviour, through intra-group interaction. However, this is simply not possible when restricted by the assumptions of the self-categorisation theory (SCT; Turner, 1985; Turner et al., 1987) account of group polarisation, based on referent informational influence (RII; Turner, 1982). This perspective posits a top-down, deductive process of social influence deriving from an individual’s self-definition as a group member. This does not allow for the possibility of bottom-up construction of identity.

Furthermore, SCT’s principle of functional antagonism states that either social or personal identity can be salient at any one time. This suggests that interpersonal influence could not occur at the same time as social influence (i.e., the effect of self-categorisation as a group member on individual thoughts and actions). Obviously, this implies that interaction in small groups must be governed either by social identity concerns, or interpersonal interaction. Yet, the combination of the data plus the theoretical arguments presented in this thesis, suggest that this is entirely possible.

Traditional research, such as that undertaken in the social identity paradigm, assumes that individuals are aware of the “shared” perceptions which are hypothesised to drive group-based action. However, our research has demonstrated that reality is not shared *a priori* but as a *process* which leads to action. The social identity approach to inter-group behaviour does not hitherto theorise about this intra-group “sharing process”. The present research adds to the body of evidence that suggests that, in fact,

the inter-personal, intra-group and inter-group aspects of group life are much less antagonistic or independent of each other than is often assumed (see also Haslam, Egghins & Reynolds, 2003; Spears et al., 2006; Turner et al., 2006).

It is appropriate at this stage to highlight the central role that social validation played as a mediator. The only meaningful (and, evolutionarily speaking, *original*) way of validating reality is through social means. Interaction has been shown to change how valid a stereotype is perceived to be, even if it does not change the content or valuation (i.e., in terms of prejudices) of that stereotype *per se*. This change in validation in turn, is sufficient to change how individuals act even in the absence of a shift in prejudice. Thus, rather than stereotypes, prejudices, and inferred social norms for action being “givens” or societally “generic” cognitive inputs to individuals, they are inductively construed or re-construed (i.e., transformed into something which could not be achieved on the basis of individual cognitive processing alone) through interaction with the in-group. In recognising that fellow in-group members constitute a social and ideological resource for socially-identified individuals to discover, negotiate and validate their belief systems (see Correll & Park, 2005; Festinger, 1950; Festinger, 1954), SCT-based approaches would allow for a more fluid, agentic and dynamic conceptualisation of social behaviour.

This has implications for how results of minimal group studies on the positive-negative asymmetry effect are understood. These studies robustly demonstrated the tendency for isolated individuals to display in-group favouritism. However, research would seem to suggest that individuals in minimal groups are reluctant to administer negative outcomes or even to punish the out-group. The “punishment” of an out-group has been shown to occur only when inter-group comparisons make it seem appropriate (Reynolds, Turner, & Haslam, 2000). To explain this relative lack of out-group punishment, it was suggested that there may be a “generic” norm against out-

group punishment in society (cf. Mummendey et al., 1992). I argue that a different approach to understanding the social conditions which give rise to perceptions of legitimacy such as described in Figure 6.1 [B], would explain the processes underlying these findings far more effectively.

This raises the question of what can actually be predicted from individual cognitions and behaviours that arise in the social vacuum of the psychology experiment. In other words, is there a useful psychological default, or a “generic” norm to which individuals defer for social action? For, if “defaults” are questioned by individuals, and necessarily socially mediated, prior research has used contrived methods of social psychological investigation by routinely using isolated individual cognitions as its unit of analysis. Turner (1982, 1991; Turner & Oakes, 1997) acknowledges that all perception is socially mediated. Research such as that into pluralistic ignorance (e.g., Miller & McFarland, 1987; 1991; Prentice & Miller, 1993) and the bystander effect (Latané & Darley, 1970) demonstrate the harmful consequences of individuals misunderstanding reality when they do not have the opportunity for social validation. Yet ironically, other research attempting to measure social reality continues to trust, and empirically depend upon, those isolated individual cognitions. These traditional methods reduce socially defined reality to individual cognitive inputs, thus making it impossible to study a key theoretical process driving social behaviour.

I argue that in order to further our understanding of social behaviour, we need a change of method (and therefore, theory) in line with the above analysis. By doing so, we may be able to take a step towards redressing the “evident mismatch between the effect that most research on inter-group bias has studied and most theories have sought to address and the most striking social problems that this research ought to be able to address,” as called for by Hewstone, Rubin, and Willis (2002; p. 594) and as quoted at the opening of this thesis. We acknowledge that we are followers in this

“metatheoretical shift”, more than leaders. In fact, the magnitude by which I propose to shift the focus is nothing compared to the relativism of researchers like Billig (e.g., 2006), Wetherell (e.g., 2007), or Antaki and colleagues (e.g., Antaki; 2006; Antaki, Condor, & Levine, 1996) —that there are no social realities outside of interaction, and that therefore, the whole concept of cognition is redundant for social psychology.

Considering Potential Alternative Explanations

Central to the ideas and explanations for empirical findings that are advanced in this dissertation, are the twin assumptions that (a) interaction is the process by which individual cognition about inter-group relations acquires its validity and therefore power, and that (b) interaction can only acquire this power to the extent that it is anchored in social consensus. Inevitably, such general assertions raise numerous questions and alternative explanations which may challenge the interpretation of results and therefore the theoretical framework. These alternatives and questions may of course be addressed empirically. To some extent, our research has done this already, or so I would argue. Some of them may also be addressed theoretically, however, or they may be logically inferred (cf. Searle, 1995). I now address some of the questions which are most frequently asked of this research.

“Isn’t this still just individual cognition?”

This question is essentially a variation on the classic Allportian charge to the group mind thesis: That there is no group process which is not, ultimately, reducible to individual cognitive processes (Allport, 1924). It is clearly the case that the phenomena we describe are grounded in, and can be described as, individual cognitive processes, in the same way that can also be understood as neurophysiological processes of the brain. However, the mere fact that the *description* of the phenomena can be made at multiple levels of abstraction (social, individual, neurological, chemical) and that each of these levels, in its own way, *explains* what happens, does not mean that each of these levels

can in isolation of the others account for the social phenomenon that we are witnessing. We argue that an emergent sense of social consensus that is inferred from interaction provides individuals with a subjective validation of social reality. The simple but important reason why individual cognition alone can never explain such a phenomenon, is that communication is logically the only way in which those individual cognitions may be validated.

“Is interaction a necessary condition for these effects to occur?”

Returning to the very beginning of this thesis, the opening quotations in the introductory chapter from Hewstone et al., (2002) and Tajfel (1972) were used to highlight the theoretical reasons for the methodological choice of studying the phenomena that are central to this research through the medium of small group interactions. Our results confirm that interaction has a significant effect on identity and action. However, one may ask whether interaction is necessary to produce these effects (see also Chapter 3). Indeed, in Chapter 4, we showed that when participants observe a staged discussion between discussants (who they believed were in-group members), similar effects were witnessed. There is a range of other means by which these effects could potentially be invoked: For example, by providing (false) feedback about in-group consensus. Ultimately however, these alternative ways of producing comparable effects do not contradict the centrality of interaction—they underline it. The psychological function of intra-group interaction is, at least in part, to assess consensus on and thereby, validate individual beliefs. It is therefore, logical that consensus feedback has similar effects.

So why study interaction? I would like to answer this question with another of my own: Why study the social functions of interaction by resorting to methods that remove interaction itself from the equation? Even if one can *model* the consequences of intra-group interaction with a simulation of it, as we did in Chapter 4, and produce

similar effects, one may inadvertently remove critical processes or dynamics from the equation, and thus never “discover” their relevance to the object of study. In view of the aim of the research, to study how and with what consequences social realities are validated through interaction, the question is nonsensical. It would only make sense to actively interfere with the content of interaction, as was done in Chapter 4, in order to test very specific hypotheses which have been formulated on the basis of prior research that preserves the ecology of the process under observation.

