

**SELF-ESTEEM, SOCIAL COMPARISON
AND DISCRIMINATION: A
REAPPRAISAL AND DEVELOPMENT OF
TAJFEL'S SOCIAL IDENTITY THEORY**

Candidate's Name: Thomas Lycan Farsides

of: The London School of Economics and
Political Science, University of London

Thesis submitted
for the degree of: Doctor of Philosophy

ABSTRACT

Six main empirical studies are reported.

Study 1 employed the "thinkaloud" procedure in the minimal group paradigm (MGP) and concludes that social categorization is insufficient to cause social identity or intergroup discrimination, and that no theoretical explanation of minimal group behaviour is adequate to explain the variety of strategies employed within that paradigm.

Study 2 employed both "Tajfel matrices" and new "allocation grids" in the MGP and concludes that two distinct forms of intergroup discrimination need to be distinguished: one which maximizes in-group profit consistent with positive in-group distinctiveness, and another which maximizes positive in-group distinctiveness by accompanying in-group profit with out-group derogation.

In Study 3 subjects completed three sets of Tajfel matrices in the MGP: individually, in "sub-groups", then again individually. Mean intergroup discriminatory behaviour polarized and mean intergroup equitable behaviour depolarized between the first and the latter two conditions. The best account of the results was concluded to be a normative one.

Study 4 demonstrated that the self-esteem hypothesis within social identity theory (SIT) is best tested using a state measure of *specific* social identity contingent self-esteem and concludes that this hypothesis has to date been both inadequately formulated and inadequately tested.

Study 5 compared predictions from SIT with those from Tesser's self-evaluation maintenance model concerning the consequences of social comparison outcomes and concludes that a modified version of the former theory is best able to account for the results obtained at both group and individual levels of comparison.

Study 6 investigated a host of issues within SIT and concludes that the theory is too simplistic in respect of many of its key notions and propositions.

A general discussion argues that a modified version of SIT can be developed which improves on Tajfel's "original" social identity theory by more adequately specifying the processes by which group phenomena are manifest.

TABLE OF CONTENTS

	Page
Title Page	1
Abstract	2
Table of Contents	4
Acknowledgements	11
Bibliography (in Volume 2)	586

CHAPTERS

Title	Page
1: Introduction - Tajfel's social identity theory	12
2: Study 1 - Thinkaloud in the minimal group paradigm	58
3: Study 2 - The nature of intergroup discrimination	108
4: Study 3 - Developing norms of intergroup behaviour	169
5: Study 4 - The nature of self-esteem	216
6: Study 5 - The consequences of social comparison: Tesser's self-evaluation maintenance model versus Tajfel's social identity theory	255
7: Study 6 - The "loose-ends" questionnaire	316
8: Conclusion - Developing social identity theory	369

APPENDICES (in Volume 2)

Title	Page
1: The Tajfel matrices - and how to score them	421
2: An example Study 1 (Thinkaloud) response booklet	422
3: Transcripts from Study 1 (Thinkaloud)	428
4: Study 1 (Thinkaloud) content analysis coding instructions	472
5: Qualitative summaries of subjects' "overall behaviour" in Study 1 (Thinkaloud)	476
6: An example Study 2 (Discrimination) allocation grid	481
7: Subjects' responses to the post-task questionnaire in Study 2 (Discrimination)	482
8: Transcripts of group discussions from Study 3 (Norms)	490
9: Allocation behaviour in Study 3 (Norms)	525
10: Questionnaire used in Study 4 (Self-esteem)	550
11: Study 5 (Social comparison) scale practice sheet	564
12: Study 5 (Social comparison) bogus measures	565
13: Study 5 (Social comparison) post-manipulation measures	570
14: Study 6 (Loose ends) "Group membership" questionnaire	577

TABLES AND ILLUSTRATIONS

Title	Page
<i>Chapter 1: Tajfel's Social Identity Theory</i>	
1.1 The consequences of perceived in-group superiority	26
1.2 The consequences of perceived in-group inferiority	27
<i>Chapter 2: Thinkaloud in the minimal group paradigm</i>	
2.1 Mean pull- and indirect in-group favouritism scores	75
2.2 Interrater agreement coefficients	77
2.3 Subjects' self-reported allocation intentions	78
2.4 Subjects' self-reported allocation motivations	82
2.5 Subjects' overall strategies and individual pull-scores	88
<i>Chapter 3: The nature of intergroup discrimination</i>	
3.1 McClintock's (1988) social value vectors	117
3.2 Group and intergroup behaviour vectors	119
3.3 Mean pull- and indirect in-group favouritism scores	130
3.4 Mean grid scores	131
3.5 Diagrammatic representation of subjects' allocations, by subject numbers	133

Tables and illustrations (cont.)	Page
3.6 Dendogram showing clusters by IP and OP grid scores	135
3.7 Dendogram showing clusters by matrix scores	138
3.8 Mean matrix and (selected) grid scores by clusters shown in illustration 3.7	140
3.9 Mean matrix and grid scores by dominant self-reported strategies	143
3.10 Diagrammatic representation of subjects' allocations, by dominant strategies	146
 <i>Chapter 4: Developing norms of intergroup behaviour</i>	
4.1 Mean pull scores by condition	181
4.2 Mean F on FAV and MIP+MJP on MD pull-scores, and mean InPref scores, by condition	182
4.3 Polarization by strategy and sub-group	184
4.4 Differences in sub-group behavioural variation in conditions 1 and 3 by strategy and sub-group	186
 <i>Chapter 5: The nature of self-esteem</i>	
5.1 Mean self-esteem changes by scale-type and manipulation	238
5.2 Number of changes in self-esteem in predicted or unpredicted directions by scale-type and manipulation	243

Tables and illustrations (cont.) **Page**

Chapter 6: The consequences of social comparison

6.1	Tesser's self-evaluation maintenance model	258
6.2	A between-group version of Tesser's self-evaluation maintenance model	262
6.3	False feedback employed in the present study	267
6.4	Changes in self-esteem as a result of upward or downward comparison with a close or distant other	290
6.5	Self-esteem change as a result of in-group superiority or inferiority on relevant dimensions	292
6.6	Changes in self-esteem as a result of in-group inferiority or superiority to the out-group and individual inferiority or superiority to in-group others	294
6.7	Changes in self-esteem as a result of in-group inferiority or superiority following comparisons with in-group or out-group others	296

Chapter 7: The loose-ends questionnaire

7.1	Mean evaluation of various charitable groups by favoured charitable group	328
7.2	Mean evaluation of various political parties by favoured political party	330

Tables and illustrations (cont.)	Page
7.3 Mean evaluation according to percentage of charitable in-group's annual target reached, when 70% of the annual target was reached the previous year	332
7.4 Mean evaluation according to percentage of charitable in-group's annual target reached, when 140% of the annual target was reached the previous year	333
7.5 Mean evaluation according to political in-group's electoral performance, when it had been beaten by the winning party by 500 votes the previous year	335
7.6 Mean evaluation according to political in-group's electoral performance, when it had beaten its nearest rival party by 500 votes the previous year	337
7.7 Mean self-esteem levels for charitable group members according to use or non-use of intergroup discrimination, positive or negative in-group distinctiveness, and presence or absence of a conflict of values	340
7.8 Mean self-esteem levels for political party supporters according to positive or negative in-group distinctiveness and presence or absence of a conflict of values	342
7.9 Mean evaluations according to own sex's earnings relative to other sex's earnings for the same work	343

Tables and illustrations (cont.)**Page**

- 7.10 Levels of use of exit and/or intergroup discrimination by charitable group members according to positive or negative in-group distinctiveness, availability or non-availability of exit, and availability or non-availability of cognitive alternatives **345**
- 7.11 Levels of use of exit and/or intergroup discrimination by political party supporters according to positive or negative in-group distinctiveness, availability or non-availability of exit, and availability or non-availability of cognitive alternatives **346**

ACKNOWLEDGEMENTS

I would like to thank the ESRC, without whom I could not have pursued any post-graduate study. There are also many others who have assisted in the production of this thesis. An infallible memory is one of many talents I do not possess and I therefore apologize anyone whose contributions to this thesis I have forgotten to mention below. I am grateful nonetheless.

A very great thank you to each and every one of the following (in no intended order):- Bobbie, my wife, and Seth, my son, who provided and continue to provide joy in and meaning to my life (and much, much more); Mamie and Costas Farsides, for many things, especially my wife; all those who took part in the studies reported in this thesis, for - on the most part - cheerful participation; George Gaskell, my Ph.D. supervisor, for Ph.D. supervision, and other members of the Social Psychology Department at the LSE, particularly Jan Stockdale, Vanessa and Fiona, for various reasons; the faculty of Keele University Psychology Department, who made participants available to me, especially the very generous Angus Gellately; all social psychologists at Sussex University, who provided support, encouragement and friendship; all those who provided valuable feedback on sections of this thesis presented at various conferences; Marco Cinnerella, for unselfishly sharing references, ideas, and respondents with me; Caroline Kelly and Steve Reicher, for excellent examination of this thesis and for suggesting a number of minor (but definitely improving) changes; and all my family and friends, who continue to be these things.

In as much as it means anything, I would like to dedicate this thesis to all those who are the recipients of prejudice and discrimination, for whom social psychology should be doing much more.

CHAPTER 1: TAJFEL'S SOCIAL IDENTITY THEORY

CHAPTER OVERVIEW

After specifying the broad aims of this thesis, this chapter begins with a detailed exposition of Tajfel's social identity theory. This is followed by a close examination of the form and implications of what might be called the "sufficiency condition" claim of that theory, namely "that the mere perception of belonging to two distinct groups - that is, social categorization per se - is sufficient to trigger intergroup discrimination favouring the in-group" (Tajfel & Turner, 1979: 38). Two arguments central to much of the material to follow in this thesis are then introduced. The first is a claim that social identity theory is rather too individualistic to provide a "genuinely social" theory of group and intergroup behaviour, and the second involves distinguishing between individuals identifying themselves "as" group members in contrast to individuals "identifying with" groups and/or their members. Finally, details are given of the specific content of each of the chapters to follow.

Overall, the chapters to follow this one have two main aims. The first is to investigate the psychological and behavioural effects of mere social categorization upon those categorized, and also the processes by which such effects come about. The second is to investigate the psychological and behavioural effects of "non-mere" social categorization, and again the processes by which such effects come about. In each case attempts are made to specify where Tajfel's social identity theory is supported, and where that theory needs or might benefit from modification.

INTRODUCTION

The global aim of this thesis is explicit in its sub-title: to reappraise and develop Tajfel's social identity theory. The basic components of that theory are straightforward. Individuals sometimes think of themselves and act toward others in terms of characteristics and situations more or less unique to them as individuals. When they do this, people can be said to be operating in terms of their personal identities and they act toward others on an interpersonal level. However, individuals also sometimes think of themselves and act toward others in terms of characteristics and situations they share with other individuals because they belong to the same groups as them. When they do this, people can be said to be operating in terms of their social identities and they act toward others on an intergroup level. All individuals, whether operating in terms of their personal or their social identities, have a need for satisfactory (or positive) self-esteem. When social identities are operative, self-esteem becomes inextricably tied up with evaluations of the groups providing those social identities ("in-groups"). In-groups are evaluated in terms of whether they are superior or inferior to other groups ("out-groups"): if they are superior they are evaluated positively and if they are inferior they are evaluated negatively. Thus, membership of superior in-groups provides positive self-esteem and membership of inferior groups provides negative self-esteem. Finally, because of the universal need for positive self-esteem, people try to belong to superior groups and/or try to make the groups they are members of positively distinct, via intergroup discrimination if necessary.

This thesis has three principal aims. The first is to challenge certain interpretations of social identity theory and certain hypotheses ostensibly derived from it. For example, it is argued that despite sometimes persistent claims

to the contrary (see below for references), social identity theory does not justify the predictions and/or assertions: (i) that perceived social categorization inevitably leads to intergroup discrimination; (ii) that strength of social identity will positively correlate with extent of intergroup discrimination; (iii) that level of self-esteem will negatively correlate with subsequent intergroup discrimination; or, (iv) that extent of intergroup discrimination will positive correlate with subsequent levels of self-esteem or change in self-esteem.

The second major aim is to test certain key aspects of social identity theory, and to use the results of those tests to suggest necessary modifications to that theory. For example, the claim that in-group evaluation is determined by the outcome of between-group social comparison outcomes is tested and it is suggested in the light of the results that in-group evaluation is affected by movement toward or away from in-group goals, as well as by attainment or non-attainment of them, and also that positive in-group distinctiveness is only one (non-necessary) such goal.

The third main aim is to identify areas in which social identity theory might benefit from modification, regardless of the validity of the theory as it stands at present. One such area concerns extending the theory beyond the strictly intergroup arena so as to include predictions concerning individual and inter-individual phenomena. A second concerns making the normative component of social identity theory much more explicit, in such a way as to justify the claim that the theory can provide a "genuinely social" social psychological account of group and intergroup phenomena.

Another way of describing the main aims of this thesis is to state that it investigates a single (compound) question which lies at the heart of Tajfel's social identity theory, namely: what are the psychological and behavioural effects of

(multigroup) social categorization, both when it occurs "in isolation" and when it occurs in combination with other factors, and by what processes do such effects come about? In particular, when and how does (multigroup) social categorization lead to (perhaps various types of) intergroup discrimination?

TAJFEL'S SOCIAL IDENTITY THEORY: A DETAILED EXPOSITION

Tajfel's social identity theory is a theory of intergroup conflict intended to supplement Sherif's "realistic conflict theory" (Tajfel & Turner, 1979: 33-34). The latter theory holds that intergroup conflict develops when members of two or more groups attempt to satisfy negatively interdependent group goals (Campbell, 1965; Sherif et al, 1961; Sherif & Sherif, 1953, 1967).¹

Tajfel maintained that realistic conflict theory needed supplementing because:-

(1) Realistic conflict theory "does not focus either upon the processes underlying the development and maintenance of group identity nor upon the possibly autonomous effects upon the in-group and intergroup behaviour of these 'subjective' aspects of group membership" (Tajfel & Turner, 1979: 34), and

(2) Empirical studies suggest that negative interdependence of group goals may be neither necessary nor sufficient for intergroup conflict to occur.

Social identity theory supplements realistic conflict theory by proposing that *some* intergroup conflict is the result of variables "inherent in the intergroup situation itself"

¹ Negative interdependence of goals means that progress toward one goal entails, requires or is identical to movement away from the other.

(Tajfel & Turner, 1979: 40). It claims that in *addition* to "realistic" or "instrumental" competition stemming from negatively interdependent group goals, there is also "social" competition which "is motivated by self-evaluation and takes place through social comparison" (Tajfel & Turner, 1979: 41).

A need for self-evaluation, or rather a need to obtain, maintain or enhance satisfactory self-esteem (i.e. a satisfactory self-concept), is the motivational engine of social identity theory (Tajfel, 1978a: 61; Tajfel, 1978b: 9; Tajfel & Turner, 1979: 40). To understand how that self-esteem need will operate in contexts where group membership is salient it is necessary to distinguish between personal and social identity.

Personal identity comprises of the cognitive, emotional and evaluative aspects of an individual's self-concept which derive from their awareness of themselves as a unique human being. Social identity consists of the cognitive, emotional and evaluative aspects of an individual's self-concept "which derive from membership of social groups that are salient to them" (Tajfel, 1978b: 14; Tajfel & Turner, 1979: 40). Social groups are salient to group members if they "have internalized their group membership as an aspect of their self-concept: they must be subjectively identified with the relevant in-group" (Tajfel & Turner, 1979: 41; see also Tajfel, 1982: 491).

Social identity is "positive or negative according to the evaluations...of those groups that contribute" to that social identity (Tajfel & Turner, 1979: 40). Evaluations of specific in-groups, in turn, "are determined with reference to specific other groups through social comparisons in terms of value-laden attributes and characteristics. Positively discrepant comparisons between in-group and out-group produce high prestige; negatively discrepant comparisons between in-group and out-group result in low prestige" (Tajfel & Turner, 1979:

40. See also Tajfel, 1978a: 83, 95; Tajfel 1978b: 16-19). It should be noted that "in-groups do not compare themselves with every cognitively available out-group: the out-group must be perceived as a relevant out-group" (Tajfel & Turner, 1979: 41, emphasis added) for in-group evaluation, social identity and self-esteem to be affected. It should also be noted that "[n]ot all between-group differences have evaluative significance (Tajfel, 1959), and those that do vary from group to group" (Tajfel & Turner, 1979: 41). If between-group differences do not have evaluative significance for in-group members in-group evaluation, social identity and self-esteem will be unaffected by such differences (Tajfel, 1982: 491). The term positive in-group distinctiveness can be used to refer to superiority of salient in-groups to relevant out-groups on dimensions valued by in-group members and the term negative in-group distinctiveness can be used to refer to inferiority of salient in-groups to relevant out-groups on dimensions valued by in-group members.²

Despite the first sentence of the preceding paragraph, two factors other than between-group differences affect in-group evaluation, social identity (derived from the in-groups in question), and (social identity contingent) self-esteem (Tajfel & Turner, 1979: 45).³ These are the perceived stability and legitimacy of such between-group differences.

The perceived stability of positive or negative in-group distinctiveness affects how secure resulting in-group

² In this thesis the terms *positive* and *negative* in-group distinctiveness will be restricted to such situations. It should be noted that in other literature these terms are frequently used more generally to refer to any in-group superiority or inferiority, regardless of in-group salience, out-group relevance and dimension value.

³ The phrase "social identity contingent" self-esteem is used as a reminder that social identity only affects self-esteem when that social identity is salient, and only *fully* determines individuals' self-esteem when *only* their social identity is salient.

evaluations, social identities and self-esteem are. If valued in-group distinctiveness is perceived to be stable, (or, in the jargon of social identity theory, "cognitive alternatives" to the present situation do not exist), resultant effects on in-group evaluations, social identities and self-esteem will be secure. If valued in-group distinctiveness is perceived to be unstable (i.e. if cognitive alternatives to the present between-group status differential do exist), resultant effects on in-group evaluations, social identities and self-esteem will be insecure (Tajfel, 1978a: 87; Tajfel, 1978b: 7; Tajfel & Turner, 1979: 45). This is important because social identity and self-esteem will only be "satisfactory" when they are both positive and secure, and will be unsatisfactory when they are either negative or insecure.

The legitimacy of in-group distinctiveness refers to whether that distinctiveness is perceived to be "related to a conflict of values, i.e. ...based on unfair advantages, various forms of injustice, exploitation, illegitimate use of force, etc." (Tajfel, 1978a: 89). Where in-group distinctiveness is perceived to be illegitimate (i.e. where it results in or is accompanied by a conflict of values) then the "usual" effects of that distinctiveness on in-group evaluation, social identity and self-esteem may be cancelled or even reversed. That is, positive in-group distinctiveness accompanied by a conflict of values may result in neutral or even negative effects on in-group evaluations, social identity and self-esteem, while negative in-group distinctiveness perceived by the inferior group to be illegitimate may result in neutral or even positive effects on in-group evaluations, social identity and self-esteem.

To summarize so far, legitimate positive in-group distinctiveness results in positive contributions to in-group evaluations, social identity and self-esteem, while legitimate negative in-group distinctiveness results in negative contributions to in-group evaluations, social identity and

self-esteem.⁴ Illegitimate valued in-group distinctiveness has less predictable effects on in-group evaluations, social identity and self-esteem, as such effects will depend on the relative importance or weight of the in-group distinctiveness compared to that of the perceived illegitimacy associated with it. Further, the effects of valued in-group distinctiveness on in-group evaluations, social identity and self-esteem will be secure if that distinctiveness is stable, and insecure if it is unstable.

As was mentioned above, the "fundamental motivational assumption" within Tajfel's social identity theory is that individuals strive to obtain, maintain or enhance satisfactory self-esteem. Such self-esteem is satisfactory if it is both positive and secure; self-esteem is unsatisfactory if it is either negative or insecure. Because of the effects of social identity on self-esteem, therefore, individuals will strive to be members of securely and legitimately positively distinct in-groups. If in-groups are negatively distinct, or if their positive in-group distinctiveness is perceived to be illegitimate and/or unstable, membership of those groups will contribute to unsatisfactory social identity and self-esteem, and the individuals concerned will be motivated to take steps to become members of securely and legitimately positively distinct in-groups.

When social identity is unsatisfactory group members have at their disposal two main possible strategies to bring about membership of securely and legitimately positively distinct

⁴ The word "contributions" here is crucial. In-group evaluations, social identity and self-esteem are each multiply determined. Positive or negative in-group distinctiveness for a single in-group from a single out-group on a single valued comparison dimension is unlikely to ever fully determine overall evaluations of that in-group (which is dependent on the outcome of multiple comparisons on multiple dimensions from multiple out-groups), let alone overall social identity (which is dependent on multiple in-group memberships) or self-esteem (which personal identity also contributes to).

groups: leaving the unsatisfactory in-group (i.e. "exit") or changing the in-group so that it becomes satisfactory (i.e. "voice"). There are various options within these two main strategies.

Exit is essentially an individualistic response to unsatisfactory social identity, in that it involves individuals leaving whichever group is contributing to that social identity. Exit can be actual, where individuals physically leave the unsatisfactory group, or psychological, where individuals physically remain in the group but psychologically distance themselves from it (Tajfel & Turner, 1979: 43). Usually, exit from an unsatisfactory group is accompanied or followed by an attempt to join (or establish) a more satisfactory group (known as "social mobility"). The "usually" in the previous sentence implies that occasionally exit alone can occur, i.e. exiting individuals simply renounce membership of the unsatisfactory group without attempting to join (or establish) a more satisfactory one (i.e. rather than "exchanging" an unsatisfactory social identity for a more satisfactory one, such individuals exit from the unsatisfactory group in order to simply operate in terms of more satisfactory personal identities).

Exit is not always an available or an attractive option. There can be barriers both to leaving the unsatisfactory group and to joining more satisfactory ones. Leaving a group may be physically impossible, e.g. where others continue to treat one as a member of the group no matter how much one tries to distance oneself from it, or it may "conflict with important values which are themselves a part of [the individual's] acceptable self-image" (Tajfel, 1978a: 64), e.g. where membership of and loyalty to the group is important to the individual, and where leaving it would seem like treachery, to the group and/or to themselves (Tajfel & Turner, 1979: 35). Similarly, joining a more satisfactory group may be physically impossible (e.g. because one does not have the requisite

"entry qualifications"), or it may be psychologically impossible (e.g. where there is a cultural prohibition on social mobility, or where others refuse to accept one as a member of the new group, or where joining the alternative group conflicts with the would-be exiting individual's acceptable self-image).

The alternative to exit in the face of unsatisfactory group membership is voice. This is essentially a social strategy in that it requires numbers of individuals to remain in the currently unsatisfactory group and to take steps to make it satisfactory. Again, there are a number of options within the strategy of voice.

The first is direct social competition. This is where group members employ intergroup discrimination in an attempt to obtain secure and legitimate positive in-group distinctiveness where such distinctiveness is presently lacking. As with exit, this will not always be an available or attractive option. First, if one's group is presently securely negatively distinct (i.e. if no cognitive alternatives to in-group inferiority exist) there is probably nothing to be gained, and potentially much to be lost, by discriminating against the presently superior group (Tajfel & Turner, 1979: 45). Antagonizing a group with superior powers of retaliation is not a sensible way to try and overcome inferiority. Second, intergroup discrimination may not be perceived as a legitimate method of trying to achieve positive in-group distinctiveness: it may conflict with values important to the in-group (and to the acceptable self-image of in-group members). Third, alternative methods of rectifying unsatisfactory self-esteem may seem safer and/or more attractive (e.g. where there are few prohibitions to exit).

The second main type of voice is called social creativity. This is another social strategy where numbers of group members remain in their presently unsatisfactory group and employ

strategies aimed at making it more satisfactory, but this time without engaging in full-scale direct social competition. This is done by "redefining or altering the elements of the comparative situation" (Tajfel & Turner, 1979: 43) using one or more of various methods, including: (i) choosing new dimensions upon which to compare the in-group with out-groups; (ii) reversing the values attached to comparison dimensions, so that previously negative in-group attributes are now claimed to be positive; (iii) choosing new out-groups to compare the in-group with; or (iv) refusing to engage in between-group comparisons at all (Tajfel, 1978b: 13).

Social creativity is often the least attractive of the options by which unsatisfactory social identity can be countered. This is because it essentially involves attempts to ignore or circumvent situations which have already been acknowledged to result in unsatisfactory group membership and which still exist. That is, social creativity is a compensatory activity, a type of "denial", accompanied by all the difficulties associated with such a strategy (Tajfel, 1978b: 17). Not only will the situation which led to unsatisfactory in-group membership still exist, the strategies attempted to minimize the negative effects of that situation may in themselves make the situation worse, e.g. by antagonizing a dominant out-group in seeming both to reject its dominance and to claim superiority to it in other ways (Tajfel, 1978b: 18).

This exposition of Tajfel's social identity theory is obviously more complex than the one offered at the beginning of this chapter. The diagrams below illustrate "how the theory works" when individuals come to perceive that a group to which they belong is either superior (Diagram 1.1) or inferior (Diagram 1.2) to some comparison group.

As mentioned above, if the in-group is not salient (Box 2 in each diagram), if the dimension of superiority or inferiority is not valued by in-group members (Box 3 in each diagram), or

if the comparison group is not considered relevant (Box 4 in each diagram), there will be no effect on in-group evaluation, social identity or self-esteem (Box 5 in each diagram).⁵ Only when the in-group is salient, the comparison dimension is valued, and the comparison group is relevant, will in-group superiority or inferiority result in *positive* (Box 6 in Diagram 1.1) or *negative* (Box 6 in Diagram 1.2) in-group distinctiveness, respectively. This suggests that members of in-groups with "irrelevant" superiority may be motivated, if the need arises, to make that superiority relevant (Box 5 of Diagram 1.1); by making the presently non-salient in-group salient, by valuing the presently unvalued dimension of in-group superiority, and/or by making the presently irrelevant out-group relevant. Unless such a need arises, however, members of irrelevantly superior groups will not be motivated to take any action. Members of in-groups with "irrelevant" inferiority, however, will be motivated to keep that inferiority irrelevant (Box 5 of Diagram 1.2); by continuing to see the in-group as non-salient, by continuing not to value the dimension of inferiority, and/or by continuing to see the out-group as irrelevant.

Before predicting the effects of *valued* distinctiveness, social identity theory needs to consider whether that distinctiveness is perceived to be legitimate (Box 7 in each diagram) and stable (Box 8 in each diagram). Only when the distinctiveness is perceived to be both legitimate and stable can it be predicted to make positive (Box 9 in Diagram 1.1) or negative (Box 9 in Diagram 1.2) contributions to in-group evaluations, social identity and self-esteem.

If positive in-group distinctiveness is perceived to be illegitimate and/or unstable then the positive effects of that

⁵ In the case of the "in-group" not being salient it is probably more correct to say that evaluation of the inferior or superior group may be affected by the comparison, but it will not be an in-group.

distinctiveness on in-group evaluation, social identity and self-esteem will be threatened (Box 10 of Diagram 1.1). In the case of perceived illegitimacy this is because the conflict of values which accompanies or results from the positive in-group distinctiveness will contribute negatively to in-group evaluations, social identity and self-esteem. In the case of instability this is because that instability makes for insecure positive contributions to in-group evaluations, social identity and self-esteem from the positive in-group distinctiveness.

Similarly, if negative in-group distinctiveness is perceived to be illegitimate and/or unstable then the negative effects of that distinctiveness on in-group evaluation, social identity and self-esteem will be unstable (Box 10 of Diagram 1.2). In the case of perceived illegitimacy this is because that illegitimacy may be accompanied by "discounting" of the in-group inferiority, which may negate or reverse the "usual" effects of negative in-group distinctiveness on contributions to in-group evaluations, social identity and self-esteem. In the case of instability this is because that instability allows for the possibility of action which will reduce, or even reverse, the negative in-group distinctiveness.

When the contributions from valued in-group distinctiveness on in-group evaluation, social identity and self-esteem are either threatened or insecure, in-group members will attempt to employ social competition, social mobility and/or social creativity to achieve satisfactory (i.e. legitimate, secure, and positive) self-esteem. Which of these actions they will take depends on both the source of the dissatisfaction and the options available to them.

If social mobility is possible (i.e. if there are no objective or psychological barriers to leaving the in-group), so that a person can easily leave an unsatisfactory group in order to obtain or operate in terms of an alternative, more

satisfactory (social or personal) identity (Box 11 in each diagram), this will be the most attractive option (Tajfel, 1978a: 64, 67). Until such time as social mobility is successfully employed, however, contributions to in-group evaluations, social identity and self-esteem from the current threatened positive (Diagram 1.1) or insecure negative (Diagram 1.2) in-group distinctiveness will themselves be positive but threatened (Box 12 of Diagram 1.1) or negative but unstable (Box 12 of Diagram 1.2).

If social mobility is not possible, and social competition in and of itself is deemed likely to make membership of the present in-group satisfactory (i.e. by achieving perceived legitimate and secure positive in-group distinctiveness), then this will be the most attractive option (because of the difficulties associated with social creativity which were outlined above) (Box 13 of each diagram). Until such time as social competition is successfully employed, however, contributions to in-group evaluations, social identity and self-esteem from the current threatened positive or insecure negative in-group distinctiveness will themselves be positive but threatened (Box 14 of Diagram 1.1) or negative but unstable (Box 14 of Diagram 1.2).

Finally, if social mobility is not possible, and social competition alone is perceived as unlikely to make membership of the present group satisfactory, then social creativity will be employed, either in conjunction with social competition or in place of it (Tajfel & Turner, 1979: 44). Until such time as social creativity is successfully employed, however, (with or without social competition), contributions to in-group evaluations, social identity and self-esteem from the current threatened positive or insecure negative in-group distinctiveness will themselves be positive but threatened (Box 15 of Diagram 1.1) or negative but unstable (Box 15 of Diagram 1.2).

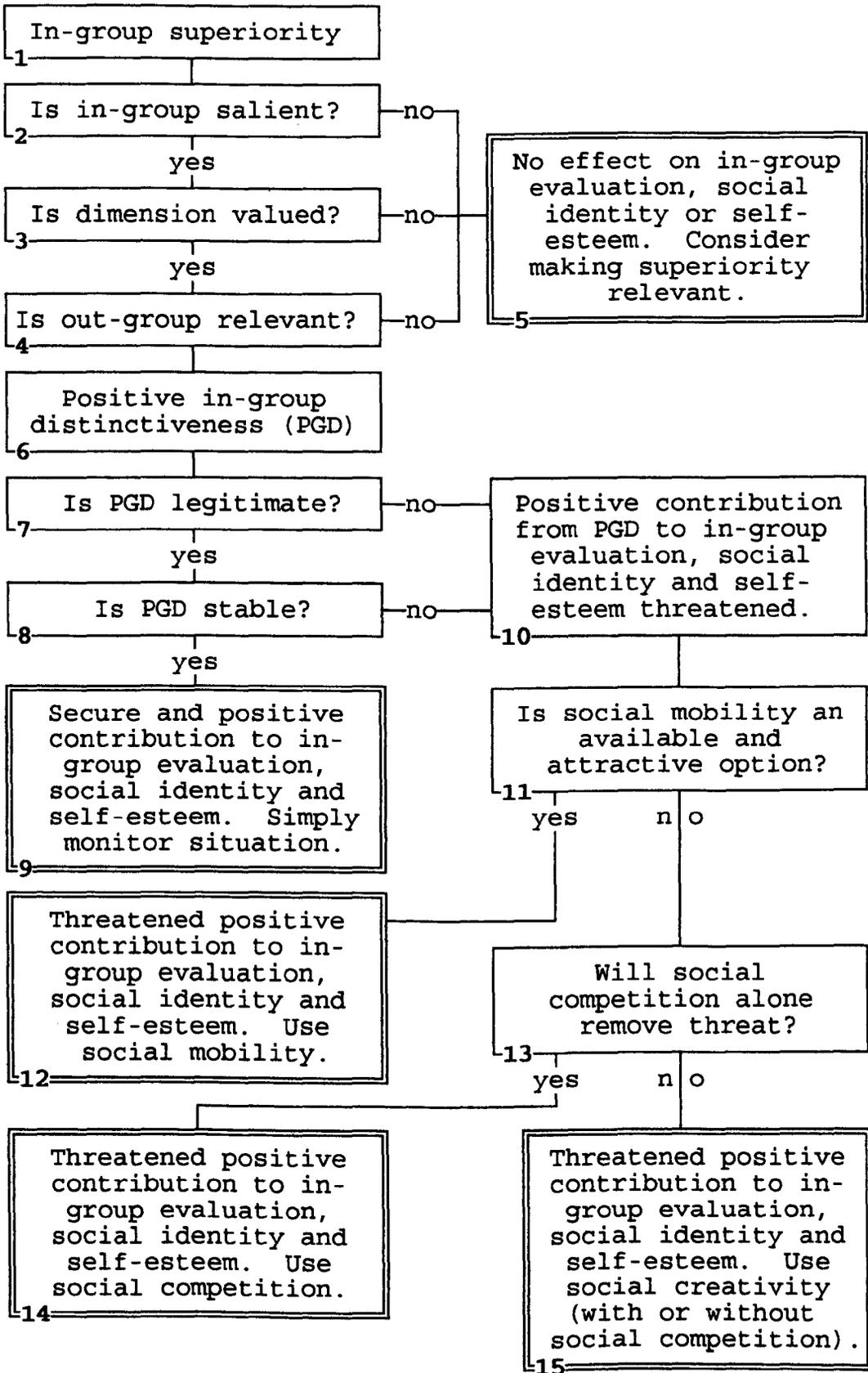


Diagram 1.1: The consequences of perceived in-group superiority

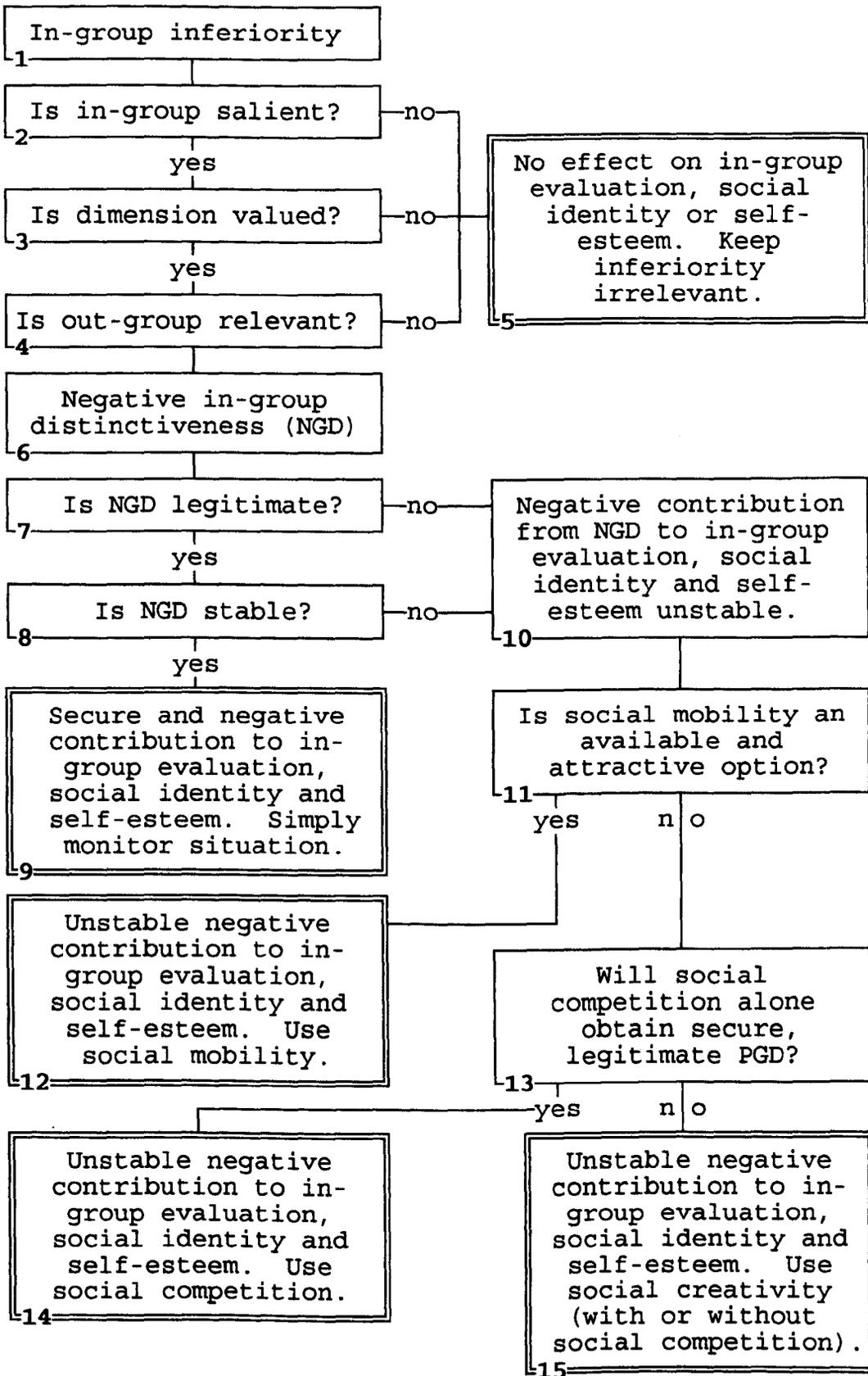


Diagram 1.2: The consequences of perceived in-group inferiority

THE "SUFFICIENCY CONDITION" WITHIN SOCIAL IDENTITY THEORY

One central issue with respect to social identity theory is its relationship with Sherif's "realistic conflict theory". Tajfel claims that one reason why the former theory needs supplementing by the latter is that negative interdependence of group goals is not a necessary condition for the emergence of intergroup conflict. Specifically, group members' awareness that they are members of one group but *not* members of another (or others) may in and of itself be a sufficient condition to generate intergroup discrimination, even when there is no perceived interdependence between group goals. Thus, what might be termed the "sufficiency condition" within social identity theory states "that the mere perception of belonging to two distinct groups - that is, social categorization per se - is sufficient to trigger intergroup discrimination favouring the in-group" (Tajfel & Turner, 1979: 38). This is an important assertion, central to several of the arguments to be made later in this thesis. Thus, it is as well to be clear that the "sufficiency condition" is widely adhered to, and also to be clear as to the exact claim being made by that condition.