Questions about the necessity of interaction in experimentation are often raised on the basis of questionable views that favour experimental robustness over ecological validity: It has become “good practice” to run experiments on isolated individuals, and research involving social interaction is seen as methodologically questionable (Haslam & McGarty, 2001). The assumption that interaction poses a problem for the robustness of experimental designs is due in part to concern over statistical independence of results. However, modern statistical techniques, such as the multilevel modelling employed in the current research, are perfectly capable of partialling out variance caused by group-level and individual-level factors (Raudenbush & Bryk, 2002; Snijders & Bosker, 1999). There may have been a time when running group research of this kind was a methodological problem, but that time is long past. In fact, given the ready availability of these statistical solutions to the problem of group dependence, one could ask the question why research practice in social psychology is still not adopting what has been made technically possible since a few decades. The desire to make inferences about group processes through the study of individuals in isolation, is much more questionable than studying individuals who are nested in a dynamic group environment. This is simply because the former does not adequately replicate the conditions of group behaviour, and therefore may inadvertently change this behaviour in fundamental ways.

A related concern with experiments involving between-participant interactions that is often raised is that the nature of such interactions entail a lot of “noise” and would imply a lack of experimental control. However, the fact that condition effects emerge despite this noise only adds confidence to those effects of experimental conditions which emerge consistently (i.e., those effects which can be replicated must be of very considerable magnitude). The final charge - that such interactions do not allow for experimental control - has more merit. It is indeed the case that group interaction confounds two processes: (a) those effects which are due to group interaction *per se*, and (b) those effects which are due to elaboration of a specific topic. This is a limitation that can easily be addressed by using an appropriate control condition, and if possible by using multiple controls. As outlined in the empirical chapters, ideally an experimental condition involving interaction should be compared with a no interaction control condition, as well as with a control condition in which the interaction considers an unrelated topic. In combination, these conditions would provide all the experimental control required.

“Isn’t this just polarisation?”

Any answer to this question requires a definition of polarisation. In the original group polarisation studies, the phenomenon was defined “narrowly” as an attitude shift, i.e., the difference between an individual’s opinion (e.g., on a 7-point Likert-type scale) before and after discussion. In the present research however, the overall effect of group discussion on attitude change was descriptively small. In fact, in none of the studies included in this thesis did a significant shift in prejudice occur due to the experimental interventions—there is simply no evidence that any significant attitude change took place. Instead, the results attest to the capacity of social interaction to validate the out-group stereotype.

“Polarisation” can also be defined more broadly to refer to a change or extrematisation of the “climate” in a group or society, or even as a marked shift in group behaviour. This broad definition is of course much more closely aligned with the everyday use of the term polarisation, although it is not what social psychology ordinarily understands by it. Our research certainly shows, across studies, a relatively consistent change in the valuation (or validity) of prejudices against out-groups, and an even more consistent change in behaviour towards those out-groups (or in the case of Chapter 5 a shift in self-stereotypical behaviours). Nevertheless, we did not find evidence for a parallel shift in individual attitudes. Thus, the social sharing of previously held attitudes resulted in an increased sense that they were subjectively valid (insofar as consensus was achieved in the discussion), but not in opinion change.

This research significantly and innovatively builds upon the group polarisation literature by linking the content of in-group and out-group *stereotypes* to social *action* and social *behaviour*. Whereas the polarisation research demonstrates the effect of group discussion on attitudes, we extend this to social action. This is significant because, as it has often been demonstrated, the link between attitudinal intentions and behaviour is not always reliable (e.g., see Hogg & Smith, 2007). Therefore, we have demonstrated that interaction could have an impact on society through driving behavioural change, rather than on the political “climate” alone.

Furthermore, polarisation studies conceptualise reality as impacting upon attitudes through discussion (Figure 6.1, [A]). The present research conceptualises the opposite: Discussion impacting upon perceptions of shared reality (Figure 6.1, [B]). In other words, the changes that occur within groups which result in polarisation are shown to feed back into the identity process. By conceptualising interaction as a tool to shape shared reality, rather than making individuals’ views strengthened in the direction in which they already tended, the present research demonstrates the ability of

intra-group interaction to guide stereotypes and behaviour in opposite directions (Studies 2.2, 5.1 and 5.2) depending on the framing of the discussion and content of the in-group norm. Thus the present research shows a *qualitative transformation of identity content* rather than a crystallisation effect.

Turner's (1991) review of prior research which tested SCT's prediction of group polarisation, focused on studies that were framed within explicit inter-group contexts. The research described in this thesis however, demonstrates that "polarisation", or rather, identity content change, was not (necessarily) driven by inferences made from the socio-structural context. Study 5.1 demonstrated that it was the fact of reaching a consensus within the discussions, that drove the effect, and this consensus was not dependent on the inter-group context. Similarly, the effects of Study 4.1 also occurred with no pre-defined inter-group relation. The original group decision studies of Lewin (1953), the risky / cautious shift findings (Fraser, Gouge, & Billig, 1971; Stoner, 1968; Wallach & Kogan, 1965) as well as Moscovici and Zavalloni's (1969) polarisation research, were not conducted within an inter-group context, and therefore, were not dependent upon one. I propose that there were additional processes, over and above the opinion shift, occurring in these studies. The research presented here demonstrates that changes at the intra-group level impact upon fundamental processes of norm formation, consensualisation, transformation of stereotypes, and social validation.

Concluding Remarks

The studies reported here gave participants the opportunity to reach out beyond the confines of traditional laboratory research, and discuss the context in which they were placed. Through using this socially interactive paradigm, I was able to show that externally imposed stereotypes are not automatically internalised by in-group members. This research therefore, represents an attempt to overcome the reductionism of the

Cartesian epistemological construal, which has for too long narrowed the spotlight of social psychological research, viewing individuals as the prime causal units of human behaviour. In advocating a different approach to studying the social psychology of groups, I hope to have started a programme of research which shows that, rather than being slaves of a static society, people are agents with the potential to collectively shape their own destinies.

REFERENCES

- Abrams, D., Wetherell, M., Cochrane, S., Hogg, M. A., & Turner, J. C. (1990). Knowing what to think by knowing who you are: Self-categorisation and the nature of norm formation, conformity and group polarisation. *British Journal of Social Psychology, 29*, 97–119.
- Ajzen, I. (1985). From intentions to actions: A theory of planned behaviour. In J. Kuhl & J. Beckmann (Eds.), *Action control: From cognition to behaviour* (pp. 11-39). New York: Springer-Verlag.
- Ajzen, I. (1991). The theory of planned behaviour. *Organizational Behaviour and Human Decision Processes, 50*, 179-211.
- Allport, F. H. (1924). *Social psychology*. New York: Houghton Mifflin.
- Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.
- Amiot, C. E., & Bourhis, R. Y. (2003). Discrimination and the positive-negative asymmetry effect: ideological and normative processes. *Personality and Social Psychology Bulletin, 29*, 597-608.
- Amiot, C. E., & Bourhis, R. Y. (2005). Discrimination between dominant and subordinate groups: The positive-negative asymmetry effect and normative processes. *British Journal of Social Psychology, 44*, 289-308.
- Amnesty International. (1996). *Rwanda and Burundi, the return home: Rumours and realities*. New York: Amnesty International.
- Antaki, C. (2006). Producing a 'cognition'. *Discourse Studies, 8*, 9-15.
- Antaki, C., Condor, S., & Levine, M. (1996). Social identities in talk: Speakers' own orientations. *British Journal of Social Psychology, 35*, 473-492.

- Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999).
When white men can't do maths: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*, 29-46.
- Asch, S. E. (1952). *Social psychology*. New York: Prentice Hall, Inc.
- Asch, S. E. (1956) Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs, 70*, 416.
- Asch, S. E. (1958). Interpersonal influence: Effects of group pressure upon the modification and distortion of judgments. In E. E. Maccoby, T. M. Newcomb & E. L. Hartley (Eds.), *Readings in social psychology*. New York: Holt & Co, Ltd.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioural change. *Psychological Review, 84*, 191-215.
- Baron, R. S., Hoppe, S. I., Kao, C. F., Brunzman, B., Linneweh, B., & Rogers, D. (1996). Social corroboration and opinion extremity. *Journal of Experimental Social Psychology, 32*, 537-560.
- BBC News. (2004). *Rwanda: How the genocide happened*. Retrieved from <http://news.bbc.co.uk/1/hi/world/africa/1288230.stm> on 07/07/08.
- Bettencourt, B.A., Dorr, N., Charlton, K., & Hume, D.L. (2001). Status differences and in-group bias: a meta-analytic examination of the effects of status stability, status legitimacy and group permeability. *Psychological Bulletin, 24*, 520-542.
- Billig, M. (2006). A psychoanalytic discursive psychology: from consciousness to unconsciousness. *Discourse Studies, 8*, 17-24.
- Bliuc, A. M., Hartley, L., & McGarty, C. (2008). The role of wedge politics in legitimising and delegitimising political action. *Proceedings of the Annual Meeting of the Society of Australasian Social Psychologists (SASP)*.

- Bluic, A., McGarty, C., Reynolds, K., & Muntele, D. (2007). Opinion-based group membership as a predictor of commitment to political action. *European Journal of Social Psychology, 37*, 19-32.
- Blumer, H. (1958). Race prejudice as a sense of group position. *Pacific Sociological Review, 1*, 3-7.
- Bobo, L. D. (1999). Prejudice as group position: Microfoundations of a sociological approach to racism and race relations. *Journal of Social Issues, 55*, 445-472.
- Bornstein, G., Crum, L., Wittenbraker, J., Harring, K., Insko, C. A., & Thibaut, J. (1983). On the measurement of social orientations in the minimal group paradigm. *European Journal of Social Psychology, 13*, 321-350.
- Bourhis, R. Y., Sachdev, I., & Gagnon, A. (1994). Inter-group research with the Tajfel matrices: Methodological notes. In M. P. Zanna & J. M. Olson (Eds.), *The Social Psychology of Prejudice: The Ontario Symposium* (vol. 7, pp. 209-232). Hillsdale, NJ: Lawrence Erlbaum.
- Breed, W. & Ktsanes, T. (1961). Pluralistic ignorance in the process of opinion formation. *Public Opinion Quarterly, 25*, 382-392.
- Brewer, M. B. (1999). The psychology of prejudice: in-group love or out-group hate? *Journal of Social Issues, 55*, 429-444.
- Brown, R. (2000). Social identity theory: past achievements, current problems and future challenges. *European Journal of Social Psychology, 30*, 745-778.
- Buhl, T. (1999). Positive-negative asymmetry in social discrimination: Meta-analytic evidence. *Group Processes and Inter-group Relations, 2*, 51-58.
- Byatt, A. S. (1991). *Possession: A Romance*. London: Vintage.
- Cameron, J. E. (2004). The three-factor model of identity. *Self and Identity, 3*, 239-262.

- Campbell, D. T. (1965). Ethnocentric and other altruistic motives. In D. Levine (Ed.), *Nebraska symposium on motivation* (pp.283-301). Lincoln: University of Nebraska Press.
- Caporael, L. R. (2001). Parts and wholes: The evolutionary importance of groups. In C. Sedikides & M. Brewer (Eds.), *Individual self, relational self, collective self* (pp. 241-258). Philadelphia: Psychology Press.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*, 752-756.
- Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olsen, & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (vol. 5, pp. 3-39). Hillsdale: Lawrence Erlbaum.
- Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 73-96). New York: Guilford Press.
- Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and re-evaluation of the role of norms in human behaviour. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (vol. 24, pp. 201-233). San Diego, CA: Academic Press.
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*, 1015-1026.
- Coch, L. C. & French, J. R. F. (1948). Overcoming resistance to change. *Human Relations*, *1*, 512-532
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37-46.

- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioural sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.
- Correll, J. & Park, B. (2005). A model of the in-group as a social resource. *Personality and Social Psychology Review*, 9, 4, 341-359.
- Dagevos, J., & Gijsberts, M. (2007). *Jaarrapport integratie 2007 [Annual report integration 2007]*. The Hague, Netherlands: Sociaal Cultureel Planbureau.
- Dambrun, M. & Guimond, S. (2001). Relative deprivation theory and hostility towards North Africans. *International Review of Social Psychology*, 14, 57-89.
- Dambrun, M., Taylor, D. M., McDonald, D. A., Crush, J., & Méot, A. (2006). The Relative deprivation-gratification continuum and the attitudes of South Africans toward immigrants: A test of the V-curve hypothesis. *Journal of Personality and Social Psychology*, 91, 1032-1044.
- De Dreu, C. K. W., Nijstad, B. A., & van Knippenberg, D. (2008). Motivated information processing in group judgment and decision making. *Personality and Social Psychology Review*, 12, 22-49.
- Descartes, R. (1656/1960). *Discourse on method and meditations*. New York: The Liberal Arts Press.
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgement. *Journal of Abnormal and Social Psychology*, 629-636.
- Diener, E. (1976). Effects of prior destructive behaviour, anonymity, and group presence on deindividuation and aggression. *Journal of Personality and Social Psychology*, 33, 497-507.
- Diener, E. (1979). Deindividuation, self-awareness, and disinhibition. *Journal of Personality and Social Psychology*, 37, 1160-1171.

- Diener, E. (1980). Deindividuation: The absence of self-awareness and self-regulation in group members. In P. B. Paulus (Ed.), *Psychology of group influence* (pp. 209-242). Hillsdale, NJ: Erlbaum.
- Devine, P. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 1, 5-18.
- Doosje, B., Haslam, S. A., Spears, R., Oakes, P. J., & Koomen, W. (1998). The effect of comparative context on central tendency and variability judgements and the evaluation of group characteristics. *European Journal of Social Psychology*, 28, 173-184.
- Drury, J. & Reicher, S. D. (2000). Collective action and psychological change: The emergence of new social identities. *British Journal of Social Psychology*, 39, 579-604.
- Eagly, A. H. & Chaiken, S. (1993). *The Psychology of Attitudes*. Orlando: Harcourt Brace Jovanovich College Publishers.
- Ellemers, N. (1993). The influence of socio-structural variables on identity enhancement strategies. In W. Stroebe & M. Hewstone (Eds.). *European Review of Social Psychology* (vol. 4, pp. 27-57). Chichester, England: Wiley.
- Ellemers, N., Doosje, B., & Spears, R. (2004). Sources of respect: the effects of being liked by in-groups and out-groups. *European Journal of Social Psychology*, 34, 2, 155-172.
- Ellemers, N., Kortekaas, P., & Ouwerkerk, J. (1999). Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European Journal of Social Psychology*, 29, 371-389.
- Erdentuğ, A. & Colombijn, F. (2002). *Urban ethnic encounters: The spatial consequences*. London: Routledge.
- Festinger, L. (1950). Informal social communication. *Psychological Review*, 57, 271-282.

- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Festinger, L., Pepitone, A., & Newcomb, T. (1952). Some consequences of deindividuation in a group. *Journal of Abnormal and Social Psychology*, 47, 382-389.
- Festinger, L., & Thibaut, J. (1951). Interpersonal communication in small groups. *Journal of Abnormal Social Psychology*, 46, 92-09.
- Fielding, K.S., & Hogg, M.A. (1997). Social identity, self-categorization, and leadership: A field study of small interactive groups. *Group Dynamics: Theory, Research, and Practice*, 1, 39-51.
- Fields, J. M. & Schuman, H. (1976). Public beliefs about the beliefs of the public. *Public Opinion Quarterly*, 40, 427-448.
- Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Fleming, M. A. & Petty, R. E. (2000). Identity and persuasion: An elaboration likelihood approach. In D. J. Terry & M. A. Hogg (Eds.), *Attitudes, behaviour, and social context: The role of norms and group membership* (pp. 171-199). Mahwah, NJ: Lawrence Erlbaum Associates.
- Fortuyn, P. (2002). *De puinhopen van acht jaar paars [The mess of eight years of purple government]*. Rotterdam: Speakers Academy.
- Fraser, C., Gouge, C., & Billig, M. (1971). Risky shifts, cautious shifts and group polarisation. *European Journal of Social Psychology*, 1, 7-29.
- French, J. R. P. & Raven, B. (1959). The bases of social power. In A. Arbor (Ed.), *Studies in Social Behaviour* (pp. 150-167). Michigan: Michigan Institute for Social Research, University of Michigan.

- Gaertner, L., Iuzzini, J., Guerrero, Witt, M., & Oriña, M. M. (2006). Us without them: Evidence for an intragroup origin of positive ingroup regard. *Journal of Personality and Social Psychology, 90*, 426–439.
- Gaertner, L. & Schopler, J. (1998). Perceived in-group entitativity and inter-group bias: An interconnection of self and others. *European Journal of Social Psychology, 28*, 963–980.
- Goldhagen, D. J. (1997). *Hitler's willing executioners: Ordinary Germans and the Holocaust*. London: Abacus.
- Greenwald, A. G. (1968). Cognitive learning, cognitive response to persuasion, and attitude change. In A. G. Greenwald, T. M. Brock, & T. M. Ostrum, (Eds.), *Psychological foundations of attitudes* (pp. 147-169). London: Academic Press.
- Guimond, S. & Dambrun, M. (2002). When prosperity breeds inter-group hostility: the effects of relative deprivation and relative gratification on prejudice. *Personality and Social Psychology Bulletin, 28*, 900-912.
- Gurr, T. R. (1970). *Why men rebel*. Princeton: Princeton University Press.
- Hardin, C. D. & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: The interpersonal context* (vol. 3, pp. 28-80). New York: Guilford.
- Harding, A. (2003). *Rwanda tries to heal its wounds*. Retrieved from http://news.bbc.co.uk/1/hi/programmes/from_our_own_correspondent/311489.stm on 07/07/08.
- Haslam, S. A. (1997). Stereotyping and social influence: Foundations of stereotype consensus. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 119-143). Oxford: Blackwell.
- Haslam, S. A. (2004). *Psychology in organizations: The social identity approach* (2nd Ed.). London: Sage Publications.

- Haslam, S.A., Eggins, R.A., & Reynolds, K. J. (2003). The ASPIRe model: Actualizing Social and Personal Identity Resources to enhance organizational outcomes. *Journal of Occupational and Organisational Psychology*, 76, 83-113.
- Haslam, S. A. & McGarty, C. (2001). A 100 years of certitude? Social psychology, the experimental method and the management of scientific uncertainty. *British Journal of Social Psychology*, 40, 1-21.
- Haslam, S. A., Oakes, P. J., Reynolds, K. J., & Turner, J. C. (1999). Social identity salience and the emergence of stereotype consensus. *Personality and Social Psychology Bulletin*, 25, 7, 809-818.
- Haslam, S. A. & Turner, J. C. (1992). Context-dependent variation in social stereotyping II : The relationship between frame of reference, self-categorization and accentuation. *European Journal of Social Psychology*, 22, 251-277.
- Haslam, S. A., Turner, J. C., Oakes, P. J., McGarty, C., & Reynolds, K. J. (1998). The group as a basis for emergent stereotype consensus. *European Review of Social Psychology*, 28, 755-776.
- Haslam, S. A., Turner, J. C., Oakes, P. J., Reynolds, K. J., Eggins, R. A., Nolan, M., & Tweedie, J. (1998). When do stereotypes become really consensual? Investigating the group-based dynamics of the consensualisation process. *European Journal of Social Psychology*, 28, 755-776.
- Haslam, S. A., Turner, J. C., Oakes, P. J., Reynolds, K. J. , & Doosje, B. (2002). From personal pictures in the head to collective tools in the world: How shared stereotypes allow groups to represent and change social reality. In C. McGarty, V. Y. Yzerbyt, & R. Spears (Eds.), *Stereotypes as explanation: The formation of meaningful beliefs about social groups*. Cambridge: CUP.
- Hastie, R. & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review*, 112, 494-508.

- Heckhausen, H., & Kuhl, J. (1985). From wishes to action: The dead ends and short cuts on the long way to action. In M. Frese & J. Sabini (Eds.), *Goal-directed behaviour: The concept of action in psychology* (pp. 134–160). Hillsdale, NJ: Erlbaum.
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53*, 575-604.
- Hinkle, S. & Brown, R. (1990). Inter-group comparisons and social identity: some links and lacunae. In D. Abrams and M. Hogg (Eds.), *Social identity theory: Construction and critical advances* (pp. 48-70). London: Harvester Wheatsheaf.
- Hobsbawm, E. (1994). *Age of extremes: The short 20th century, 1914-1991*. New York: Pantheon.
- Hogg, M. A., Coopershaw, L., & Holzworth, D. W. (1993). Group prototypicality and depersonalised attraction in small interactive groups. *Personality and Social Psychology Bulletin, 19*, 452-465.
- Hogg, M. A., & Hains, S. C. (1998). Friendship and group identification: A new look at the role of cohesiveness in groupthink. *European Journal of Social Psychology, 28*, 323-341.
- Hogg, M. A. & Smith, J. R. (2007). Attitudes in social context: A social identity perspective. *European Review of Social Psychology, 18*, 89-131.
- Hogg, M. A. & Turner, J. C. (1987). Inter-group behaviour, self-stereotyping and the salience of social categories. *British Journal of Social Psychology, 26*, 325-340.
- Homans, G. C. (1961). *Social behaviour: Its elementary forms*. London: Routledge Paul.
- Inzlicht, M. & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 1*, 365-371.
- Isenberg, D. J. (1986). Group polarisation: A critical review and meta-analysis. *Journal of Personality and Social Psychology, 50*, 1141–1151.

- Jackson, L. (2006). Riot and revenge. *4 Corners, Australian Broadcasting Corporation*.
- Retrieved from <http://www.abc.net.au/4corners/content/2006/s1588360.htm> on 23/11/07.
- Jamieson, J. P. & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology*, *93*, 4, 544-564.
- Jost, J. T. & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consensus. *British Journal of Social Psychology*, *33*, 1-27.
- Judd, C. M., McClelland, G. H., & Culhane, S. C. (1995). Data analysis: Continuing issues in the everyday analysis of psychological data. *Annual Review of Psychology*, *46*, 433-465.
- Kashima, Y. (2000). Maintaining cultural stereotypes in the serial reproduction of narratives. *Personality and Social Psychology Bulletin*, *26*, 594-604.
- Kelley, H. H. & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. New York: Wiley.
- Kelman, H. C. (1961). Processes of opinion change. *Public Opinion Quarterly*, *25*, 57-78.
- Krull, J. L. & MacKinnon, D. P. (1999). Multilevel mediation modelling in group-based intervention studies. *Evaluation Review*, *23*, 418-444.
- Kwok, W. W., Wright, B., & Kashima, Y. (2007). Constructing inter-group relationships in social communication. *Japanese Psychological Research*, *49*, 121-135.
- Lamm, H. & Myers, D. G. (1978). Group-induced polarisation of attitudes and behaviour. *Advances in Experimental Social Psychology*, *11*, 145-195.
- Latané, B. & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.
- Le Bon, G. (1895/1947). *The crowd: A study of the popular mind*. London: Ernest Benn.
- Lewin, K. (1947). Frontiers in group dynamics: Concept, method and reality in social science; social equilibria and social change. *Human Relations*, *7*, 5-41.