Several commentators report that social categorization (which, as the term is used, involves at least two groups of which any individual is a member of only one) is sufficient to give rise to certain forms of intergroup discriminatory behaviour. Thus, Turner (1983a: 351) refers to the "well replicated finding that social categorization *is alone* sufficient for intergroup discrimination" (emphasis added). Similarly, Wilder (1986: 311) refers to an "exhaustive literature [which] indicates that the *mere* categorization of persons into an ingroup and an outgroup is sufficient to foster bias (i.e. ingroup favouritism at the expense of the outgroup) (emphasis added). Turner repeats that "imposing social categorizations upon people even on an explicitly random basis...produces

discriminatory intergroup behaviour" (Turner et al., 1987: 27-28) (emphasis added). Hogg & Abrams (1988: 51) agree that "social categorization - the discontinuous classification of individuals into two distinct groups - is sufficient to generate intergroup discrimination". Bagby & Rector (1992: 398) also concur in stating that "the mere categorization of individuals is all that is necessary to create ingroup favouritism" (emphasis added). And Espinoza & Garza (1985: 381) cite Tajfel (1978) to argue that "[a]ny social categorization which creates identifiable, distinct social groups (e.g., ethnicity, gender, Democrat/Republican) can be sufficient to evoke spontaneous biases in perception, attitude, and behaviour favouring one's group (Tajfel, 1978)" (emphasis added).

Billig & Tajfel (1973: 28-29) are perhaps more careful than the above commentators in that they clearly argue, consistent with the exposition of social identity theory given above, that it is not so much (multigroup) social categorization per se which is referred to by the "sufficiency condition" (as it is possible for external imposition of social categorization without those categorized even being aware of such categorization), but that it is "an individual's act of categorizing his social world into distinct social groups, into 'them' and 'us', [which] can be, at least in our societies, a sufficient condition for introducing in his behaviour certain forms of ingroup favouritism and of discrimination against the out-group" (emphasis altered). Thus, it is not (multigroup) social categorization per se which is held to be sufficient for the emergence of intergroup discrimination, it is the recognition of such social categorization by the individual(s) categorized (i.e. multigroup social categorization plus self-categorization into one of those groups on the part of those categorized) which can be sufficient to promote intergroup discriminatory behaviour on the part the categorized individual(s). (Although it is, in fact, a robust finding that externally

imposed multigroup social categorization regularly results in mean intergroup discriminatory behaviour on the part of those categorized.)⁶

It is also important to be clear, despite possible appearances to the contrary in some of the quotes above, that the claim of the "sufficiency condition" is not that (recognized, multigroup) social categorization is *generally* or *universally* sufficient to promote intergroup discrimination (i.e. that social categorization will *always* result in intergroup discrimination): it is that "*mere*" social categorization (i.e. social categorization "*per se*", "*alone*", or "*in isolation*") is sufficient (when accompanied by self-categorization into those categories by the people categorized) to promote intergroup discrimination. The exposition above makes it clear that many other factors (e.g. the availability of cognitive alternatives, a lack of opportunities for social mobility, etc.) are held to be necessary by social identity theory before intergroup discrimination can be confidently predicted to follow "non-mere" social categorization. Further, the phenomenon of out-group favouritism (i.e. the opposite to intergroup discrimination in favour of the in-group) - which itself requires recognized multigroup social categorization - was cited by Tajfel as another of the reasons why Sherif's "realistic conflict theory" was in need of supplementation. Thus, social categorization (however conceived) is not held by social identity theory to be universally sufficient for intergroup discrimination in the sense that whenever social categorization occurs intergroup discrimination will follow. Rather, the claim is that (recognized, multigroup) social

⁶ For further references to the "sufficiency condition" see, among many others, Allen & Wilder (1975: 971, 1979: 74); Bornstein et al. (1983: 321); Brown (1986: 544); Brown (1988: 254); Lemyre & Smith (1985: 660); Locksley et al. (1980: 774); Mummendey et al. (1992: 125); Oakes & Turner (1980: 295); Platow et al. (1990: 222); Sachdev & Bourhis (1984: 36, 1985: 415, 1987: 278); St. Claire & Turner (1982: 307); Turner (1975: 5); Turner (1980: 131); Turner (1981b: 75); Worchell & Norvell (1980: 764).

categorization "alone" (i.e. in the absence of any other relevant psychological stimuli) is sufficient to promote intergroup discrimination. (Although if the "sufficiency condition" proves to be true there would perhaps be some reason to suppose that social categorization does, in fact, always promote intergroup discrimination, but that other factors can attenuate, cancel, reverse, or indeed accentuate, such an effect.)

Now that the form of the "sufficiency condition" held within social identity theory has been spelled out, the crucial question is, of course, is it true? Is it in fact the case that mere (recognized, multigroup) social categorization is sufficient to promote intergroup discrimination and, if so, to what extent is it true, and what may we conclude from such a truth?

There can be no denying that it is a very robust finding (see Mullen et al, 1992 for a review, including some exceptions) that the nearest thing possible to "mere" social categorization⁷ does in fact lead to mean attitudinal and/or behavioural in-group favouritism on the part of the individuals categorized. Such in-group favouritism takes at least two forms: on average subjects allocate more "goods" to in-group members than to out-group members, and, again on average, subjects also sacrifice possible "goods" for in-group members when doing so achieves a certain degree of superiority for the in-group (and/or its members) over the out-group (and/or its members) with respect to the "goods" received by each. Thus, there can be no doubt that the average behaviour which follows mere social categorization is discriminatory between in-group and out-group in favour of the former. To

⁷ i.e. arbitrary or random categorization into one or the other of two novel categories with: anonymity of category membership; no direct interaction or communication between those categorized; and no obvious rational or instrumental link between the social categorization and possible subsequent behaviour.

the extent therefore that this average individual behaviour is construed as *intergroup* behaviour, social identity theory's "sufficiency condition" must be taken to be valid: mere (recognized, multigroup) social categorization must be accepted as a sufficient cause of intergroup discrimination (with the implication that non-mere social categorization is also likely promote a tendency toward intergroup discrimination).

Social identity theory does, of course, accept that the behaviour following mere social categorization should be construed as intergroup discrimination, and in doing so finds justification for the "sufficiency condition". That is, it accepts such findings as support for its claim that social categorization *per se* (understood as group members' *perceptions* that they are members of one group and not of another) is sufficient to promote intergroup discrimination (thereby also justifying the claim that social identity theory is a necessary supplement to Sherif's "realistic conflict theory").

What is more, social identity theory offers an account of how mere social categorization promotes intergroup discrimination, which runs roughly as follows. Individuals become aware of their membership of a group. To the extent that they do this their social identity as group members becomes salient and their self-esteem becomes dependent upon their evaluation of their in-group. And, as in-groups are held to be evaluated positively to the extent that they are perceived as positively distinct from out-groups (and negatively to the extent that they are perceived as negatively distinct from out-groups), group members therefore try to perceive the in-group from which their social identity is derived as positively distinct. In the situation of social categorization *per se*, the *only* route to positive in-group distinctiveness (and away from negative in-group distinctiveness) is to favour the in-group

over the out-group.⁸ Mere social categorization promotes intergroup discrimination, therefore, because: (i) individuals strive for satisfactory self-esteem; (ii) social categorization entails that individuals' social identities are salient (and therefore their self-esteem is dependent upon their evaluation of the in-group, which is in turn dependent on perceptions of positive in-group distinctiveness); and, in the situation of social categorization per se, (iii) intergroup discrimination in favour of the in-group is the only available route to positive in-group distinctiveness and positive in-group evaluation, which alone will satisfy their social identity and self-esteem needs.

An important implication of social identity theory's interpretation of the effects of social categorization per se is that to the extent that individuals perceive they are members of one social category and not of another (i.e. to the extent that they recognize the mere multigroup social categorization), each and every individual will be predicted to engage in in-group preference. This is because recognizing the multigroup social categorization will result in individuals' social identities becoming salient, and the only means available to them for satisfying their resultant social identity contingent self-esteem needs will be to employ in-group preference in an attempt to obtain positive in-group distinctiveness and thereby positive evaluations of their in-groups and of themselves as in-group members. This is a straightforward consequence of social identity theory's account of why mere social categorization is sufficient to promote intergroup discrimination. Therefore, any instance of an individual not employing in-group preference following mere

⁸ It can also be noted that in the "standard" minimal group paradigm (i.e. in which social categorization per se is investigated), there is: only one salient in-group, only one comparison dimension, only one out-group, an unstable between-group situation (i.e. the possibility of cognitive alternatives), no in-group norms to potentially provide a conflict of values when engaging in in-group favouritism, and no possibility of exit.

social categorization has to be "explained away" by social identity theory. Further, apart from having to "explain away" any behaviour following mere social categorization which is non-discriminatory, social identity theory cannot give any explanation of (i.e. is "silent" about) such behaviour. It is clearly the case that social identity theory should be "allowed" to "explain away" a certain amount of non-discriminatory behaviour following mere social categorization, (e.g. by claiming that the non-discriminating individuals were not aware of the social categorization), but, equally clearly, threats to social identity theory's "sufficiency condition" increase proportionally to the extent of such behaviour. If there is "too much" non-discriminatory behaviour following mere social categorization, especially if there is evidence that the categorized but non-discriminatory subjects are aware of their categorization, then: (i) social identity theory's claim that mere (recognized, multigroup) social categorization is sufficient to engender intergroup discrimination is threatened; and, (ii) social identity theory's adequacy as an account of the behaviour following from mere social categorization will be in doubt.

It is important to emphasise that the description above is of social identity theory's account of the effects of *mere* social categorization. When social categorization is "non-mere" social identity theory still holds that: (i) individuals strive for satisfactory self-esteem; and, (ii) to the extent that social identities are salient individuals' self-esteem will be (in part) determined by their evaluations of their in-groups, which in turn depends on their perceptions of valued in-group distinctiveness. However, as already noted, social identity theory does not hold that social categorization will *generally* or *universally* promote intergroup discrimination, as when such categorization does not occur "in isolation" it will not be the case that showing favouritism towards a particular in-group relative to a particular out-group on a specific dimension is the *one and only* route to satisfying self-esteem

and social identity needs. One reason that such discrimination will not be the only route to satisfying such needs is that alternative strategies may be available (e.g. exit, social creativity). Further, such discrimination may not even be an available, attractive, or indeed a necessary possibility (e.g. because an in-group is already securely and legitimately positively distinct, because an inferior in-group lacks cognitive alternatives, or because intergroup discrimination would entail a conflict of values, etc.).

Similarly, although social identity theory predicts that the extent of in-group membership salience will positively correlate with the extent of in-group favouritism shown following mere social categorization (as increasing in-group salience increases the extent to which self-esteem is determined by social identity, in-group evaluation, and positive in-group distinctiveness, and in-group favouritism is the only route to such things when social categorization is mere), it does not predict such a correlation when social categorization is non-mere (as there will be a variety of strategies available for satisfying social identity and self-esteem needs, of which any single given possible instance of in-group favouritism - if available - may be only one).⁹

In this section it has been argued that social identity theory holds that mere (multigroup) social categorization is sufficient to result in each and every individual who recognizes the categorization engaging in in-group preference, and that the more salient the mere social categorization, the stronger the discrimination. However, consistent with the exposition given in the previous section, social identity

⁹ See, among others, Brown et al. (1986: 275), Hinkle & Brown (1990: 62), and Kelly & Kelly (1994: 64) for examples where social identity theorists at least seem to interpret that theory as saying that there will be a correlation between (non-mere) in-group salience and in-group favouritism, and Turner (1987: 30) for a denial of that interpretation, similar to the one made here.

theory is clear that when social categorization does not occur in isolation a whole host of considerations other than the social categorization (in and of itself) need to be taken into account in order to predict the behaviour of the socially categorized individuals.

IS SOCIAL IDENTITY THEORY INDIVIDUALISTIC?

One of the central arguments developed in this thesis is that social identity theory does not fully live up to its claim of being a "genuinely social" social psychological theory, and so it is as well to "flag up" this line of argument now. This argument is based on a claim that a genuinely social theory of group and intergroup behaviour must include a conceptualization of group-level phenomena (by which individuals must be guided for their behaviour to be considered as "genuine" group behaviour).

Without going too far into in-depth ontological musings, it is possible to consider groups as entities "in their own right", with their own "goals", "aims", etc.. To give a relatively simple example, a group-level "goal" of certain feminist groups is to overcome biases in society in order to achieve equality of opportunity for people of both sexes/genders. Similarly, a group-level goal of the Anti-Apartheid movement is to end apartheid.

Groups may be considered as "entities in their own right", but most thinkers now accept that individuals are the only entities with genuine powers of thought, volition, and action. Thus, groups cannot "act on their own behalf": they have to have individuals act on their behalf.

At an individual-level of analysis it is possible to distinguish between behaviour people engage in "purely for

themselves" and behaviour they engage in "on behalf of" someone or something else (whether or not the latter behaviour also satisfies individual desires and/or needs). An example of the latter sort of behaviour is when individuals do things "on behalf of" particular in-groups, e.g. when people go to war to fight "for" their country.

Importantly, to the extent that individuals (be they in-group members or not) wish to genuinely act "on behalf of" a particular group, they have to know what acting on behalf of that group entails. They can, to be sure, engage in action which they believe will *benefit* the group, but genuinely acting *on behalf of* the group (i.e. doing what the group - were it a sentient being - would "want" them to do) requires that they know what the group's "needs", "aims", "desires", "prescriptions", "goals", etc. are.¹⁰

At the individual-level of analysis, therefore, people's behaviour can be considered as (at least attempted)¹¹ *group* behaviour if those people are trying to do what they perceive the *group* "wants" or "expects" them, as group members, to do (as revealed by, for example, the group's manifesto, code of conduct, *raison d'être*, or whatever). Group behaviour is thus motivated by individuals' desires to act on behalf of a group:

¹⁰ An analogous example of this distinction at the inter-individual level might be where a parent does something to benefit their child, regardless of what the child wants themselves (e.g. disciplines them), as compared to where a parent does something "on behalf of" their child, perhaps even when the parent believes that this will not "benefit" the child (e.g. where they buy the child something the parent disapproves of but the child desires).

¹¹ Possible indications of the veracity of such an analysis are found whenever spokesmen or women (speaking on behalf of a group, irrespective of what they personally believe) deny that the actions of particular individuals are in fact on behalf of the group the spokesperson represents, such as when representatives of the British army or government claimed that any British soldiers serving in Northern Ireland adopting a "shoot to kill" policy were not acting on behalf of the British army/government.

the motivating desires being individual variables (i.e. individuals wanting to act on behalf of the group), while the means by which their behaviour is "guided" or "steered" is a social or a group-level variable (i.e. one or more individuals doing what the group "wants").

Social identity theory conceives of (certain forms of) group behaviour as motivated by individuals striving to achieve or maintain positive in-group distinctiveness in order to ultimately achieve or maintain satisfactory self-esteem. Here, as in the account sketched just above, the fundamental motivation for group behaviour is an individual desire. But whereas such individual motivation was directed at achieving *group ends* in the account offered above (as well as individual ends: i.e. *wanting* to achieve group ends), for social identity theory the motivation is *exclusively individual*. That is, for social identity theory, there is no sense in which individuals discriminate (or whatever) because the group "wants" them to (i.e. they do not discriminate "on behalf of" the group): they discriminate solely because *they* want to (i.e. wholly "for themselves").

To put the same point in a slightly different way, social identity theory does not conceive of individuals discriminating (or whatever) because positive in-group distinctiveness is a *group-goal*, they discriminate because such distinctiveness is a goal each *group-member* has: positive in-group distinctiveness is a widespread individual goal (i.e. among members of the group), rather than a goal which genuinely exists at a group-level. Social identity theory is clearly social to the extent that positive in-group distinctiveness is held to be a goal which group members will *share* because of *common* group membership, and is also social to the extent that such common goals are held to lead to uniformities of behaviour by group members. It is argued here however that this is not genuine group behaviour because group members are acting solely *for themselves*: they are not acting

"on behalf" of anyone or anything else (i.e. the group "as a whole" and/or other group members).¹²

If this analysis is accepted, one implication of it is that genuine group behaviour (including intergroup discrimination) is simply not possible in the situation of social categorization per se. Without a history, minimal groups cannot have established or pre-existing group-level goals, norms, etc., to "guide" or "direct" the behaviour of in-group members.¹³ And, without intragroup interaction, minimal group members cannot generate goals, norms, etc. for the group, so again the mere group membership does not allow group members to ascertain how they might act "on behalf of" their in-groups. Even if they want to act on behalf of their in-groups, therefore, in-group members have no clues as to what the in-group "wants" or "expects" them to do.