- Lewin, K. (1953). Studies in group decision. In D. Cartwright & A. Zander (Eds.). *Group Dynamics: Research and Theory* (pp. 287-301). New York: Row, Peterson & Co.
- Liska, A. E. (1984). A critical examination of the causal structure of the Fishbein/Ajzen attitude-behaviour model. *Social Psychology Quarterly*, 47, 61-74.
- Lott, A. J. & Lott, B. J. (1965). Group cohesiveness as interpersonal attraction. *Psychological Bulletin*, 64, 259-309.
- Luus, E. C. A. & Wells, G. L. (1994). The malleability of eyewitness confidence: co-witness and perseverance effects. *Journal of Applied Psychology*, 79, 714-723.
- Lyons, A. & Kashima, Y. (2003). How are stereotypes maintained through communication? The influence of stereotype sharedness. *Journal of Personality and Social Psychology*, 85, 989-1005.
- Mackie, D. M. (1986). Social identification effects in group polarisation. *Journal of Personality and Social Psychology*, 50, 720 - 728.
- Mackie, D. M. & Cooper, J. (1984). Attitude polarisation: The effects of group membership. *Journal of Personality and Social Psychology*, 46, 575-585.
- Mackie, D. M., Gastardo-Conaco, M. C., & Skelly, J. J. (1992). Knowledge of the advocated position and the processing of in-group and out-group persuasive messages. *Personality and Social Psychology Bulletin*, 18, 145-151.
- Mackie, D. M., Worth, L. T., & Asuncion, A. G. (1990). Processing of persuasive in-group messages. *Journal of Personality and Social Psychology*, 58, 812-822.
- Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-confirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, 42, 236-243.
- Marx, D. M. & Stapel, D. A. (2006). Understanding stereotype lift: On the role of the social self. *Social Cognition*, 24, 776-791.

- Maxwell, K. (2002). Friends: The role of peer influence across adolescent risk behaviours. *Journal of Youth and Adolescence*, *31*, 267-277.
- McGarty, C. & Penny, R. E. C. (1988) Categorization, accentuation and social judgement. *British Journal of Social Psychology*, *27*, 147-157.
- McGarty, C., Yzerbyt, V. Y., & Spears, R. (2002). Social, cultural and cognitive factors in stereotype formation. In C. McGarty, V. Y. Yzerbyt, & R. Spears (Eds.), *Stereotypes as explanations: The formation of meaningful beliefs about social groups* (pp. 1-15). Cambridge: Cambridge University Press.
- McGarty, C., Turner, J. C., Hogg, M. A., David, B., & Wetherell, W. S. (1992). Group polarisation as conformity to the prototypical group member. *British Journal of Social Psychology*, *31*, 1–20.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). London: Sage.
- Miller, D. T. & McFarland, C. (1987). Pluralistic ignorance: when similarity is interpreted as dissimilarity. *Journal of Personality and Social Psychology*, *53*, 298-305.
- Miller, D. T. & McFarland, C. (1991). When social comparison goes awry: The case of pluralistic ignorance. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research*. Hillsdale, NJ: Lawrence Erlbaum.
- Moscovici, S. & Zavalloni, M. (1969). The group as a polariser of attitudes. *Journal of Personality and Social Psychology*, *12*, 125-135.
- Mummendey, A. & Otten, S. (1998). Positive-negative asymmetry in social discrimination. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology*, *9*, 107-143.
- Mummendey, A., Simon, B., Dietze, C., Grünert, M, Haeger, G., Kessler, S., Lettgen, S., & Schäferhoff, S. (1992). Categorisation is not enough: Inter-group

- discrimination in negative outcome allocation. *Journal of Experimental Social Psychology*, 28, 125-44.
- Myers, D. G. (1982). Polarising effects of social interaction. In H. Brandstatter, J. H. Davis & G. Stocker-Kreichgauer (Eds.), *Group decision making* (pp. 125-157). London: Academic Press.
- Myers, D. G. & Bishop, G. D. (1970). Discussion effects on racial attitudes. *Science*, 169, 3947, 778-779.
- Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative social influence is underdetected. *Personality and Social Psychology Bulletin*, 34, 7, 9132-923.
- Núñez, R. E. (1997). Eating soup with chopsticks: Dogmas, difficulties and alternatives in the study of conscious experience. *Journal of Consciousness Studies*, 4, 143-166.
- Oakes, P.J., Haslam, S.A, Morrison, B. & Grace, D. (1995). Becoming an ingroup: Re-examining the impact of familiarity on perceptions of group homogeneity. *Social Psychology Quarterly*, 58, 52-61.
- Pettigrew, T. F. & Tropp, L. R. (2006) A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90, 751-783.
- Petty, R. E., Briñol, P., & Tormala, Z. L. (2002). Thought confidence as a determinant of persuasion: The self-validation hypothesis. *Journal of Personality and Social Psychology*, 82, 722-741.
- Petty, R. E. & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IA: Brown.
- Petty, R. E. & Cacioppo, J. T. (1986a). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.
- Petty, R. E. & Cacioppo, J. T. (1986b). The elaboration likelihood model of

- persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, 19, 123-205.
- Poortvliert, P. M., Janssen, O., Van Yperen, N. W., & Van de Vliert, E. (2007). Achievement goals and interpersonal behavior: How mastery and performance goals shape information exchange. *Personality and Social Psychology Bulletin*, 33, 1435-1447.
- Postmes, Baray, Haslam, Morton, & Swaab, (2006). The dynamics of personal and social identity formation. In T. Postmes, & J. Jetten (Eds.), *Individuality and the group: Advances in social identity* (pp. 215-236). London: Sage.
- Postmes, T., Haslam, S. A., & Swaab, R. (2005). Social influence in small groups: An interaction model of social identity formation. *European Review of Social Psychology*, 16, 1-42.
- Postmes, T. & Smith, L. G. E. (in press). Why do the privileged resort to oppression? A look at some intra-group factors. *Journal of Social Issues* [accepted for publication October 2007].
- Postmes, T. & Spears, R. (1998). Deindividuation and anti-normative behaviour: A meta-analysis. *Psychological Bulletin*, 123, 238-259.
- Postmes, T., Spears, R., & Cihangir, S. (2001). Quality of decision making and group norms. *Journal of Personality and Social Psychology*, 80, 918-930.
- Postmes, T., Spears, R., & Lee, M. (2002). Inter-group differentiation in computer-mediated communication: effects of depersonalization. *Group Dynamics: Theory, Research and Practice*, 6, 3-16.
- Postmes, T., Spears, R., Lee, T., & Novak, R. (2005). Individuality and social influence in groups: Inductive and deductive routes to group identity. *Journal of Personality and Social Psychology*, 89, 747-763.

- Postmes, T., Spears, R., Sakhel, K., & de Groot, D. (2001). Social influence in computer-mediated communication: The effects of anonymity on group behaviour. *Personality and Social Psychology Bulletin*, 27, 1243-1254.
- Prentice, D. A. & Miller, D. T. (1993). Pluralistic ignorance and the perpetuation of social norms by unwitting actors. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (pp. 161-178). San Diego, CA: Academic Press.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage..
- Reicher, S. (1984). The St. Pauls's 'riot': An explanation of the limits of crowd action in terms of a social identity model. *European Journal of Social Psychology*, 14, 1-21
- Reicher, S. (1987). Crowd behaviour as social action. In J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell (Eds.), *Rediscovering the social group: A self-categorisation theory* (pp. 171-202). Oxford: Blackwell.
- Reicher, S. (1996). 'The Battle of Westminster': Developing the social identity model of crowd behaviour in order to explain the initiation and development of collective conflict. *European Journal of Social Psychology*, 26, 115-34
- Reicher, S. (1997a). Collective psychology and the psychology of the self. *BPS Social Section Newsletter*, 36, 3-15
- Reicher, S. (1997b). Social identity and social change: Rethinking the context of social psychology. In W. P. Robinson (Ed.), *Social groups and identities: Developing the legacy of Henri Tajfel* (pp. 317-336). London: Butterworth.
- Reicher, S., Haslam, S. A., & Hopkins, N. (2005). Social identity and the dynamics of leadership: Leaders and followers as collaborative agents in the transformation of social reality. *Leadership Quarterly*, 16, 547-568.

- Reicher, S., Hopkins, N., & Condor, S. (1997). Stereotype constructions as a strategy of influence. In R. Spears, P. J. Oakes, N. Ellemers and S. A. Haslam (Eds.), *The Social Psychology of Stereotyping and Group Life* (pp. 94-117). Oxford: Blackwell.
- Reicher, S., Spears, R., & Postmes, T. (1995). A social identity model of deindividuation phenomena. *European Review of Social Psychology*, 6, 161-198.
- Reynolds, K. J., Oakes, P. J., Haslam, S. A., Nolan, M., & Dolnik, L. (2000). Responses to powerlessness: stereotypes as an instrument of social conflict. *Group Dynamics: Theory, Research and Practice*, 4, 275-290.
- Reynolds, K. J., Turner, J. C., & Haslam, S. A. (2000) When are we better than them and they worse than us? A closer look at social discrimination in positive and negative domains. *Journal of Personality and Social Psychology*, 78, 64-80.
- Rosenthal, R., & Rosnow, R. L. (1985). *Contrast analyses: Focused comparisons in the analysis of variance*. Cambridge, England: Cambridge University Press.
- Rudé, G. (1964). *The crowd in history: A study of popular disturbances in France and England, 1730-1848*. New York: Wiley.
- Ryan, M. K., Iyer, A., Hersby, M. D., & Kulich, C. (in prep.) *Responding to negative identities: The independent role of the three factors of identity*.
- Sassenberg, K., Kessler, T., & Mummendey, A. (2003). Less negative = more positive? Social discrimination as avoidance or approach. *Journal of Experimental Social Psychology*, 39, 48-58.
- Sassenberg, K. & Postmes, T. (2002) Cognitive and strategic processes in small groups: Effects of anonymity of the self and anonymity of the group on social influence. *British Journal of Social Psychology*, 41, 463-480.
- Searle, J. R. (1995). *The construction of social reality*. London: Free press.

- Sechrist, G. B. & Stangor, C. (2001). Perceived consensus influences inter-group behaviour and stereotype accessibility. *Journal of Personality and Social Psychology*, 80, 645-654.
- Shapiro, J. R. & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, 11, 107-130.
- Shaw, M. W. (1976). *Group dynamics: The psychology of small group behavior*. London: McGraw-Hill
- Sherif, M. (1935). A study of some social factors in perception. *Archives of Psychology*, 27, 1-60.
- Sherif, M. (1956). Experiments in group conflict. *Scientific American*, 195, 5, 54-58.
- Sherif, M. (1967). *Group conflict and cooperation: Their social psychology*. London: Routledge and Kegan Paul.
- Sherif, M. & Hovland, C. I. (1961). *Social judgement: Assimilation and contrast effects in communication and attitude change*. New Haven: Yale University Press.
- Sherif, M. & Sherif, C. W. (1956). *An outline of social psychology* (2nd ed.). New York: Harper & Brothers.
- Sidanius, J., & Pratto, F. (1999). *Social dominance: An inter-group theory of social hierarchy and oppression*. New York, NY, US: Cambridge University Press.
- Smith, J. R. & Louis, W. R. (in press). Do as we say and as we do: The interplay of descriptive and injunctive group norms in the attitude-behaviour relationship. *British Journal of Social Psychology*.
- Smith, L. G. E. & Postmes, T. (in press). Intra-group interaction and the development of norms which promote inter-group hostility. *European Journal of Social Psychology* [accepted for publication August 2007].

- Snijders, T. A.B. & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modelling*. London: Sage.
- Spears, R., Ellemers, N., Doosje, B., & Branscombe, N. R. (2006). The individual within the group: Respect! In T. Postmes & J. Jetten (Eds.), *Individuality and the Group: Advances in Social Identity* (pp. 175-195). London: Sage.
- Spears, R., Lea, M., Corneliusen, R. A., Postmes, T., & Ter Haar, W. (2002). Computer-mediated communication as a channel for social resistance: The strategic side of SIDE. *Small Group Research*, 33, 555-574.
- Spears, R., Lea, M., & Lee, S. (1990). De-individuation and group polarization in computer-mediated communication. *British Journal of Social Psychology*, 29, 121-134.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-28.
- Stangor, C., Sechrist, G. B., & Jost, J. T. (2001). Changing racial beliefs by providing consensus information. *Personality and Social Psychology Bulletin*, 27, 4, 486-496.
- Stasser, G. & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, 53, 81-93.
- Staub, E. (1996). Cultural-societal roots of violence: The examples of genocidal violence and of contemporary youth violence in the United States. *American Psychologist*, 51, 117-132.
- Staub, E. (1997). *The roots of evil: The origins of genocide and other group violence*. Cambridge: Cambridge University Press.

- Staub, E., & Rosenthal, L. H. (1994). Mob violence: Cultural-societal sources, instigators, group processes, and participants. In L. D. Eron, J. H. Gentsy, & P. Schlegel (Eds.), *Reason to hope: A psychosocial perspective on violence and youth* (pp. 281-313). Washington, DC: American Psychological Association.
- Steele, C. M. (1997) A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613-629.
- Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 64, 885-896.
- Sternberg, R. J. (1990). *Wisdom: Its nature, origins, and development*. Cambridge: Cambridge University Press.
- Stoner, J. A. F. (1968). Risky and cautious shifts in group decisions: The influence of widely held values. *Journal of Experimental Social Psychology*, 4, 442-459.
- Stott, C. & Drury, J. (2000). Crowds, context and identity: Dynamic categorization processes in the 'poll tax riot'. *Human Relations*, 53, 247-273.
- Stott, C. & Drury, J. (2004). The importance of social structure and social interaction in stereotype consensus and content: Is the whole greater than the sum of its parts? *European Journal of Social Psychology*, 34, 11-2
- Stott, C., & Reicher, S. (1998). Crowd action as inter-group process: Introducing the police perspective. *European Journal of Social Psychology*, 28, 509-529
- Struch, N. & Schwartz, S. H. (1989). Inter-group aggression: Its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology*, 56, 364-373.
- Swaab, R., Postmes, T., van Beest, I., & Spears, R. (2007). Shared cognition as a product of, and precursor to, shared identity in negotiations. *Personality and Social Psychology Bulletin*, 33, 187-199.