Tajfel's original interpretation of minimal group behaviour attempted to solve this problem by postulating that there were "generic norms" in society, such that group members (*qua* group members) displayed both intergroup fairness and in-group solidarity (i.e. in-group preference, leading to intergroup discrimination) because that is what they believed society

¹² It may be illustrative to speculate what happens according to each account when individuals strive to achieve positive in-group distinctiveness, but fail. According to social identity theory group members will care about this failure only because it means that they personally are not deriving positive social identity from the situation: the "suffering" of in-group others or the group itself will be simply an irrelevant adjunct to what they themselves feel. According to the account being developed here, in contrast, individual group members will suffer precisely because the in-group and/or other in-group members suffer.

¹³ The one possible exception to this will occur if it proves to be the case that all groups have particular goals (such as positive in-group distinctiveness), simply by dint of being groups. This will make social identity theory a particular instance of a generic norm theory (see below), and, to be accepted, will require that empirical evidence be obtained that it is indeed the case that such generic norms exist.

wanted or expected of them. Such a generic norm interpretation was later abandoned, however, because it seemed to open up the possibility of explaining any behaviour post hoc, simply by claiming that such behaviour was conforming to some previously unidentified generic norm.

Tajfel turned instead to social identity theory, which substituted intra-individual needs for societal norms as the steering force behind intergroup behaviour. According to the argument being developed here, each of these accounts focuses primarily on the wrong level of analysis to best understand group behaviour: generic norm theory on the societal level and social identity theory on the intra-individual level, whereas the present argument suggests that the "privileged" level of analysis for group and intergroup behaviour is the group level.¹⁴

The implication of this line of reasoning for interpretation of "standard" minimal group behaviour is as follows. It is true that at an individual level social categorization per se does result in minimal group subjects on average discriminating between in-group and out-group in favour of the in-group. This is clearly "intergroup discrimination" in the sense that, on average, subjects discriminate between the groups (in favour of the in-group). This is, however, a *mean* of (relatively widespread) *individual* behaviour. It is not the case that the "groups" are discriminating against each other, or rather, that the individuals are discriminating between the groups "on behalf of" their own groups. For individual behaviour (such as in-group preference) to be considered as an instance of group behaviour (such as

¹⁴ Which is not to deny that societies and individuals can greatly influence the goals and norms of particular groups. It should also be noted that the argument sketched in this section (and developed a little more fully in the final chapter of this thesis) does not entail or require that social identity theory and/or generic norm theory are wrong. It simply means that neither is well suited to form the basis of an adequate theory of group and intergroup behaviour.

intergroup discrimination), it is necessary for the individuals concerned to be acting "on behalf of" groups, so that it can legitimately be claimed at a group level of analysis that it is the groups "themselves" that are behaving in a certain way (such as discriminating against each other). However, to be able to legitimately talk of group behaviour, it is necessary also to be able to talk about group-level goals (or whatever) which "guide" or "steer" the behaviour of individuals acting "on their behalf" (as opposed to merely taking the mean of group members' individual behaviour or motivation and calling this group behaviour or motivation).¹⁵

IDENTIFICATION "AS" AND IDENTIFICATION "WITH"

In some of the arguments to follow it will be useful to distinguish between identity and identification "as" on the one hand, and identity and identification "with" on the other. As has already been mentioned, social identity theory defines social identity as "those aspects of an individual's self-image that derive from the social categories to which he perceives himself as belonging" (Tajfel & Turner, 1979: 40). Thus, social identity as used within social identity theory concerns individuals identifying themselves "as" members of one or more particular social categories or groups (and

¹⁵ This distinction is illustrated by the difference between the statements that "Britain declared war on Germany" and "Over half of British people went to Germany on holiday this year". It can be noted that the former statement can be true even when only one individual in-group member with appropriate authority acts. It can also be noted that although it is possible to rephrase the latter statement as "Britain went to Germany on holiday this year", such a rephrasing, properly understood, still refers to the behaviour of (British) in-group members, not to the behaviour of the group as such. Thus, Chamberlain indulged in group behaviour because he acted "on behalf of" Britain (and the British), but British people are not usually acting "on behalf of" Britain when they go on holiday, no matter how much uniformity of behaviour they show, (although they may of course be perceived to be acting on behalf of their country).

evaluating that membership). That is, social identity is simply self-categorization as a member of one or more groups, together with the evaluative and emotional appraisals of such self-categorizations.

This can be differentiated from a sense of identity in which people identify "with" someone or something else. This sense of the term refers to people "caring" about others to the extent that they want for those others what the others are perceived to want for themselves. Thus, social identity in this sense refers to individuals "caring" about the groups identified "with": to the phenomenon of individuals wanting for the identified-with groups and/or group-members what they perceive the groups and/or group members (as group members) want for themselves.

To return to the main topic of the previous section, social identity theory conceives of group members indulging in certain forms of intergroup discrimination because they identify themselves as group members, and therefore discriminate in order to satisfy the evaluative aspects of those self-categorizations (i.e. to satisfy their self-esteem needs as group members). In contrast, the bare-bones of the group theory sketched above suggests that it is individuals identifying *with* groups which is most important in predicting and explaining group behaviour, to a degree regardless of whether or not the identifying individuals identify themselves as in-group members (although identifying oneself as an in-group member may well incline one to identify *with* the group).

THE CONTENTS OF THIS THESIS

Social identity theory relies heavily on the "sufficiency condition". If it is not the case that mere (recognized, multigroup) social categorization is sufficient to promote

intergroup discrimination then the claim that social identity theory is needed as a supplement to Sherif's "realistic conflict theory" is seriously undermined.¹⁶ Further, to the extent that mere (multigroup) social categorization does not result in individuals engaging in in-group preference, there will be a threat to social identity theory's claim that individuals with salient social identities necessarily search and strive for positive in-group distinctiveness because this is the only means by which they may satisfy their social identity and self-esteem needs. And, to the extent there are instances of mere (multigroup) social categorization not leading to in-group favouritism, especially when the individuals concerned recognize the social categorization, this also threatens social identity theory's explanation of such in-group favouritism: if salient social identity does not reliably result in in-group preference following social categorization per se, a postulated motivation to achieve satisfactory social identity cannot easily provide an adequate explanation of such in-group favouritism when it does occur.

Because of its centrality to social identity theory, several of the studies reported in this thesis investigate the "sufficiency condition" of social identity theory. The minimal group paradigm (i.e. the situation in which the nearest thing possible to social categorization per se takes place) therefore forms the starting point for empirical investigations reported in this thesis.

¹⁶ **NB:** For ease of exposition, and in line with the vast majority of literature within social identity theory, the terms "in-group preference", "in-group favouritism", and "intergroup discrimination (in favour of the in-group)" will be used interchangeably throughout this thesis unless specified to the contrary. In line with the arguments outlined above, however, it should be realised that "intergroup discrimination" at an individual level (i.e. in the sense of individuals showing in-group preference) is not considered in this thesis as necessarily an instance of group or intergroup behaviour, although it is clearly an instance of group members discriminating between their in-group and an out-group.

Chapter 2 describes a minimal group experiment in which subjects were asked to "think aloud" whilst making their allocations to in-group and out-group members. If the thinkaloud procedure does not interfere with processes which usually occur within the minimal group paradigm, then we would expect from social identity theory that the social categorization "per se" would result in all or at least the majority of socially categorized subjects engaging in in-group favouritism via a sequence of social categorization, social identity (i.e. identification "as" an in-group member), between-group social comparison, and an attempt to obtain positive in-group distinctiveness via in-group favouritism. The thinkaloud procedure was included primarily to see if any evidence could be found for why subjects behave as they do following mere social categorization.

Unfortunately, but nevertheless interestingly, the addition of the "thinkaloud" procedure to the standard minimal group paradigm resulted in subjects failing to employ the mean in-group favouritism usually obtained following mere social categorization. Unless the thinkaloud either prevented the subjects from being aware of the social categorization (i.e. failing to perceive two groups of which they were individually members of only one), or made available to subjects some means other than intergroup discrimination for meeting their social identity contingent self-esteem needs, this seems to suggest social categorization per se is not, in fact, sufficient to cause individuals to engage in in-group favouritism.

Further, there was very little evidence from subjects' protocols for social identity theory's hypothesized sequence from social categorization to intergroup behaviour. Despite the fact that every subject was aware that they were socially categorized, not one subject showed any evidence of having attempted to predict or affect the between-group outcome of everybody's allocations, or of having been motivated by social identity contingent self-esteem needs. However, it should

also be noted that - where there was any at all - there was also very limited, or at best mixed, empirical support for alternative theoretical explanations of minimal group behaviour, such as provided by generic norm theory (Branthwaite et al., 1979; Tajfel, 1970; Tajfel & Billig, 1974; Tajfel et al., 1971; Wetherell, 1982), equity theory (Ng, 1981, 1986), the behavioural interaction model (Rabbie et al., 1989), intergroup accentuation theory (Doise & Sinclair, 1973), self-categorization theory (Turner et al., 1987), or the demand characteristics explanation (Gerard & Hoyt, 1974).

There was some evidence, however, that where in-group favouritism did occur it was accompanied by subjects identifying with their imposed in-groups (in the sense of *caring about* their in-groups and/or their fellow in-group members, rather than in the sense of *recognizing* or *identifying themselves as members of* their in-groups), although not all subjects who identified with their imposed in-groups engaged in in-group preference. This suggests that while (recognized, multigroup) social categorization may not be sufficient to engender in-group favouritism, social identification with an in-group may be a necessary, but not a sufficient, condition of such discrimination (although the possibility that the discrimination caused the social identification with the in-group cannot be ruled out). Nevertheless, the sheer diversity of the behaviours adopted by the subjects in this study, and also of the explanations the subjects gave of those behaviours, suggests that the most sensible conclusion to reach may be simply that it is not possible to make any confident predictions concerning the effects of social categorization per se on subsequent group and intergroup behaviour, other than to make some bland comment that people behave in such a way as to try to make sense of their situation. Such a comment is relatively uninteresting because of the wide variety of possible means available to do this (of which in-group identification and in-

group favouritism in order to obtain positive in-group distinctiveness is only one).

Chapter 3 examines the effects of mere social categorization on subsequent behaviour from another angle. Specifically, the question was asked: what sort of intergroup discrimination (if any) does social categorization per se engender? In the same year that the classic exposition of social identity theory was published (i.e. Tajfel & Turner, 1979), Brewer (1979) reported that the intergroup discrimination exhibited by subjects in the minimal group paradigm typically takes the form of in-group preference rather than out-group derogation. That is, following mere social categorization subjects tend on average to favour the in-group over the out-group, but do not actually derogate the out-group whilst doing so. Thus, the sort of intergroup discrimination which social categorization results in seems to be of a rather benign sort. To adopt and adapt a rather outdated slogan, minimal group subjects tend to say "you're OK, we're better" rather than "we're OK, you're rubbish". This is clearly a different level of intergroup discrimination to that dealt with by empirical investigations of Sherif's realistic conflict theory (where out-group members were often verbally, and sometimes physically, abused).

Further, several commentators have claimed that even the intergroup discrimination which does typically occur in the minimal group paradigm may be merely an artifact of the procedures employed within that paradigm (Gerard & Hoyt, 1974; Schiffmann & Wicklund, 1992). That is, rather than employing in-group favouritism in an attempt to obtain positive in-group distinctiveness and thereby satisfactory in-group evaluation, social identity and self-esteem, subjects employ in-group preference simply because it seems the only meaningful behaviour available to them, especially when presented with allocation matrices which may "suggest" that this form of intergroup discrimination is expected or appropriate (Locksley et al., 1980: 776).

Part of the difficulty in trying to decide between these alternative possibilities stems from the "Byzantine complexity" of the Tajfel matrices which are typically used to measure intergroup behaviour and/or attitudes in the minimal group paradigm (Bornstein et al., 1983b: 376). Regardless of whether or not these matrices suggest, invite or force in-group favouritism, the design and scoring of them has been regularly criticised for making interpretation of subjects' allocation strategies extremely difficult (e.g. Aschenbrenner & Schafer, 1980; Bornstein et al, 1983a; Brewer & Silver, 1978; Hyland, 1979; Mummendey & Schreiber, 1983, 1984a).

The study reported in this chapter compared subjects' minimal group behaviour as indicated by the Tajfel matrices with that indicated by a new measure of social behaviour (termed "allocation grids") inspired by McClintock's (1988) social value vectors. These grids are argued to have a number of advantages over Tajfel matrices, only two of which will be mentioned here. First, they allow allocations of "bads" as well as or instead of "goods". Second, to a great extent they avoid constraining allocations such that a particular allocation to one member of a recipient pair does not entail a particular allocation to the other member of that pair (as is the case with the Tajfel matrices). Use of such a measure enables investigation of whether subjects in the minimal group paradigm employ in-group preference and/or out-group derogation when given a relatively free choice.

Results from this study were consistent with those usually obtained from minimal group studies in that significant mean in-group favouritism was indicated by the Tajfel matrix scores. However, examination of the grid scores revealed that only a sub-set of the minimal group subjects engaged in discrimination in favour of the in-group. Other subjects employed predominately fair or random allocation strategies. This result again suggests that mere social categorization is

not sufficient to engender discrimination in favour of the in-group, at least at an individual level.

Further, where in-group favouritism was employed by subjects, two distinct forms of such discrimination were discernable. "Strong" in-group favouritism involved subjects obtaining maximum in-group superiority by combining maximum in-group profit with maximum out-group derogation, but "weak" in-group favouritism involved subjects employing in-group profit and in-group superiority without out-group derogation. Therefore it appears that minimal group subjects do not always employ in-group favouritism following social categorization, but a proportion of those who do tend to do so "with a vengeance". The challenge for social identity theory is to identify which form of in-group favouritism (if any) will be employed in which circumstances. Clearly, social categorization per se is not a sufficient explanation either of in-group favouritism or of minimal group behaviour more generally.

Chapter 4 presents a study which attempts to go some way toward meeting the aforementioned challenge. A crucial implication from each of the previous studies is that social categorization per se (even when recognized) is not sufficient to engender any form of group (or intergroup) behaviour, let alone intergroup discrimination. This is because, as identified by the full exposition of social identity theory, "genuine" group behaviour requires much more than simple (recognized, multigroup) social categorization. It requires, at a minimum, that "a collection of individuals who perceive themselves to be members of the same social category...share some emotional involvement in this common definition of themselves, and achieve some degree of consensus about the evaluation of their group and of their membership of it" (Tajfel & Turner, 1979: 40). It might be added that a further requirement is achievement of some sort of consensus concerning how to go about achieving such evaluation

(remembering that intergroup discrimination is a *social*, that is a collective, strategy).

The study reported in Chapter 4 deliberately violated one of the procedural criteria of the minimal group paradigm, namely the one prohibiting intragroup communication. In the absence of existing group norms, it is difficult to see how consensus can be achieved concerning appropriate in-group behaviour without such intragroup communication. It is particularly difficult to understand how social categorization per se could bring about such consensus, a consensus which seems to be a necessary requirement for genuine group behaviour.

Drawing heavily on classic studies within social psychology which demonstrate the effects of group membership on group members' behaviour (e.g. Sherif's norm studies and the group polarization paradigm), the study reported in Chapter 4 made the standard minimal group paradigm the first of three within-subject conditions. In the second condition subjects were asked to make a second set of allocations to in-group and out-group others, but this time collectively in "sub-groups" of three members of the same (previously) minimal groups. Subjects were tape-recorded whilst reaching their allocation decisions in this condition. The subjects were then asked in the third condition to individually make a final set of allocations to in-group and out-group others (not including fellow members of their "sub-groups").

In condition 1 subjects employed significant mean in-group favouritism and significant mean intergroup fairness, thereby replicating "standard" minimal group paradigm results. Further, mean in-group favouritism polarized and mean intergroup fairness depolarized from condition 1 to conditions 2 and 3. Thus, it appears that social categorization combined with intragroup communication accentuates in-group favouritism and attenuates intergroup fairness. However, these sample-

mean results conceal considerable variation at the sub-group and the individual levels of intergroup behaviour.

As in the previous two studies, substantial numbers of minimal group subjects did not employ in-group favouritism in condition 1. Several other individual strategies were adopted, including fairness and randomness. Once again, the conclusion must be that social categorization per se is insufficient to engender in-group favouritism, at least at an individual level.¹⁷

More interestingly, in-group favouritism did not polarize from condition 1 to conditions 2 and 3 in all of the eight sub-groups, and not all sub-groups engaged in significant in-group favouritism in these two latter conditions. Thus, social categorization plus intragroup communication was not sufficient to engender in-group favouritism, either at the individual or the sub-group levels, any more than social categorization per se was. Indeed, it is argued in the discussion section of Chapter 4 that the best account that can be given of (individual, sub-group and sample) behaviour in conditions 2 and 3 is a normative one. Thus, rather than attempting to explain group behaviour in these latter conditions in terms of the "structural variables" of social categorization and intragroup communication per se, it is argued that the content and the processes of that communication need to be examined to ascertain (i) whether sub-group norms were consensually agreed and adopted, and (ii) what those sub-group norms were.

The argument that is beginning to be developed is clearly moving some way away from the simplicity of Tajfel's social identity theory, in as much as it is claimed that to understand, explain and predict genuine group behaviour

¹⁷ The same is true at a sub-group level, as particular sub-groups in condition 1 also did not employ significant mean in-group favouritism.

(including intergroup discrimination) it is necessary to consider much more than a simple desire to attain satisfactory self-esteem via positive in-group distinctiveness following social categorization. In Chapter 5 it is nevertheless claimed that social identity theory's motivational hypothesis has been both misrepresented and inadequately tested. Before rejecting that "self-esteem hypothesis" in favour of an alternative one, then, it behoves us to determine and test the best version of that hypothesis possible.