- Tajfel, H. (1972). Experiments in a vacuum. In H. Tajfel & Israel, J. (Eds.), *The context of social psychology: A critical assessment* (pp. 69-119). London: Academic Press in co-operation with the European Association of Experimental Social Psychology.
- Tajfel, H. (1978). Social categorization, social identity and social comparison. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 61-76). London: Academic Press.
- Tajfel, H. (1981). Social stereotypes and social groups. In J. C. Turner & H. Giles (Eds.), *Inter-group Behaviour* (pp. 144–167). Oxford: Blackwell.
- Tajfel, H., Billig, M. G., Bundy, R. F., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1, 149-177.
- Tajfel, H. & Turner, J. C. (1979). An integrative theory of inter-group conflict. In W.G. Austin & S. Worchel (Eds.), *The Social psychology of inter-group relations* (pp. 33-47). Monterey, CA: Brooks/Cole.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of inter-group behaviour. In S. Worchel & W. G. Austin (Eds.), *Psychology of inter-group relations* (pp. 7–24). Chicago: Nelson Hall.
- Terry, D. J., & Hogg, M. A. (1996). Group norms and the attitude-behaviour relationship. A role for group identification. *Personality and Social Psychology Bulletin*, 22, 776-793.
- Thibaut, J. W. & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.
- Thomas, E., & McGarty, C. (in press). The role of efficacy and moral outrage norms in creating the potential for international development activism through group-based interaction. *British Journal of Social Psychology*.
- Thompson, E. P. (1971). The moral economy of the English crowd in the eighteenth century. *Past and Present*, 50, 76-136.

- Thompson, M. S., Judd, C. M., & Park, B. (2000). The consequences of communicating social stereotypes. *Journal of Experimental Social Psychology, 36*, 567-599.
- Tilly, C., Tilly, L., & Tilly, R. (1975). *The rebellious century: 1830-1930*. Cambridge, MA: Harvard University Press.
- Turner, J. C. (1982). Towards a cognitive redefinition of the social group. In H. Tajfel (Ed.), *Social identity and inter-group relations* (pp. 15-40). Cambridge: Cambridge University Press.
- Turner, J. C. (1983). The measurement of social orientations in the minimal group paradigm – some comments. *European Journal of Social Psychology, 13*, 351-367.
- Turner, J. C. (1985). Social categorisation and the self-concept: A social cognitive theory of group behaviour. In E.J. Lawler (Ed.), *Advances in Group Processes* (vol. 2, pp. 77-122). Greenwich, CT: JAI Press.
- Turner, J. C. (1987). Introducing the problem: Individual and group. In J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell, (Eds.), *Rediscovering the social Group: A self-categorisation theory* (pp. 1-41). Oxford: Blackwell.
- Turner, J. C. (1987b). A self-categorisation theory. In J. C. Turner, M. A., Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell (Eds.), *Rediscovering the social group: A self-categorisation theory* (pp. 42–67). Oxford: Blackwell.
- Turner, J. C. (2005). Explaining the nature of power: A three-process theory. *European Journal of Social Psychology, 35*, 1-22.
- Turner, J. C. (1991). *Social influence*. Milton Keynes: Open University Press.
- Turner, J. C. (1999). Some current issues in research on social identity and self-categorisation theories. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 68-89). Oxford, England: Basil Blackwell.

- Turner, J. C., Hogg M. A., Oakes P. J., Reicher, S., & Wetherell, M. S. (1987).
Rediscovering the social group: A self-categorisation theory. Oxford: Blackwell.
- Turner, J.C. & Oakes, P.J. (1997). The socially structured mind. In C. McGarty and S.A. Haslam (Eds.), *The Message of Social Psychology* (pp. 355-373). Oxford: Blackwell.
- Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, 20, 454-463.
- Turner, J.C., Reynolds, K. J., Haslam, S.A., Veenstra, K. E. (2006). Reconceptualizing personality: Producing individuality through defining the personal self. In T. Postmes & J. Jetten (Eds.), *Individuality and the Group: Advances in Social Identity* (pp. 11-36). London: Sage.
- Turner, J. C., Wetherell, M. S., & Hogg, M. A. (1989). Referent informational influence and group polarisation. *British Journal of Social Psychology*, 28, 135–147.
- Turner, M. E., Pratkanis, A. R., & Samuels, T. (2003). Identity metamorphosis and groupthink prevention: examining Intel’s departure from the DRAM industry. In S. A. Haslam, D. van Knippenberg, M. J. Platow, & N. Ellemers (Eds.), *Social identity at work: Developing theory for organisational practice* (pp. 117-136). Philadelphia, PA: Psychology Press.
- Turner, R. H. & Killian, L. M. (1972). *Collective behaviour*. Englewood Cliffs: Prentice-Hall Inc.
- Van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin*, 134, 504-535.
- Van Zomeren, M., Spears, R., Leach, C., & Fischer, A. H. (2004). Put your money where your mouth is! Explaining social action tendencies through group-based anger and group efficacy. *Journal of Personality and Social Psychology*, 87, 649-664.

- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.
- Vinokur, A. & Burnstein, E. (1974). Effects of partially shared persuasive arguments on group induced shifts: A group problem-solving approach. *Journal of Personality and Social Psychology*, *29*, 305-315.
- Wallach, M. A. & Kogan, N. (1965). The roles of information, discussion, and consensus in group risk taking. *Journal of Experimental Social Psychology*, *1*, 1-19.
- Walton, G. M. & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*, 456–467.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063-1070.
- Wetherell, M. S. (1987). Social identity and group polarization. In J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell, *Rediscovering the social group: A self-categorization theory* (pp. 142-170). Oxford, England: Blackwell.
- Wetherell, M. S. (2007). A step too far: Discursive psychology, linguistic ethnography and questions of identity. *Journal of Sociolinguistics*, *11*, 661-681.
- Wilder, D. A. (1990). Some determinants of the persuasive power of in-groups and out-groups: Organization of information and attribution of independence. *Journal of Personality and Social Psychology*, *59*, 1202 – 1213.
- Wilkinson, L. & The APA Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594-604.
- Wittenbaum, G. M., Hubbell, A. P., & Zuckerman, C. (1999). Mutual enhancement: Toward an understanding of the collective preference for shared information. *Journal of Personality and Social Psychology*, *77*, 967-978.

- Wright, S.C., Taylor, D.M., & Moghaddam, F.M. (1990). Responding to membership in a disadvantaged group: from acceptance to collective protest. *Journal of Personality and Social Psychology*, 58, 994-1003.
- Zander, A. (1979). The psychology of group processes. *Annual Reviews in Psychology*, 30, 417-51.
- Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order vs. deindividuation, impulse, and chaos. In W. J. Arnold & D. Levine (Eds.), *Nebraska symposium on motivation* (pp. 237-307). Lincoln: University of Nebraska Press.

Appendix 3.1

Definitions of codes and example quotes, Studies 3.1 and 3.3

Research Question		Code	Definition	Example Quotes
1. Were participants more likely to stereotype the out-group after discussion compared to individual reflection?	1.1	ST	Stereotyped the out-group	“Supply us with food such as Chinese shops, Indians and kebab shops”
	1.2	RST	Resisted stereotyping the out-group	“Everyone that comes from outside the United Kingdom not just poor people in war”
2. What was the affective tone of the description?	2.1	HOS	Hostile	“Some immigrants are merely here to deal drugs and rake money off the state” “Many...are quite ungrateful because coming to this country is a privilege not a right”.
	2.2	SYM	Sympathetic	“They must be scared to come over from a foreign country, to be so scared to have to leave home...just ordinary people like you and me who are desperate”
	2.3	AMB	Ambivalent	
3. What was the content of the stereotype of the out-group?	3.1	WELF	Immigrants take advantage of the welfare system in the UK	“Some do nothing and live in paid-for houses whilst supported by benefits”
	3.2	POOR	Immigrants are poor	“They usually come from poorer countries”.
	3.3	BETL	Immigrants come to the UK to seek a better quality of life	“They come over from their home for a better life, or a chance of a better life”
	3.4	INTREL	There are hostile inter-group relations between immigrants and UK citizens.	“Their relationships to the British people are generally fairly edgy, especially in urban areas where clusters of different ethnicities are found. This