Abrams & Hogg (1988: 320) identify two corollaries of social identity theory's self-esteem hypothesis. The first is that "successful intergroup discrimination...elevates self-esteem" and the second is that "depressed or threatened self-esteem promotes intergroup discrimination" (Hogg & Abrams, 1990: 33). It is argued in this thesis that there are two primary faults with such a formulation. The first has already been implicitly dealt with in the full exposition of Tajfel's social identity theory above. This is that successful in-group favouritism will only elevate self-esteem under very specific conditions (e.g. when in-group favouritism is successfully used to obtain or enhance, rather than to maintain or protect, secure and legitimate positive in-group distinctiveness), and, similarly, self-esteem needs will only motivate in-group favouritism under certain circumstances (e.g. when such discrimination is perceived as likely to achieve secure and legitimate positive in-group distinctiveness and more attractive means of meeting self-esteem needs are unavailable). These issues are more fully explored in Chapter 7 below.

Chapter 5 is concerned with addressing the second main problem with Abrams & Hogg's formulation of the corollaries of the self-esteem hypothesis. This is that they do not adequately specify that it is not self-esteem per se which motivates and is affected by in-group favouritism: it is *social identity contingent self-esteem*. Self-esteem, as the term is usually

employed, operationalized and measured, refers to an individual and global psychological trait. That is, individuals are held to have unique and relatively stable levels of self-esteem which is derived from their "overall" evaluation of themselves, and such self-esteem levels vary across individuals. However, this "general" "trait" self-esteem is not the sort of self-esteem which social identity theory's self-esteem hypothesis refers to.

The self-esteem hypothesis within Tajfel's social identity theory is concerned with "state" self-esteem derived from "specific" group membership in particular intergroup situations. In-group favouritism which is successful in achieving or increasing (or making more salient) secure and legitimate positive in-group distinctiveness is predicted to elevate that part of in-group members' self-esteem which is at that moment (commonly) determined by membership of the successfully discriminating in-group. Similarly, self-esteem which is threatened or depressed because of membership of a particular in-group which is momentarily not securely and legitimately positively distinct is predicted to promote in-group favouritism (under certain circumstances) because of a need to make that (common) aspect of in-group members' self-esteem more satisfactory.

Chapter 5 reports a study in which Rosenberg's (1965) self-esteem scale was modified to produce four measures of different "types" of self-esteem: one measuring "general trait" self-esteem, one measuring "general state" self-esteem; one measuring "trait" self-esteem derived from a particular (national) group membership (i.e. a "specific trait" self-esteem measure); and one measuring "state" self-esteem derived from that particular (national) group membership at a particular moment (i.e. a "specific state" self-esteem measure). It was predicted that only the last of these measures would be sufficiently sensitive to capture changes in self-esteem resulting from making positive or negative aspects

of a particular group membership more salient than they had been a few moments before. This hypothesis was supported. It was also predicted that there would be a positive correlation between subjects' level of identification with the particular in-group of interest (their self-defined national group) and the change in self-esteem resulting from the experimental manipulation. This hypothesis was not supported. The main conclusion drawn in this chapter was that the most appropriate measure for testing the self-esteem hypothesis within social identity theory is a "state" and "specific" one which measures self-esteem derived from a particular group membership at a particular moment (i.e. an SSSE measure), and that because no studies to date appear to have employed such a measure social identity theory's self-esteem hypothesis has not yet been adequately tested.

Chapter 6 reports a study which compares predictions derived from social identity theory with predictions derived from Tesser's self-evaluation maintenance model (Pleban & Tesser, 1981; Tesser, 1980, 1984, 1986, 1988; Tesser & Campbell, 1980, 1982, 1983; Tesser & Collins, 1988; Tesser & Paulus, 1983; Tesser & Smith, 1980; Tesser et al, 1984, 1988) concerning the effects of between-individual and between-group social comparison outcomes.

Social identity theory claims that when social identity is salient social comparison outcomes which reveal new or enhanced (legitimate and secure) positive in-group distinctiveness will raise self-esteem, and social comparison outcomes which reveal new or "enhanced" (legitimate and secure) negative in-group distinctiveness will lower self-esteem. It also makes the "dynamic" predictions that when social identity is salient subjects will: (i) wish to aid the performance of in-group others; (ii) become increasingly close to in-group others when the in-group is (legitimately and securely) positively distinct, but will distance themselves from in-group others when the in-group (legitimately and

securely) negatively distinct; and, (iii) increase the relevance of comparison dimensions when the in-group is (legitimately and securely) positively distinct on those dimensions, but will decrease the relevance of comparison dimensions when the in-group is (legitimately and securely) negatively distinct on those dimensions. Social identity theory, as an explicitly intergroup theory, is relatively "silent" about the effects of between-individual social comparison outcomes.

The self-evaluation maintenance model predicts that personal inferiority to another on a self-relevant dimension will deflate self-esteem while personal inferiority to another on a self-irrelevant dimension will elevate self-esteem, and that the closer one is to the comparison other, the greater the changes in self-esteem will be. The model also makes the "dynamic" predictions that following upward social comparisons on self-relevant dimensions individuals will wish to: (i) avoid increasing the performance differential in the other's favour; (ii) distance themselves from the superior other; and, (iii) reduce the self-relevance of the comparison dimension. The model is relatively silent about: (i) the effects of downward individual social comparisons; (ii) the "dynamic" consequences of upward social comparisons on self-irrelevant dimensions (apart from wishing to increase closeness to the superior other); and, as an explicitly interpersonal theory, (iii) the effects of between-group social comparison outcomes.

Unadulterated, social identity theory received considerably more support than the self-evaluation maintenance model from the study reported in this chapter. Not only are social identity theory's predictions more exhaustive than the self-evaluation maintenance model's, most of the effects of social comparison outcomes (ostensibly between-individual as well as explicitly between-group) could be explained in terms of between-group status differentials.

It was nevertheless argued that social identity theory could be profitably modified to explicitly incorporate predictions about: (i) the effects of between-individual social comparisons; and, particularly, (ii) the interactive effects of between-individual and between-group social comparison outcomes.

Chapter 7 reports the results of a questionnaire designed to re-examine some of the issues raised above and to "mop up" a few outstanding ones. The issue of different "sorts" of in-group favouritism was the first to be examined, in conjunction with an argument that social identity theory also needs to differentiate different "sorts" of out-group. It was argued that the term "out-group" should be restricted to those groups against which in-groups have a particular interest in employing "strong" discrimination, either because of an in-group "culture" of such out-group derogation, or because of situational factors which promote such an in-group attitude. Groups against whom in-groups have no such orientation should be referred to simply as "non-in-groups". Against these groups it was predicted that in-group members would employ only "weak" discrimination (i.e. in-group preference without out-group derogation). Each of these predictions was supported, suggesting that (recognized, multigroup) social categorization may lead only to in-group preference without out-group derogation, and that "something else" needs to be added to the equation before "strong" intergroup discrimination/social competition can be predicted; this "something else" being an in-group norm dictating such behaviour.

The next issue to be addressed was the relationship between positive or negative in-group distinctiveness and in-group evaluation, social identity and self-esteem. Social identity theory claims that the mere fact of in-group superiority or inferiority more or less determines these factors, but it was argued in this chapter that this is far too simplistic. In

particular it was argued: (i) that in-groups (and membership of them) can be, and often are, evaluated by means other than direct social comparisons with other groups, such as via comparing the in-group against "objective" or temporal criteria; and, (ii) that progress toward or away from desired in-group states (of which superiority to other groups is only one) affect in-group evaluations, social identity and social identity contingent self-esteem, as well as the mere attainment or non-attainment of such in-group goals. These propositions received empirical support, suggesting that social identity theory needs to incorporate a more sophisticated account of the determinants of in-group evaluation, social identity and self-esteem.

The self-esteem hypothesis within social identity theory was the final main issue examined, this time from the standpoint that current formulations and tests of that hypothesis are inadequate because they neglect the importance of the stability and legitimacy of valued in-group distinctiveness. With respect to corollary 1 of that hypothesis it was argued that it is not intergroup discrimination per se (or even "successful" intergroup discrimination) which elevates self-esteem, it is new or enhanced legitimate and stable positive in-group distinctiveness, however such in-group distinctiveness is brought about. This proposition received considerable empirical support.

With respect to corollary 2 of the self-esteem hypothesis it is argued that it is not depressed or threatened self-esteem per se which promotes intergroup discrimination, it is depressed or threatened social identity contingent self-esteem, and then only when the proposed intergroup discrimination is perceived as an available and attractive route to secure and legitimate positive in-group distinctiveness. This proposition received mixed and largely post hoc support. It was argued that this was because the notion of legitimate intergroup discrimination was not

adequately incorporated into the questionnaire's design. As in Chapter 5, it was concluded that to date social identity theory's self-esteem hypothesis has not been adequately formulated or tested.

The final chapter begins with a summary of the main empirical results reported earlier in the thesis, and of the principal conclusions drawn from those results. The implications of those results and conclusions for social identity theory are considered, as are the potential theoretical and methodological limitations associated with them. Some attempt is then made to sketch a "modified" version of social identity theory which is consistent with contemporary research (including that presented in this thesis), much of which seems to present challenges to various aspects of the original theory. The main similarities and differences between the "modified" and the "original" versions of social identity theory are then highlighted, and the chapter ends with suggestions for future research in this area.

CHAPTER 2: STUDY 1 - THINKALOUD IN THE MINIMAL GROUP PARADIGM

CHAPTER OVERVIEW

This chapter reports a study in which a "thinkaloud" procedure (cf. Ericsson & Simon, 1984) was employed within the minimal group paradigm (cf. Tajfel et al., 1971). Subjects who were categorized into one or the other of two novel and explicitly arbitrary categories were instructed to make point allocations to in-group and out-group others whilst being tape-recorded verbalizing their thoughts and feelings.

Unlike in most minimal group studies, subjects in the present study did not engage in mean sample-level in-group favouritism. Indeed, subjects in the present study employed a wide variety of individual allocation strategies, with only 3 (of 22) subjects engaging in significant amounts of in-group preference. Unless the thinkaloud procedure interfered in an important way with processes which usually occur in the minimal group paradigm (a possibility rejected in this chapter), this suggests that mere social categorization is not sufficient to promote intergroup discrimination.

Neither did all socially categorized subjects in the present study indicate that they identified with their imposed minimal groups once they had identified themselves as group members. This suggests that mere social categorization (including awareness of group membership) may also be insufficient to promote identification with in-groups.

Finally, because a wide variety of strategies were employed in the present study, both across individuals and within individuals across allocations, it is suggested that no current theory purporting to explain minimal group behaviour is sufficiently broad and flexible to do so.

INTRODUCTION

The "sufficiency condition" within social identity theory (see Chapter 1) states "that the mere perception of belonging to two distinct groups - that is social categorization per se - is sufficient to trigger intergroup discrimination favouring the in-group" (Tajfel & Turner, 1979: 38). Justification for this claim comes from the results of experiments within the minimal group paradigm (Billig & Tajfel, 1973; Brewer, 1979; Tajfel 1970; Tajfel & Billig, 1974; Tajfel et al., 1971). In this paradigm explicitly arbitrary or random criteria are used to privately assign each experimental subject to one of two (or more) novel categories and then subjects are asked to individually distribute "goods" (e.g. points, money or evaluations) to pairs of other people about whom they know nothing except their group membership. Results typically show that, on average, subjects allocate more goods to in-group than to out-group members and that, again on average, subjects have a preference for in-group members receiving *more* goods than out-group members, as opposed to in-group members receiving the maximum possible goods. That is, mere social categorization seems sufficient to promote intergroup discrimination, just as claimed by the "sufficiency condition" within social identity theory.

Prior to this experimental discovery, the dominant social-psychological theory of intergroup conflict came from Sherif, who hypothesized that intergroup conflict stems from group members' perceptions of negatively interdependent group goals (e.g. Sherif et al, 1961).¹ If, as seems reasonable, social categorization per se does not entail perceptions of negatively interdependent group goals, then mere social categorization being sufficient to promote intergroup discrimination means that perceived negatively interdependent

¹ Two groups' goals are negatively interdependent if movement by one group toward its goals necessarily puts the other group's goals further out of its reach.

group goals cannot be a necessary precondition of intergroup conflict.

A simplified account of Tajfel's social identity theory explains subjects' minimal group behaviour as follows. Subjects recognize the imposed social categorization and their individual positions as members of only one of the two resultant social categories. Subjects' self-esteem then becomes dependent on their evaluation of their own social category (i.e. their "in-group") and their evaluation of themselves as members of that social category (i.e. their "social identity"). Groups are evaluated in comparison to other groups, with superior groups being evaluated positively and inferior groups negatively. Subjects in the minimal group paradigm therefore indulge in intergroup discrimination in an attempt to achieve in-group superiority, positive evaluation of the in-group and its members, and thus positive social identity and self-esteem (and/or to avoid in-group inferiority, negative evaluation of the in-group and its members, and thus negative social identity and self-esteem).

Competing accounts of minimal intergroup behaviour come from generic norm theory, equity theory, the behavioural interaction model, the intergroup accentuation hypothesis, self-categorization theory, and a demand characteristics explanation.

Generic norm theory (e.g. Branthwaite et al., 1979; Tajfel, 1970; Tajfel & Billig, 1974; Tajfel et al, 1971; Wetherell, 1982) holds that within all societies there are "generic" norms which members of those societies learn, internalize, and act upon in appropriate circumstances. Common to most societies are generic norms of "fairness" and of "groupness". A societal norm of fairness means that (when appropriate) society members should distribute goods equally among all (equally deserving) recipients and a societal norm of groupness means that (when appropriate) group members should

favour in-group members over out-group members. Both norms are said by generic norm theory to be activated within the minimal group paradigm: the former because goods are being distributed among people who equally deserve (or do not deserve) goods, and the latter because they are being distributed between in-group and out-group members. Thus, mere social categorization results in intergroup discrimination because it activates the groupness norm, but discrimination is rarely maximal as the task of distributing goods among "equal" others simultaneously activates the fairness norm. This means that generic norms of fairness and groupness are simultaneously working in opposite directions within the minimal group paradigm, so that the resultant behaviour is best understood as a "compromise" between fairness and in-group favouritism.

Equity theory asserts simply that people desire equitable outcomes whenever goods are distributed. An equity theory explanation of minimal group behaviour (e.g. Ng, 1981, 1986) rests on the proposition that subjects display in-group favouritism in an attempt to compensate for similar discrimination by out-group members so that an equitable outcome is achieved overall.

The behavioural interaction model (e.g. Rabbie, 1992; Rabbie et al., 1989) asserts that subjects in minimal intergroup situations are ultimately motivated by "economic" self-interest, striving to obtain maximum possible "goods" for themselves (as individuals) via instrumental intragroup cooperation and instrumental intergroup competition. Subjects in the minimal group paradigm aim to further the goals of in-group members rather than out-group members in the hope and belief that in-group members will be more likely than out-group members to reciprocate in order to further their own selfish economic interests. This belief is held because the social categorization results in more perceived "common fate" with in-group than with out-group members, which in turn leads

subjects to feel more positive interdependence with the former rather than the latter. (Two individuals are positively interdependent if each can only move toward or away from their own goal if the other does likewise.) In-group favouritism is not maximal, however, because subjects realise that their self-interest can be furthered by out-group members' allocations as well as in-group members', even if out-group members are deemed less likely than in-group members to engage in reciprocal goal satisfaction.

The intergroup accentuation hypothesis holds that whenever social categories are superimposed on people who were considered previously only as individuals there will be a tendency for those people to accentuate inter-category differences (and also, perhaps, intra-category similarities). Applied to the minimal group paradigm (e.g. Doise & Sinclair, 1973) this hypothesis suggests that social categorization has the effect of making subjects wish to behaviourally accentuate intergroup differences, which they do via intergroup discrimination.

Self-categorization theory (e.g. Turner et al., 1987) claims that people's actions as group members are guided by a desire to be as similar as possible to a "prototypical" in-group member. The prototypical in-group member is one who maximizes the ratio of difference to out-group members over similarity to other in-group members. Applied to the minimal group paradigm, self-categorization theory suggests that subjects imitate the expected non-maximal in-group favouritism of the prototypical in-group member.

Finally, the demand characteristics explanation of minimal intergroup discrimination (e.g. Gerard & Hoyt, 1974; Schiffmann & Wicklund, 1992) argues that characteristics of the minimal group situation virtually demand intergroup discrimination. Minimal group subjects are held to have nothing to base their allocation decisions on except their

group memberships and it is supposed that it seems obvious to them that they should favour the in-group, otherwise why would the experimenter bother to put them into one of the groups?

Numerous experimental studies have been conducted to try and substantiate and/or undermine the theoretical positions above, but very little, if anything, has been done in the way of obtaining subjects' accounts of their behaviour in the minimal group paradigm. The study reported here attempts to remedy this. Subjects took part in a standard minimal group paradigm experiment with one major modification: subjects were asked to "think aloud" throughout and were tape-recorded doing so.

Ericsson & Simon (1980, 1984) suggest that verbalizing thought affects cognitive processes (and therefore, presumably, behaviour) only if the thoughts verbalized would not have been attended to if they were not verbalized. That is, subjects who think aloud will think and act in exactly the same way as subjects who do not think aloud, unless the thinking aloud itself causes subjects to attend to thoughts which they would not have attended to were they not thinking aloud.

If it can be assumed that thinking aloud in the minimal group situation does not cause subjects to attend to thoughts they would not otherwise have attended to, such a procedure offers a useful research device for investigating why subjects act as they do within the minimal group paradigm.² Each theoretical explanation of minimal group behaviour suggests that various motives, means and ends might be accessible to consciousness by subjects during the allocation of goods. According to all of the theories, for example, subjects should be aware of the social categorization and should in some way accept their

² This is an assumption, however, and there is also some debate as to the adequacy of the thinkaloud method generally. See Nisbett & Bellows, 1977; Nisbett & Wilson, 1977; Vesper & Voss, 1985. Nevertheless, given the present state of understanding about minimal group behaviour, the method certainly seems worth trying.

membership of one group and their non-membership of the other(s). Beyond this, of course, each theory would lead one to expect subjects to vocalize somewhat different thoughts.

If social identity theory is valid, for example, one might expect subjects, as well as identifying themselves as members of one of the minimal groups but not of the other, to demonstrate some sort of identification with the in-group and its members which would lead them express a concern that members of the in-group should end up receiving more goods than members of the out-group.³ Whether or not subjects would be aware of the direct and indirect instrumental reason for this desired in-group superiority (i.e. positive evaluation of the in-group and positive self-esteem, respectively) is less certain.

According to generic norm theory we might expect subjects to be aware of a "conflict" between a desire to be fair and another desire to be loyal to the in-group. Again, it is less clear whether or not subjects would be aware that these desires stemmed from internalized societal expectations.

Equity theory suggests that subjects thinking aloud during a minimal group paradigm experiment would express a desire for an equitable outcome, as well as a belief that this required intergroup discrimination in order to compensate for expected in-group favouritism by out-group members.

There seems no reason to suppose that subjects would be unaware of their primary motive according to the behavioural interaction model. This model suggests that subjects will be quite clear that they are acting as they are - that is, cooperating with in-group members and competing with out-group members - purely for indirect personal economic gain. We can also expect that subjects will express a belief that in-group

³ See Chapter 1, pp. 41-21, for the distinction between identity with and identity as.

members will be more likely to cooperate with them than will out-group members.

Doise's intergroup accentuation hypothesis is that minimal group subjects discriminate in an attempt to behaviourally differentiate between the social categories. It is perhaps more reasonable to assume that subjects will be conscious of this than to assume that subjects will have no idea how they are acting or why they are doing so.

Self-categorization theory is more problematic. It is not clear according to that theory whether people attend to their cognitive determination, and behavioral emulation, of the prototypical in-group member. It seems unlikely that they would. (Under what situation would one expect an account of a social action to include a reference to prototypical in-group members, by whatever name?) However, even under a self-categorization theory explanation one might still expect some sort of declaration of similarity to in-group members and difference to out-group members and/or an expression of a desire to act like an in-group member and not like an out-group member.

If the demand characteristics explanation is correct, it might be predicted that minimal group subjects would feel that in-group favouritism was pretty much the only option they had and/or was expected of them (forced upon them?) by the experimenter and/or the experimental situation.

Obtaining minimal group subjects' accounts of their behaviour has another potential benefit, quite apart from potentially distinguishing between the validity of the various theoretical positions expounded above. This is to investigate claims that the principal dependent measures used in the minimal group paradigm, i.e. the Tajfel matrices, are susceptible to some of the more potentially damaging criticisms that have been made of them.

The first criticism come from Aschenbrenner & Schaefer (1980: 396) who suggest that the pull-scores derived from the Tajfel matrices may not "provide a representative description of the subjects' behaviour."⁴ So, for example, if on a particular occasion a subject chooses to allocate 15 points to an in-group member and 17 points to an out-group member (on matrix type 2 in Appendix 1), this will contribute +4 (from a possible range of 0-12) to a pull score of MD on MIP+MJP (Maximum Difference in favour of the in-group on Maximum In-group Profit+Maximum Joint Profit, the emphasis should be noted). It is difficult to see how giving more points to an out-group member than to an in-group one can justifiably contribute so positively to such a score. Nevertheless, obtaining subjects' accounts may throw some light on how subjects themselves construe such a choice.

Mummendey et al. (1992) unusually asked subjects to use Tajfel matrices to distribute "bads" rather than "goods" and found no significant in-group favouritism. This suggests that the Tajfel matrices may lack content validity,⁵ such that subjects are forced into making choices from a relatively narrow range of possibilities. As Tajfel himself notes, giving less goods to out-group members than to in-group members may have a very different psychological significance than actively taking away goods from (and/or giving "bads" to) out-group members (Tajfel et al, 1971: 174). Subjects' accounts may reveal whether subjects are simply being "nicer" to in-group than to out-group members (as the behavioural interaction model might suggest) or are being as "nasty" as they can to out-group members within the limited range of behaviours available to

⁴ See Appendix 1 for the Tajfel matrices most often used in the minimal group paradigm, and also for how pull-scores are derived from them.

⁵ "Content validity refers to the degree to which a measure covers the range of meanings included within the concept" (Babbie, 1989: 125), the concept being, in this case, intergroup behaviour.

them (as the intergroup accentuation hypothesis might suggest).

Bornstein et al. (1983a: 342) note that some strategies are structurally mutually inhibitory using the Tajfel matrices, while others are operationally entailed. If a subject wishes to maximize the overall goods obtained by recipients on a matrix which measures MIP+MD on MJP (Maximum In-group Profit+Maximum Difference in favour of the in-group on Maximum Joint Profit), for example, they are forced to be as inequitable as it is possible to be on that particular allocation. It is possible that subjects may wish to pursue some strategies but do not wish to do so at any cost. Thus a 'pull' toward MIP+MD from MJP may be an artefact of subjects wishing to pursue both MJP and fairness. Subjects' accounts should throw some light on such matters.

A related criticism by Bornstein et al. (1980: 342) is that subjects may be forced into 'second-choice' strategies by the constraints imposed by the design of the Tajfel matrices. On one matrix type (i.e. Type 3), for example, all possible allocations "maximize" (and/or "minimize") the joint goods obtained by the recipient pair. What does a subject do if they are attempting to maximize joint profit in such a situation, and does it affect their subsequent actions? Again, subjects' accounts should be illuminating.

Overall then there seems to be a strong case for obtaining subjects' accounts of what they think they are doing in the minimal group paradigm, what they hope to achieve by acting as they do, problems they encounter, and what they do about such problems. Specifically, such accounts should enable each of several theoretical explanations of minimal group behaviour to be examined from the subjects' points of view, and they may also suggest whether or not certain key criticisms of Tajfel-matrix dependent measures are valid.

METHOD

Participants

26 second-year psychology students from Keele University volunteered to be subjects in a study which was introduced as "A qualitative examination of decision-making." 4 subjects were excluded from the analysis: one because they showed signs of being familiar with the minimal group paradigm, two because their audio-tapes were blank (see below), and one because their audio-tape could not be located. Of the remaining 22 subjects 7 were male and 15 female. Ages ranged from 19 to 30 years of age, with a mean average of 21.5 years (mode = 20 years, median = 20.5 years).

Procedure

Experimental sessions of 2-4 subjects were conducted over four consecutive days according to a time-table filled out at the time of volunteering. Upon arriving at an experimental session (except in the first session) subjects were immediately asked whether they had heard anything about the study from others who had already taken part. No-one claimed to have heard about the study, and several said that previous participants had refused to talk about the study, even when asked about it.

Subjects were then told that each of them would soon be given a unique code number and would be randomly categorized into one of two groups, Group W or Group X. All subjects allocated to Group W would have a code number in the 40s and all subjects allocated to Group X would have a code number in the 70s. Allocation to groups, the subjects were told, would be carried out by the experimenter selecting at random one envelope for each subject from a collection of such envelopes, each one of which contained a unique code number and social category membership information. At this stage one subject

checked that the envelopes did indeed contain different code numbers and the possibility of being assigned to either group. Another subject then "shuffled" the envelopes.

Next, subjects were informed that the nature of the decision-making task they would be required to perform involved allocating points to pairs of other people about whom they would know nothing except their individual code number and their group membership. It was stressed that they would not know who was in which group and that each subject would make their allocations individually in separate rooms. It was particularly noted that subjects were seen in separate sessions only because of limited availability of rooms and that there was no more likelihood of people in their session being in the same group as them (or in the other group) than there was of people in other sessions being in their group (or in the other group).

Subjects were told that they would never be allocating points to themselves, but that points allocated to them by others would be recorded (hence the need for individual code numbers) and that money they received for participating would be directly proportional to the number of points that they accrued. Although no direct points/money exchange rate was mentioned it was made clear that the more points a person had allocated to them, the more money they would receive at the end of the week.

Subjects were told that the main purpose of the study was to discover the basis upon which they made their decisions. To this end consent was gained from subjects for them to be tape-recorded whilst giving a running commentary on everything they were thinking throughout the procedure, no matter how seemingly trivial or unimportant. (No subject withheld their consent.) It was emphasized that the experimenter had no preconceptions concerning how people would make their decisions, but that he was genuinely interested in finding out. He was

particularly interested, they were told, in what they were trying to do, why, any problems they encountered, what they did about any such problems, and what they thought the outcome of their decisions might be. Nevertheless, it was stressed that the experimenter was interested in anything and everything subjects were thinking.

Use of the Tajfel matrices was demonstrated and all questions were answered to the best of the experimenter's ability. The experimenter then 're-shuffled' the envelopes and randomly selected one envelope for each subject.

Upon receiving their envelopes subjects were directed to individual rooms where they were instructed to open their own envelope and follow the instructions contained within. Inside each envelope was a booklet containing (as well as a unique code number and group membership information) a repeat of most of the verbal instructions the subjects had received - including worked examples of how to use the Tajfel matrices, a set of 18 Tajfel matrices, a short demographics questionnaire, and a blank page inviting written comments.⁶

Three matrix types were used, copies of which can be found in Appendix 1. Each matrix type was presented six times; twice with two in-group recipients, twice with two out-group recipients, and twice with one recipient from each group. For each recipient-pair on each matrix type one recipient appeared on the top-row on the first presentation and on the bottom row for the second presentation. Each matrix appeared on a separate page. In an attempt to minimize or avoid possible

⁶ Following the Tajfel matrices were a number of "allocation grids" the use of which had also been explained to the subjects. These were included merely as a pilot for a later study (see next chapter) and they will not be discussed further. An example of the response booklets subjects received (showing, for reasons of space, only one of the eighteen allocation matrices, and none of the allocation grids) can be found in Appendix 2.

order-effects, matrix types and recipient-pair patterns were presented in a different sequence for each subject.

Tape-recorders were switched on by subjects in accordance with printed instructions to do so. The instructions also stressed the need to think aloud as much as possible, an exhortation that was repeated on each page of the response booklet. In addition, subjects were occasionally instructed to read printed matter aloud. The first such instruction was for subjects to read aloud their unique code number and group membership.

After completing all eighteen matrices subjects were invited to make any comments (written or spoken) they wished concerning the experiment they had just taken part in, and were then instructed to turn off the tape-recorder and return their response booklets and audio-tapes to the experimenter. He then answered any questions they had (within the limits of the experimental procedure, i.e. that did not involve explanation of possible or expected results); told them they would be fully debriefed and paid at the end of the week, and extracted a promise from them not to tell anyone yet to do the study anything about what they had been asked to do (except in the final session). (No subject declined to make such a promise.) At the end of the week the experimenter fully debriefed the subjects (in a class given by him); fully answered any questions; and gave each subject £2, regardless of points allocated to them in the experiment. Subjects were also invited to make any complaints they might have to any one of several people (e.g. their personal tutor). No such complaints were made to the experimenter's knowledge.

Overall, the design of the study was that of a 'standard' minimal group paradigm experiment, with the main modification of a think aloud procedure. In line with Billig & Tajfel (1973) social categorization was explicitly (and genuinely) random in order to investigate the effects of social

categorization per se on subsequent intergroup behaviour, unconfounded by any possible perceptions of similarity subjects might have perceived they had with in-group others (other than that resulting from social categorization, see Farsides, 1993a). Social categorization was also carried out in such a way that subjects could not possibly feel they had any control over it (see Langer, 1975).

Analysis

Tajfel matrices were used to calculate pull-scores and indirect in-group favouritism scores (see Appendix 1 for details). Seven of the audio-tapes were then transcribed in full. Subjects were selected to have their tape transcribed in full if (i) they obtained relatively high pull-scores on one or more strategy and (ii) their tapes were particularly clear and therefore could be accurately transcribed with ease. The rest of the audio-tapes to be used (15) were carefully listened to by the experimenter and all material deemed by him to be relevant to the task in hand was also transcribed. All transcribed material can be found in Appendix 3.

A coding frame was designed and the qualitative data was subjected to a series of content analyses. The first sections of the coding frame each used a single matrix presentation as the coding unit. A third section involved the experimenter giving a short qualitative précis of each subject's "overall" behaviour and then using each such summary as a single coding unit. The coding frame is reproduced in Appendix 4. In the first part (Section A: "allocation intentions") eleven primary categories were generated from the relevant research literature (see introduction) and each subject's expressed aim during each allocation was assigned to the appropriate category. If a subject's account of their behaviour seemed to significantly invoke two or more categories a new combinational category was generated and incorporated into the

coding frame. The primary categories were:- personal profit (subject claims reason for allocation is to obtain personal goods); personal status (subject claims reason for allocation is to obtain more goods than other subjects); in-group profit (subject claims reason for allocation is to obtain goods for the in-group and/or its members); in-group status (subject claims reason for allocation is to obtain superiority for ingroup and/or its members over the out-group and/or its members); joint profit (subject claims reason for allocation is to obtain goods for all recipients); fairness (subject claims reason for allocation is to be fair, just and/or equitable); intergroup accentuation (subject claims reason for allocation is to emphasize, exaggerate or extend intergroup differences and/or intragroup similarities); prototypicality (subject claims reason for allocation is to mimic an in-group prototype); none or arbitrary (subject explicitly states that there is no rational or teleological reason for allocation); other (subject gives a specific reason or rationale for allocation other than those contained in other categories); and, absent (subject gives no explanation for their allocation).

The second principal part of the content analysis (Section B: "operative factors") involved marking the presence or absence of a subject's reference to one or more of eleven factors which the various theories considered in the introduction to this chapter might suggest would be at work during the minimal group paradigm.⁷ These factors were:- personality (subject claims allocation motivated or determined by dispositional characteristic); experimenter expectation (subject claims allocation made because it was expected by the experimenter); norm(s) (subject claims that allocation is in line with one or more norms appropriate to the situation); reciprocity (subject claims they are making their allocation in the hope and/or

⁷ Thus, Section A concerned what subjects reported they were trying to do on each allocation, and Section B concerned the reasons subjects gave for doing it.

expectation that others will reciprocate); in-group identity (subject identify with the group in expressing a commitment to and/or a concern about the fate of the in-group and/or its members); alternation (subject claims that the allocation is made in part to redress earlier actions and thereby pursue a single coherent 'overall' strategy); consideration of others' actions (subject wonders about and/or predicts other subjects' behaviour); overall outcome (subject considers likely outcome of all the subjects' allocations, e.g., how many goods the in-group will receive relative to the out-group); structural 'restriction' (subject indicates that they type of allocation they choose is directly affected by the matrix type or a change in matrix type); self-esteem (subject claims that a motivation for their choice is psychological well-being); and, personal gain (subject expresses a belief that their behaviour is consonant with a strategy whereby they will personally gain economically from their participation in the experimental procedure).

The third aspect of the content analysis (Section C: "overall strategy") involved the experimenter writing a qualitative summary of each subject's "overall" behaviour and then deciding which of the eleven categories (plus any combinational categories generated from them) used in the "subject's intentions" section would be most suitable as a single categorical summary of each subject's "overall" strategy. The experimenter's qualitative summaries are shown in full in Appendix 5.

RESULTS

Sample pull and indirect in-group favouritism scores

Average pull-scores and indirect in-group favouritism scores for the 22 subjects included in the analysis were as follows:-

<u>Matrix-score^a</u>	<u>Mean</u>	<u>(SD)</u>
MIP+MJP on MD	5.591	(5.738)*
F on FAV	4.273	(5.742)*
MJP on FAV	2.000	(6.133)
FAV on F	0.909	(5.089)
MD on MIP+MJP	-0.136	(4.223)
FAV on MJP	-0.636	(5.892)
IndFAV1	1.250	(2.939)
IndFAV2	1.114	(4.149)

N = 22. Asterisked items are significantly different from zero ($p < 0.005$) by two-tailed one sample t-test.

Table 2.1: Mean pull and indirect in-group favouritism scores

The present study did not replicate the usual minimal group paradigm experiment in that social categorization did not lead here to any significant mean sample-level in-group favouritism. Only two scores were significantly different from zero: the pull of MIP+MJP on MD and the pull of F on FAV.

^a See Appendix 1 for scoring rules.

Reliability coefficients for the content analysis

Once the coding frame had been produced two coders used it separately to content analyze the qualitative material. Differences between the coders were noted and returned to them. Coders then re-examined the material to see if they had made a simple mistake which they wished to correct (without knowing how the other coder had coded each particular item). Once this had been done inter-rater reliability checks were conducted using Krippendorff's (1980: 138-140) agreement coefficient. This measure provides a "uniform measure that is comparable across numerous situations" as (i) it "corrects for small sample sizes" and (ii) the percentage of expected agreement between raters "is computed from the proportion with which a category is used, both coders taken together" (p. 138). The agreement coefficient (when multiplied by 100) shows the percentage by which observed agreements between coders are above chance.

For the first part of the content analysis (i.e. Section A: the "allocation intentions" section) the agreement coefficient was 0.896. That is, agreement between the two coders was almost 90% above that which would be expected by chance. The second part of the content analysis (i.e. Section B: the "operative factors" section) required a separate agreement coefficient to be calculated for each factor considered. These are shown below.

<u>Factor</u>	<u>Agreement coefficient</u>
Personality	0.564
Experimenter expectation	0.004
Norm(s)	0.665
Reciprocity	0.856
In-group identity	0.816
Alternation	0.805
Consideration of other's actions	0.831
Overall outcome	0.005
Structural restriction	0.706
Self-esteem	0.005
Personal gain	0.419

Table 2.2: Interrater agreement coefficients

Agreement coefficients were high for most factors, less good for factors concerning personality and personal gain, and poor for experimenter expectation and self-esteem.⁹

No agreement coefficient was calculated for the third section of the content analysis (i.e. "overall behaviour"), as only the experimenter completed and coded qualitative summaries of each subject's behaviour overall.

⁹ It can be noted that, generally speaking, the fewer the instances of an operative factor in the present study, the worse the agreement coefficient. Thus, experimenter expectation and self-esteem were cited as operative factors very seldom by subjects in the present study.

Once reliability had been checked coders discussed how to revolve any remaining differences. On the very few occasions when agreement could not be reached the experimenter decided on the appropriate coding for reporting the results below.

Content analysis concerning allocation intentions

Nineteen categories were used to represent allocation intentions, eleven 'primary' and eight 'combinational.' The number of subjects claiming those allocation intentions, and the number of allocations claimed by subjects to have been motivated by each intention, are shown in the table below (continued, with a key, overleaf), broken down by matrix type and recipient pattern. In the table the top number in each cell represents the number of that type of allocations made and the bottom number represents the number of subjects involved in making those allocations. Cells left blank indicate that no allocations were made following that strategy.