			causes conflict.”
3.5	EMP	Immigrants come to the UK to seek employment.	“Immigrants normally come into Britain in search of work”
3.6	AS	Immigrants are Asian.	“Mostly Asians and people from the middle east”
3.7	CIT	Immigrants live in cities.	“Mainly move to major cities”
3.8	FLE	Immigrants are refugees from their country of origin.	“They come from other countries usually suffering from war or poverty”
3.9	ANY	An immigrant could be anyone who has moved from his or her country of origin to another country.	“They come from all walks of life: every country, every race and every ethnic background”
3.10	THIR	Immigrants originate from third world countries.	“They come from 3rd world countries”
3.11	REL	Immigrants are religious.	“Private and enclosed people with strong religious beliefs”
3.12	ETH	Immigrants have a different ethnic origin to UK citizens.	“They are different coloured people”
3.13	ILL	Immigrants can be in the UK illegally.	“Could be illegal but have gained access to our country”
3.14	LOW	Immigrants do low paid jobs.	“Come into the country and help us by doing some of the labour work, low paid jobs”
3.15	BEN	Immigration is beneficial to Britain.	“They fill many jobs in the UK from doctors/medical staff to bin men”
3.16	EXP	Immigrants are exploited in Britain.	“Work long hours and live in squalid conditions with terrible pay”
3.17	CUL	Culture clash	“...They don’t respect our rules and they are trying to make everyone go under their rules in our country” “Usually intimidating to true citizens and have too much freedom to exploit their own beliefs. If they

				go to a different country, they should live by the ways of that country”
				“I think it’s unfair to grant the requests of some immigrants... when they do not fit in with the morals of British society”
	3.18	INTEG	Poor cultural integration	“Sometimes immigrants group together and form ‘ghettos’. “The relation to British people is very segregated as many people are discriminated against(wrongly or <i>rightly</i>)” “They should come and accept our ways as we are made to accept theirs”
	3.19	EQ	All people should be treated equally, immigrants and UK citizens alike	“Nobody should be more important than anyone else”
	3.20	TER	Terrorism	“Some terrorists sneak past the system and put British society at risk from terrorist attacks” “They are probably very friendly, apart from the terrorists”
4. Did participants formulate a plan to solve any problems immigration causes?	4.1	POS PLAN	Formed a plan for positive (benevolent) action	“Encourage multi-faith and multi-racial activities to bring people together”
	4.2	NEG PLAN	Formed a plan for negative (hostile) action.	“To combat immigrants we should limit the numbers let in to the country, tighter controls” “...not letting so many into the country or making them have certain qualifications before they can apply for jobs”
	4.3	NO PLAN	Did not form a plan.	

Appendix 5.1

Definitions of codes for responses to the question, “Why is it true that men are better than women at maths?” Study 5.1

Focus	Code	Definition	Example quotes	
1. Traits	1.1	SPAT	Men have better spatial awareness than women	“Men are considered better at spatial awareness – better able to visualise”
	1.2	SIN	Men are better than women at(focusing on) single tasks	“Men have better focus on one task”
	1.3	SAT	Women have shorter attention spans than men	“Women have shorter attention spans”
	1.4	LOG	Men are more logical/factual/rational than women	“Men are more logically minded”
	1.5	NUM	Men are more numerically intelligent than women	“Men are better with numbers”
	1.6	BET	Women are better (than men) at things other than maths	“Women are... better at other things”
	1.7	CRE	Women are more creative than men	“Women are more creative”
	1.8	EMO	Women are more emotional (and therefore more irrational) than men.	“In tests men don’t panic as much as women so are better at the methodological approach of maths” “Men do not think about feelings as much and so find it easier to think straightforwardly”
	1.9	VERB	Women are more verbally intelligent than men	“Women are better with words”
2. Behaviours	2.1	OTH	Women have more important, other things to think about/do	“Women have more important things to think about”
	2.2	ED	Men in particular prefer maths as an academic subject or are more	“Men are pushed harder in maths due to ‘manlier’ jobs being more maths based”

			likely to advance in their maths education than women.	“Traditionally men were better educated, women expected to stay at home”
	2.3	MULTI	Women multi-task.	“Women are better at multi-tasking”
	2.4	COMP	Men are more competitive than women	“Men are more competitive – like solving problems etc”
3. Roles	3.1	PROF	Maths is involved in male professions	“[Men] have more professional jobs like engineers which require advanced mathematical skills”
	3.2	MROL	There are famous male-role models who are good at maths.	
4. Theories	4.1	SOC	Societal/cultural cause e.g. gender stereotyping	“Perception of a male dominated world can lead to women lacking the chance to practice their mathematical skills” “Cultural assumption that men are better at maths than women”
	4.2	BRA	Physical cause, i.e., differences in male and female brain function	“Men and women use parts of the brain differently. Maybe their brains work more laterally” “The area of the brain that deals with maths might be more developed in men”
	4.3	EVO	Evolutionary cause	“Evolution - Men have had to use maths skills e.g., angles more than women”
5. Denial	5.1	EQ	Maths ability of men and women is equal	“Women aren’t actually worse it’s often seen like that”

Appendix 5.2

Definitions of codes for responses to the question, “Why is it not true that men are better than women at maths?” Study 5.1

Focus		Code	Definition	Example quotes
1. Traits	1.1	PHY	Men are more physical than women.	“Men are more physical”
	1.2	ATT	Women are more attentive than men.	“Women can concentrate better” “Women have a longer attention span” “On average, women tend to have longer concentration levels than men so tend to get better grades in maths”
	1.3	INT	Women are more (numerically) intelligent than men.	“Women are more intellectual”
	1.4	IMPRESS	Women mind less than men about impressing their friends so they do not mind being perceived as intellectual.	“Less interested in impressing their friends don’t mind being branded a geek”
	1.5	VERB	Women are more verbally intelligent than men.	“Women communicate better”
	1.6	THO	Women are more thorough than men.	“Women...are more precise in their work”
2. Behaviours	2.1	EX	Women do better in exams/at school than men.	“Girls do better at school” “Women get better exam results”
3. Roles	3.1	FEM	Women use maths professionally as well as men.	“There are more female maths teachers” “The number of women in maths related jobs in increasing” “In a family environment women tend to do shopping and bills more than men so have to

				use maths more often with budgeting and have more practice” “Carol Vorderman!”
4. Theories	2.3	ROL	There are famous female role models who are good at maths.	
	4.1	SOC	The gender difference is based on social expectation/stereotypes.	“Social prejudices & opportunities may encourage different sexes to go for different subjects etc” “Women have always been equally as good at maths but haven’t had the chance to use it due to a male dominated society”
	4.2	NUR	Intelligence is determined by nurture, not nature (i.e., environmental factors, not gender).	“Environment determines your intelligence not gender”
	4.3	DEMAT	Women mature/develop earlier than men.	“Girls mature quicker – more likely to try harder at school as they realise how important it is”
5. Denial	4.4	BIAS	Maths testing is biased.	“Test may be biased”
	5.1	EQ	Maths ability of men and women is equal.	“Girls do just as well as boys at maths at school” “Women are just as intelligent as men”
	5.2	DED	Deny that maths education favours men.	“Taught the same. No different teaching determined by sex.”
	5.3	DBRA	Deny that there is a physical difference between genders which affects mathematical ability.	“No biological reason – no difference in brain activity”. “In evolution terms women are now moving into male roles so the brain capacity would have to equalise”
	5.4	DEVO	Deny that men have an evolutionary advantage.	“There is no evolutionary advantage to men being better so they aren’t”

5.5	RST	Resistance to the stereotype.	“Can’t judge all men and women together”
5.6	IND	There is individual variation in mathematical ability, not group (gender) variation.	“It is down to the individual not gender of the person as to their maths ability.”
5.7	GEN	No genetic reason for gender difference in mathematical ability.	“Genetics do not discriminate between sexes”