Matrix	Type 1			Type 2			Type 3		
	ii	oo	io/ oi	ii	oo	io/ oi	ii	oo	io/ oi
Strategy									
1 Personal Profit								1 1	
2 Personal Status									
3 In-group Profit			3 2	3 2		3 2			2 2
4 In-group Status		1 1	2 2		1 1	1 1			3 3
5 Joint Profit	5 4	5 4	4 4	7 6	5 3	6 5			
6 Fairness	9 8	5 4	10 8	5 3	5 5	4 3	12 8	12 9	3 2

7 Intergroup Accentuation		1 1							2 2
8 Proto-typicality									
9 Random/None	10 7	14 11	10 9	12 9	13 10	13 10	17 11	15 13	13 11
10 Other Strategy		4 4	2 2	4 4	2 2	2 2	1 1		
11 Absent Reason	15 11	13 8	14 12	10 6	13 10	11 8	14 10	14 13	18 14
12 1 & 3						1 1			
13 1 & 7						1 1			
14 3 & 6	1 1								
15 4 & 7						1 1			
16 5 & 6		1 1		2 2	1 1	2 1		1 1	1 1
17 5 & 7				1 1	1 1				
18 5 & 8					1 1				
19 6 & 7	3 3				1 1			2 2	1 1

Key:- ii = two in-group recipients; oo = two out-group recipients; io/oi = one recipient from each group; Type 1 = matrix measuring pull of FAV on MJP and vice versa; Type 2 = matrix measuring pull of MD on MIP+MJP and vice versa; and Type 3 = matrix measuring pull of FAV on F and vice versa (see Appendix 1 for details). For intentions:- 1 = Personal profit; 2 = Personal status; 3 = In-group profit; 4 = In-group status; 5 = Joint profit; 6 = Fairness; 7 = Intergroup accentuation; 8 = Prototypicality; 9 = Random, arbitrary or none; 10 = Other intention; 11 = Absent reason. Other intention numbers are combinations of these (See Appendix 4 for details).

Table 2.3: Subjects' self-reported allocation intentions

No reason was given for almost a third of all allocations, with most subjects making at least one allocation without verbalizing what they were doing or why.

For another third of allocations subjects were clearly making random or arbitrary decisions or explicitly claimed not to have any reason for the allocation they made. A vast majority of subjects made at least one such 'intentionless' allocation.

An identical majority of subjects made at least one allocation which they said was motivated by fairness. 16% of all allocations were so motivated: the most common of all stated "rational" intentions. Fairness was however the most variable intention across matrix types and recipient patterns, with fairness far more common on intergroup allocations on matrix type 1 (10 allocations, 8 subjects) than on matrix type 2 (4 allocations, 3 subjects) or matrix type 3 (3 allocations, 2 subjects) during intergroup allocations. For same-group allocations, on the other hand, subjects were more likely to report pursuing fairness during allocations on matrix type 3 (24 allocations, 17 subjects) than on matrix types 1 (14 allocations, 12 subjects) or 2 (10 allocations, 8 subjects).

Half the subjects made at least one allocation which they said was motivated by joint profit, although only 8% of all allocations were explained with reference to such a motive. Bearing in mind that all possible allocations on matrix type 3 yield the same overall total, it is perhaps not surprising that attempts to maximize joint profit were restricted to matrix types 1 and 2 (Type 1: 14 allocations, 9 subjects; Type 2: 18 allocations, 8 subjects; Type 3, no allocations).

Half of the subjects also claimed that at least one of their allocations was motivated by an intention not covered by the coding frame, although a mere 4% of all responses were motivated by such "other" considerations. These "other" intentions were varied and included giving the recipient with the highest code number the highest points (1:10-11)¹⁰; minimum joint profit (1:16, 13:9); minimum (4:11) or restricted (5:18) out-group profit; giving the bottom-row recipient high numbers (10:7); biggest point difference (14:13); giving similar points to similar code numbers (15:9)

¹⁰ The first number in these brackets refers to the subject, the other to the particular allocations of interest made by the subject. See Appendix 3 for details.

and dissimilar points to dissimilar code numbers (15:11); and not upsetting recipients (22:17).

11 allocations were explained by subjects as being motivated by in-group profit intentions, although only 4 subjects reported using such a strategy. In-group profit intentions were, moreover, far more common on intergroup allocations (8 allocations, 6 subjects) than they were on in-group allocations (3 allocations, 2 subjects), with no such intentions being expressed during out-group allocations.

5 subjects reported being motivated by intentions to achieve in-group status, but only for a total 8 allocations. An intention to strive for in-group status was never claimed during in-group allocations and such claims were at least twice as common during intergroup allocations (6 allocations, 4 subjects) as during out-group allocations (2 allocations, 2 subjects).

A combined strategy of maximum joint profit and fairness was cited by 3 subjects over 8 allocations and a combined strategy of fairness and arbitrariness was reported by 5 subjects over 7 allocations. No other intention was mentioned by more than 2 subjects or was reported as operative during more than 3 allocations. Intentions of striving for individual status and/or attempting to mimic an in-group prototype were not mentioned by any subjects.

Overall then the first part of the content analysis suggests that the predominant "rational" strategies subjects used in the present study were (in descending order of importance):- fairness, joint profit, "other", in-group profit, and in-group status. In-group profit and in-group status were more commonly reported on intergroup allocations than on same-group ones, while fairness was more common on matrix type 1 than on types 2 and 3 for intergroup allocations and more popular on

matrix type 3 than on types 1 and 2 for same-group allocations.

Content analysis concerning operative factors

The frequency with which the factors in Section B of the content analysis were operative during the present study are shown in the table below, broken down by matrix type and recipient pattern.

Matrix	Type 1			Type 2			Type 3		
	ii	oo	io/ oi	ii	oo	io/ oi	ii	oo	io/ oi
Personality	2	1	1	2		1			
Expt. Expec.									
Norm						3		1	
Reciprocity	1					2			
Identity	1	1	3	5	3	4			5
Alternation	4	1	4	2	1	2	1	1	1
Others' Beh.	2			1	1	2			
Outcome									
Restriction	6	4	2			2	6	3	4
Self-esteem	1				1	1			
Self Gain	1	1		1	1	3		1	

Key: ii = Two in-group recipients, oo = Two out-group recipients, io = In-group recipient on top-row of matrix and out-group members on bottom-row of matrix, oi = Out-group member on top-row of matrix and in-group member on bottom-row of matrix (see Appendix 1 for details).

Table 2.4: Subjects' self-reported allocation motivations

Personality/disposition was given as a reason for 7 allocations. The dispositions referred to were superstition (1:1+7, 4:10), fairness (13:1, 15:1), generosity (2:5), and insane greediness (14:5).

A couple of subjects wondered if discriminating on behalf of the in-group might be what was expected in the present study (5:7, 11:13). Several more subjects expressed feelings that perhaps they ought to show such discrimination (2:2, 11:11, 21:4, 22:10), while others believed that they should behave fairly (7:3, 10:1). In all of these cases, however, it either was not clear that subjects had actually decided to behave as they thought they "ought", or it was not clear that the "ought" expressed by the subject did indeed stem from a perceived experimenter expectation (as opposed to being expected by the in-group, a norm or whatever). Thus no allocations were considered by the coders to be unambiguously motivated by subjects conforming to perceived experimenter expectations.

Four allocations were explained with reference to conformity to norms. Two norms were rooted in non-minimal in-groups; one suggesting that students should obtain "the most money possible" for "other students" and another that law students should be fair (4:1). One norm revolved around the belief that all members of a minimal in-group would allocate maximum possible points to other in-group members (21:18) and the final norm prescribed that minimal in-group members "should show some loyalty" (4:3).

Reciprocity was only explicitly attempted on three occasions. One subject attempted to be fair to others in the hope that others would be fair to her (12:8), another subject maximized group profit in the hope that others would do the same (14:12), and third maximized in-group profit, again with the hope that others would do likewise with the same hope (21:18). Other subjects expressed a belief or hope that others would act as they did, but did not forward this as a reason for their own actions (e.g. 4:5, 14:1).

Identification with minimal in-groups was indicated by 6 subjects during a combined total of 22 allocations. On

intergroup allocations identification with minimal in-groups was indicated 12 times, 3 times on matrix type 1 (5:11, 21:7, 21:17), 4 times on matrix type 2 (4:3, 21:13+18, 22:9), and 5 times on matrix type 3 (5:4, 10:9, 21:4+10, 22:18). Identification with minimal in-groups was expressed during in-group allocations 6 times, once on matrix type 1 (21:6), 5 times on matrix type 2 (4:13+16, 5:17, 10:10, 21:15), and not at all on matrix type 3. On out-group allocations identification with minimal in-groups was indicated only 4 times, once on matrix type 1 (22:10), and 3 times on matrix type 2 (2:13, 4:11, 5:12).

Use of alternation was reported on 17 occasions by 7 subjects. One subject was responsible for almost half of these reported uses of alternation, employing as she did a rather indirect route to fairness: making a random allocation on the first presentation of a particular recipient pair and then trying to even out any inequalities on the pair's second presentation (12:8,10,12,13 and 15-18). A more common use of alternation was to compensate for inequalities structurally entailed by the subject pursuing a primary strategy of maximum joint profit (2:6, 11:11+12, 19:9+10, see also 4:4). One subject, however, used alternation to compensate for intergroup (10:12) and interindividual (10:13) inequalities he imposed himself during earlier allocations (see also 14:6).

Subjects reported considering others' behavioural strategies on 6 occasions. One subject hoped both that others were acting as fairly as she was (4:5) and that other in-group members were doing the same as her in trying to maximize in-group profit (4:13). Another subject similarly hoped that others were engaging in maximum joint profit as he was (14:1+12). A third subject hoped that others were allocating him more than he was allocating to them (17:14) and a fourth believed that other in-group members would be allocating maximum possible points to all other in-group members, including herself (21:18).

The overall between-group outcome of everybody's allocations was never considered by any subject.

10 subjects felt "restricted" in their allocation behaviour at least once during the study.¹¹ On intergroup allocations restriction was felt 8 times, twice on matrix type 1 (11:1+11), twice on matrix type 2 (1:11, 21:13), and four times on matrix type 3 (8:8, 10:9, 14:3, 21:10). On in-group allocations restriction was claimed 12 times, six times on matrix type 1 (1:7, 2:4, 4:2, 12:4+8, 14:1), six times on matrix type 3 (1:9, 2:8, 8:5, 10:6, 14:10, 19:12), and not at all on matrix type 2. Similarly, on out-group allocations subjects reported restriction 7 times, four times on matrix type 1 (4:10+14, 10:7, 11:8), three times on matrix type 3 (11:10, 14:15+18), and not at all on matrix type 2.

Of the restrictions reported 3 were due to superstition, avoiding "unlucky" allocations which gave 13 points to one or both recipients (1:7, 4:2, 10+14). Two further instances of restriction occurred where subjects wanted to make particular allocations to particular individual recipients and were thereby "forced" to make unintentional allocations to the second recipients (10:9, 21:13). More seriously, many subjects felt that their preferred strategies were structurally frustrated by the design of the matrices. On 8 occasions subjects attempting to maximize joint profit were unable to do so because of the design of matrix type 3 (2:8, 8:5+8, 11:10, 14:10, 15+18, 19:12). On matrix type 1, on the other hand, at least two subjects felt unable to select desired allocations which would maximize joint profit because to do so would entail unacceptable inequality between the two recipients, as maximum joint profit and fairness are structurally negatively interdependent along half the matrix

¹¹ It should be remembered that "structural restriction" is a short-hand label for subjects claiming that their behaviour was affected (i.e. restricted, modified, or whatever) by a matrix design and/or a change in matrix types.

(11:1,8+11, 14:1). Similar strategy frustrations were suffered on at least a further 5 occasions (1:9+11, 10:6, 12:4+8, 21:10). Finally, subjects amended old strategies or adopted new ones as a result of changing matrix types on at least three occasions (2:4, 10:7, 14:3).

Only two subjects claimed that they were making or avoiding particular allocations for the sake of their psychological well-being. One subject twice insisted that he was "desperately wanting to be fair" (14:1-2), his desperation suggesting that he would be distressed were he to make unfair allocations. On the other hand another subject, a member of Group W, reported that he did "definitely prefer giving more points to the members...of Group W as opposed to Group X" (10:10), although he does go on to say that he is "not really bothered whether Group X gets more than Group W but I just think it adds...something interesting to the experiment...a bit of *spice*" (10:10-11). One subject, however, reported feeling "pretty, pretty horrible" about having displayed in-group favouritism in some of her allocations (21:18).

Six subjects reported over 8 occasions that their allocations were direct or indirect attempts toward personal economic gain. One gave one recipient the maximum possible number of individual points in case he was that recipient, even though he immediately reminded himself that subjects never made allocations to themselves (1:4). A second subject was concerned not to allocate too many points to others in case this had the consequence of depriving himself of possible money (5:17). A third used the fact that he was "skint" to justify a strategy of maximum joint profit (8:4,6+7). A fourth employed maximum joint profit in the hope that others would reciprocate (14:12) and a fifth employed maximum in-group profit for the same reason (21:18). A sixth subject allocated "average" points to others in the hope they would allocate more than average points to him (17:14). It should also be noted that several other subjects were concerned with

the personal economic consequences of their own and others' allocations, although they did not explicitly claim that such concerns motivated or explained their own allocation behaviour (1:3, 2:3, 12:8?, 13:5, 17:3, 19:15, 22:18).

Structural restriction was then the most common of the operative factors considered in Section B of the content analysis, with subjects being unable to follow favoured strategies being the most frequently voiced complaint. Largely in response to this, reported alternation was also relatively common. In-group identity was relatively frequent too, especially during intergroup allocations. Relatively few allocations were explained in terms of subjects attempting to make money for themselves or by referring to subjects' personalities. Others' behaviour was rarely considered by subjects and no subject ever wondered what the overall between-group outcome of everyone's allocations would be. Very few allocations were explained by subjects in terms of them conforming to norms or experimenter expectations, and few subjects attempted to enter implicit mutually advantageous reciprocal relationships with other subjects.

Content analysis of subjects' overall strategies

Qualitative summaries of each subject's overall account of their behaviour (see Appendix 5) were each coded into a single category, using the same coding frame as used in the "allocation intentions" section of the analysis reported above. Individuals' overall behaviour codes are shown in the table below alongside each subject's individual contribution to each pull score.

Sub. No.	FAV on MJP	MD on MIP+ MJP	FAV on F	MJP on FAV	MIP+ MJP on MD	F on FAV	Overall Behaviour Code
1	0	0	4	-12	12	- 2	9
2	0	0	0	12	12	12	1/5/6
3	- 5	- 1	2	3	9	2	9
4	12	5	0	0	7	12	1/4
5	- 4	5	8	- 4	3	4	1/3
6	- 5	- 8	- 6	5	4	4	9
7	0	6	0	0	6	12	5/6/9
8	- 6	- 6	0	6	6	12	5/6/9
9	-10	- 1	- 3	2	11	- 3	9
10	- 7	- 5	12	5	- 7	0	9
11	- 6	0	0	6	12	12	5/6
12	0	0	0	0	6	8	6
13	6	- 6	- 6	- 6	6	6	9
14	- 9	0	0	3	12	4	1/5/6
15	4	- 6	1	8	6	- 1	9
16	0	0	3	0	0	3	6/9
17	1	- 1	- 6	1	- 1	4	1/6
18	4	- 1	5	2	7	3	9
19	0	6	- 5	12	6	7	5/6/9
20	- 3	6	- 3	- 9	2	5	9
21	12	0	12	0	12	0	3/9
22	2	4	2	10	- 8	-10	9
Mean	-0.636	-0.136	0.909	2.000	5.591	4.273	
SD	5.892	4.223	5.089	6.133	5.738	5.742	

Overall the pull scores of MIP+MJP on MD and F on FAV are significantly different from zero ($t_{21} = 4.57$, two-tailed $p < 0.001$ & $t_{21} = 3.49$, two-tailed $p < 0.005$, respectively).

Key:- MIP = Maximum In-group Profit; MD = Maximum Difference in favour of the in-group = MIP + MD; FAV = In-group FAVouritism; MJP = Maximum Joint Profit; F = Fairness. For overall behaviour codes 1 = Personal profit; 2 = Personal status; 3 = In-group profit; 4 = In-group status; 5 = Joint profit; 6 = Fairness; 7 = Intergroup accentuation; 8 = Prototypicality; 9 = Random, arbitrary or none; 10 = Other intention; 11 = Absent reason. Other overall behaviour codes are combinations of these (see Appendix 4 for details).

Table 2.5: Subjects' overall strategies and individual pull-scores

Ten subjects (1, 3, 6, 9, 10, 13, 15, 18, 20 and 22) employed a random or arbitrary strategy overall, or employed no overall strategy at all. Although many of these subjects made large individual contributions to one or more pull scores these tended to cancel each other out. That is, no pull score was significant at the 5% level for the ten subjects considered as a sub-sample.

Five more subjects (7, 8, 16, 19, 21) did employ overall strategies but appeared to do so arbitrarily, merely to be able to apply a decision rule to their allocations (e.g. 7:8+16, 8:4, 16:6, 19:16, 21:7, see also 9:2+18). These subjects' individual contributions to pull scores were more or less in line with their instrumentally adopted strategies.

A further five subjects (2, 4, 5, 14, 17) similarly employed overall strategies instrumentally, this time seemingly motivated by economic self-interest (e.g. 2:3+14, 4:5+13, 5:15+17, 14:12, 17:3+14). These subjects' individual contributions to overall pull scores were less obviously in line with their instrumental strategies. Considered as a sub-sample these five subjects made small contributions to most pull scores, but made a mean contribution of 6.6 (SD = 5.683) to the pull of MIP+MJP on MD and a mean contribution of 7.2 (SD = 4.382) to the pull of F on FAV.

One subject (12) adopted an overall fairness strategy and another (11) adopted a combined overall strategy of maximum joint profit plus fairness. These subjects gave no indication that these strategies were being followed instrumentally and their individual contributions to pull scores were consistent with the strategies they were pursuing.

The most popular overall strategies for subjects in the present study then were random, arbitrary or explicitly absent strategies. Even when "rational" strategies were adopted these were usually pursued instrumentally, either to provide

some sort of decision-making guide when making allocations or to try and indirectly obtain personal profit, the second most popular overall strategy. On the whole overall behaviour codes and individual contributions to pull scores were consistent.

DISCUSSION

Non-discrimination

The most striking result from the present study is the almost complete lack of in-group favouritism as evidenced by the pull scores, indirect in-group favouritism scores, subjects' declared allocation intentions, and subjects' overall behaviour codes. Thus the present study does not support the "sufficiency condition" of social identity theory which it was set up to examine, i.e. that mere (recognized, multigroup) social categorization is sufficient to engender intergroup discrimination.

One explanation for this might be that the thinkaloud procedure brought to consciousness thoughts not usually attended to by minimal group subjects and that this resulted in a different pattern of behaviour from that which would usually follow social categorization per se. An obvious example of such thought processes would be those connected with evaluation apprehension (although subjects showed little sign of caring what the experimenter thought of them, even those who did indulge in in-group favouritism, and evaluation apprehension would surely have lead to strategies other than arbitrariness and/or economic self-interest). This is not the first study in which social categorization has failed to result in significant mean in-group favouritism, however, and in most minimal group paradigm experiments a significant minority of subjects do not employ such in-group favouritism

anyway. There seems little need to conclude, then, that subjects in the present study attended to thoughts other than those usually attended to by minimal group subjects.

A second possibility is that social categorization did lead to in-group favouritism but such discrimination was not detected by the dependent measures used. Necessarily there is no evidence for such a claim, but it is in any case somewhat unlikely bearing in mind the fact that subjects were audio-taped from before they made their first allocation until after they had made their last - and were given explicit instructions to verbalize any and all thoughts they had throughout.

A third possibility, consistent with social identity theory, is not so easily dismissed. This is that social categorization did not lead to mean in-group favouritism in the present study because the thinkaloud procedure altered the usual salience of the between-group situation (i.e. the social categorization). There are two possible ways of making such a claim. The first is simply that the thinkaloud attenuated the salience of subjects' in-group membership and of the between-group social categorization. This may be argued to have occurred because concentrating on their own thought processes and speaking them aloud (to the experimenter, as it were) made subjects' personal identities more salient than is usual in the minimal group paradigm, and because personal and social identities may be functionally mutually antagonistic (Turner, 1987: 49), this increase in personal identity entailed a decrease in the salience of subject's social identities and of the between-group social categorization more generally. Thus, the subjects in this study engaged in less in-group favouritism than is usual in the minimal group paradigm because their identities and their self-esteem needs were not as bound up in their minimal group memberships or in between-group social comparison outcomes as they would have been had the thinkaloud not been used.

The second way in which the thinkaloud procedure may have altered the salience of in-group membership and social categorization concerns the possibility that in "speaking to" the experimenter, subjects *re-categorized* the situation from a between- (minimal) groups one to a situation in which the subjects compared themselves (individually or as "the subjects") with the experimenter. To the extent that subjects did this, it would again be expected that discrimination by subjects between the minimal groups would be less than in "standard" minimal group experiments, or perhaps even non-existent.

These two explanations of the present study's results are not mutually exclusive. Subjects may have had their personal identities made more salient and/or may have recategorized the situation, either as an interpersonal one between themselves and the experimenter, or as an individual-group one between the experimenter and a group comprising of all of the subjects together. Nor do either of these possibilities exclude the further possibility that for other subjects the thinkaloud did not affect the salience of their minimal group memberships or of the between-group situation. Thus, it could be argued that the absence of mean sample-level in-group favouritism was not surprising: different identities and different social situations were salient for different subjects and therefore each subject tried to meet their identity and self-esteem needs in different ways, hence the variability of intergroup behaviour obtained.

There is in fact some evidence supporting such an interpretation. Almost half of the subjects (subjects 3, 8, 9, 14, 15, 16, 17, 19 and 20) made little or no reference in their decision making to the fact that there were two groups. Other subjects showed awareness of the between-group social categorization but made no reference to the fact that they were members of one group and not of the other (e.g. 1:4, 6:10, 7:7, 12:2, 18:10). At least two more subjects had to

remind themselves of their group membership (e.g. 13:5, 22:14). And still more subjects demonstrated confusion as to which group they were members of (e.g. 1:4, 5:4+11). Finally, the subjects who did individually employ in-group favouritism (4, 5 and 21) all also indicated identification with their in-groups. Thus, it seems that there is considerable evidence supporting the claim, consistent with social identity theory's "sufficiency condition", that where in-groups and the between-group situation were salient (as indicated by identification with in-groups), in-group favouritism was indeed the result, but where such salience was absent, so too was in-group favouritism.

There are however several obstacles to accepting such an interpretation. Least telling, perhaps, is that the interpretation requires explaining the heterogeneity of behaviour obtained in the present study via individual differences in the subjects' reactions to the situation experienced by all of them (i.e. social categorization plus thinkaloud). But relying on individual difference explanations is exactly what social identity theory tries to get away from (Tajfel 1978a: 27-28), and allowing too heavy a reliance on them gives the theory an air of unfalsifiability (Abrams, 1992: 62-63).

Further, there is very little evidence from their protocols (see Appendix 3) that any subject categorized their situation as an individual-group one between the experimenter and themselves (i.e. the subjects, individually or collectively). Even if subjects did so categorize, it is far from clear what sort of behaviour this might promote: how might subjects achieve positive distinctiveness from an experimenter? Maximum joint profit (and therefore maximum "loss" by the experimenter) is perhaps the most likely candidate, but as has already been mentioned in the results section, on only 8% of all allocations did subjects report using such a strategy.

Most importantly, great efforts were made to ensure that subjects *would* be aware of the (between-groups) social categorization and of their social identities as minimal in-group members,¹² and at least some subjects (who did not engage in in-group favouritism) were indeed quite clear about both the social categorization and their place within it (e.g. 2:2, 4:2, 10:9, 11:5, 21:3). To the extent that this strongly suggests that subjects' minimal group social identities were salient (even if other identities were also salient, which is nearly always the case, both inside and outside the minimal group paradigm -see Tajfel, 1978a: 38-45), then social identity theory suggests that the subjects would be motivated to engage in in-group favouritism in order to ensure that these salient aspects of their identities *would be* positively evaluated (or at least would not make *negative* contributions to their self-esteem, which would be likely to occur if the subjects did not discriminate but out-group members did).¹³

Finally, although it is true that the three subjects (4, 5 and 21) who employed significant amounts of in-group favouritism also indicated identification *with* the group (as well as identification *as* in-group members), an equal number of subjects (2, 10 and 22) expressed such identification with the

¹² The front covers of subjects' response booklets contained only information about the particular subject's code number and group membership; subjects were required to read aloud a sentence which repeated information about their code number and group membership; subjects were required to read out the code number and group membership of each recipient before every allocation; and after every allocation subjects had to complete a sentence saying how many points they had allocated to each recipient - identified by their code number and group membership.

¹³ Remembering that when minimal group memberships are salient in the minimal group paradigm in-group favouritism is the only route to positive in-group distinctiveness, in-group evaluation, social identity, and self-esteem (see Chapter 1).

group but did not engage in in-group preference.¹⁴ Thus, it again appears not to be the case that mere (recognized, multigroup) social categorization is sufficient to promote in-group favouritism, even when accompanied by identification with one's in-group.

How the various theories fare in the present study

Social identity theory suggests that mere social categorization results in intergroup discrimination because of each subject's need to perceive their minimal in-group as superior to the out-group in order to obtain positive evaluations of the in-group and of their social identity derived from it (i.e. a positive social identity as an in-group member). The first problem with this account is clear and has been extensively considered above: mere (recognized, multigroup) social categorization did not result in in-group favouritism (either at the mean sample level or reliably at the individual level).

The second problem (also briefly touched on above) is that social categorization (accompanied by identification as an in-group member) did not prove sufficient to engender social identification with in-groups either. Most socially categorized subjects did not report or indicate any identification with their minimal in-group membership. Only six subjects (e.g. 2:13, 4:3, 5:4, 10:9, 21:6 and 22:18) indicated identifying with the in-group at some point during their allocations, and five of them were at some stage confused about, abandoned, or even actively refused to make such identifications (e.g. 2:2, 5:11, 10:12, 21:11, 22:14).

¹⁴ Also, two of the subjects who each identified with their in-groups and engaged in in-group preference claimed to be following the strategies they did primarily as an indirect means of obtaining personal profit. This is clearly inconsistent with social identity theory.

Of these six subjects who expressed identification with their in-groups, three (subjects 4, 5 and 21) accounted for 16 out of the total of 22 expressions of such identity. These three were the only subjects who received overall behaviour ratings which included in-group favouritism as a component and together they obtained high mean intergroup discrimination pull scores (FAV on MJP = 6.67 (9.24), MD on MIP+MJP = 3.33 (2.89), FAV on F = 6.67 (6.11), MJP on FAV = -1.33 (2.31), MIP+MJP on MD = 7.33 (4.51), F on FAV = 5.33 (6.11)).¹⁵ This suggests that, consistent with social identity theory, identifying with one's in-group may be necessary (but not sufficient) to engender intergroup discrimination (e.g. 4:11+16, 5:4, 21:4, see also 10:9, 22:10), even if social categorization is clearly not in itself sufficient to engender either identification with in-groups or in-group favouritism.¹⁶

Social identity theory considers social identity as an in-group member to be a cause of intergroup discrimination because when people identify themselves as in-group members their self-esteem becomes dependent upon how that in-group is evaluated. Such people therefore discriminate in order to achieve or protect positive in-group distinctiveness so that the in-group can be positively evaluated and they can enjoy positive self-esteem. In this study, however, no subject wondered what the overall between-group outcome of everybody's allocations would be, not even the subjects who did engage in in-group favouritism. Additionally, no subject expressed satisfaction at engaging in discrimination, but one who did

¹⁵ For the six subjects who expressed any identification with minimal in-groups the mean pull scores were as follows:- FAV on MJP = 2.5 (7.994), MD on MIP+MJP = 1.5 (3.937), FAV on F = 5.67 (5.715), MJP on FAV = 3.83 (6.274), MIP+MJP on MD = 3.17 (8.93), F on FAV = 3.00 (8.367).

¹⁶ The claim can only be that identification with in-groups may be necessary to promote in-group favouritism as there is a possibility that the direction of causality is the other way around, i.e. that in-group favouritism results in identification with in-groups.

discriminate said that she felt "pretty, pretty horrible about doing that" (21:18) which is in direct opposition to motivational predictions from social identity theory (see Abrams & Hogg, 1988).¹⁷

Finally, all three of the subjects who discriminated received "overall behaviour" codings which suggested that their discrimination was instrumental not toward positive in-group distinctiveness, social identity and self-esteem, but was rather instrumental either toward economic self-interest (subjects 4 and 5) or so that the subject concerned had some sort of decision-rule to guide their allocations (subject 21).

Overall then social identity theory receives little support from the present study. Social categorization proved sufficient neither for identification with in-groups nor for intergroup discrimination. Further, although social identity may have been necessary for in-group favouritism, it was not sufficient, and the discrimination employed did not seem to either be motivated by, or necessarily result in, positive self-esteem. Additionally, no subjects considered or wondered whether their behaviour (discriminatory or otherwise) might contribute to positive or negative in-group distinctiveness.

Generic norm theory claims that subjects in the minimal group paradigm are trying to balance competing demands of conforming to a generic norm of in-group preference and conforming to another generic norm of fairness. On only four occasions during the present study were norms evoked by subjects as explanations of their allocation behaviours. Two could conceivably be considered instances of a generic norm of groupness (4:3, 21:18), but the other two stem from particular (non-minimal) groups subjects belonged to: law students (4:1)

¹⁷ One expressed pleasure at having at being able to use in-group favouritism as a decision-making guide, but she made it clear that positive in-group distinctiveness was not her goal (10:10).

and students in general (14:12). The vast majority of subjects made no reference to norms as explanatory entities and none ever claimed to be trying to balance the competing demands of a norm of fairness and another norm demanding intergroup discrimination.

It is of course possible that generic norms were operative without subjects being aware of that fact. The real problem for generic norm theory is the variety of strategies followed, both across subjects and by individual subjects during their own series of allocations. The theory would need to incorporate many more norms than simply fairness and discrimination, and it would then need to specify why different subjects conformed to different norms, and also why the same subjects conformed to different norms at different stages during their allocations. It is doubtful that all or any such norms could legitimately be called generic. The present study therefore counts rather more against than in favour of generic norm theory.

Equity theory suggests that intergroup discrimination occurs in the minimal group paradigm because subjects are trying to counter anticipated in-group favouritism on the part of out-group members in order to achieve an equitable outcome overall. Subjects in the present study rarely consider others' actions, however, and even when they do they do not show any signs of expecting out-group members to display in-group favouritism (absolutely or relative to other in-group members), still less of letting such expectations dictate their own behaviour. It is more common for subjects to hope or believe that other subjects are acting in the same way as themselves (4:5+13, 14:1+12, 17:14, 21:18). Also, although most subjects showed a concern for equality or parity (i.e. "fairness") at some point, only two showed signs of being motivated by a concern for equity (12:1+9, 21:11), one of whom certainly did not let such concerns dictate her behaviour (21:12+14). Third, as already noted, subjects showed no

interest in the overall between-group outcome of everybody's allocations. Equity theory therefore also receives little support from the present study.

The behavioural interaction model claims that minimal group subjects pursue economic self-interests via instrumental cooperation with in-group members and instrumental competition against out-group members because they perceive more positive interdependence with the former than with the latter as a result of social categorization making subjects feel more common fate with in-group than with out-group members.

The motivational basis for the model is supported by the present study in so far as many subjects are fairly clearly motivated by economic self-interest in their allocations and/or consider the economic consequences to themselves of their own and others' allocations (e.g. 1:3-4, 4:13, 5:15+17, 8:4, 13:5, 14:12, 17:3+14, 19:15, 21:2, 22:18), with at least one subject (temporarily) feeling that without the opportunity of allocating to himself there is no reason to prefer any allocation strategy to any other (2:3).

A variety of perceived interdependencies were evident in the present study and although these were not all predicted by the behavioural interaction model, the model can provide a reasonable post-hoc explanation of both these interdependencies and of the various strategies employed by the subjects.

It has already been noted that the imposed social categorization in the present study did not have uniform effects. Some subjects perceived more positive interdependence with in-group than with out-group members, as is usual in the minimal group paradigm. These subjects would be expected to perceive more common fate and identify more with in-group than with out-group members and they would also be expected to pursue instrumental cooperation with the former

and instrumental competition with the latter (e.g. 4:3,11,13,15+16 and 21:4,6,7,10,15,17+18).

If for some reason a subject thought that there was only a finite amount of money available to be shared between in-group and out-group, however, and similarly only a finite amount of money to be shared between in-group members, such a subject would be expected to perceive negative interdependence between both the in-group and the out-group and between himself and other in-group members. S/he would then be expected to perceive more common fate, and identify more, with in-group members than with out-group members, but to employ both instrumental in-group favouritism and instrumental *intragroup* discrimination in favour of him or herself (e.g. 5:4,15,17+18).

If on the other hand the emphasis on the randomness of the social categorization caused subjects to perceive more or less equal positive interdependence with all other subjects, these subjects would be expected to perceive common fate, and to identify with, in-group and out-group members alike, and therefore to employ instrumental cooperative strategies with everyone, regardless of group membership (e.g. 7:7, 11:8, 14:all, 17:14).

Finally, if the randomness of the social categorization instead caused subjects to perceive no interdependence relationship with any other subjects, these subjects would be expected to perceive no common fate and identify with no-one, and they would consequently be expected to employ random and/or arbitrary allocation strategies (e.g. 1:4, 3:16, 8:1+2, 9:2, 13:2, 20:1, 22:1,2).

While this explanation of the results of the present study is perhaps superior to any which could be gleaned from the other theories considered here, difficulties remain for the behavioural interaction model. First, the explanation

•

offered is post-hoc and suffers from not being able to account for the fact that the social categorization resulted in differing perceptions of interdependence in the first place. Second, as has already been noted, few subjects considered the likely actions of others and fewer considered the likelihood of others engaging in mutually beneficial reciprocity strategies when deciding how to act themselves, both of which would be expected if subjects were truly concerned with exploiting perceived interdependencies. Third, and perhaps most damaging, the behavioural interaction model cannot explain why a substantial minority of subjects seemed to adopt strategies not as indirect methods of satisfying economic self-interest but rather as guides to decision-making. The behavioural interaction model offers only a post-hoc and incomplete account of the present study's findings, therefore, and the results of the present study offer rather mixed support for the model.

Intergroup accentuation theory clearly receives very little support from the present study. Two subjects each made one allocation based on accentuating or reflecting intergroup differences (6:10, 22:18), one of whom also made a single allocation based on accentuating or reflecting intragroup similarity (6:8). Both of these subjects, however, received random overall behaviour ratings, suggesting that each of the aforementioned allocations were simply one of a range of arbitrary strategies used by the subjects. None of the other 20 subjects, several of whom pursued consistent and meaningful strategies, mentioned intergroup accentuation.

Self-categorization theory received absolutely no support from the present study. No subjects considered an in-group prototype when deciding how to make decisions. Nor did they seem concerned with trying to act as much like in-group members and/or as little like out-group members as possible.

Several subjects considered the possibility that they were "meant" to display in-group favouritism in their allocation decisions. Three subjects wondered if discriminating in favour of the in-group was "part of the experiment" (e.g. 5:7, 11:13, 22:10). Two more felt that they "ought" to display in-group favouritism (2:2, 21:4), while one simply asked "how am I meant to do this?" (13:6). Such expressions would provide fairly strong evidence of an experimenter expectation explanation of minimal group discrimination were it not for two things. First, subjects rarely followed their speculations of how they "ought" to act with consistent in-group favouritism in order to please the experimenter (e.g. 2:2, 5:15, 11:16, 13:7, 21:11, 22:14). Second, much of the possible support for an experimenter expectation/demand characteristics explanation is equally consistent with other theoretical positions. A subject might claim that she "should" behave in a certain way, for example, in order to (i) achieve positive in-group distinctiveness, (ii) conform with a generic norm, (iii) achieve an equitable out-come, (iv) reciprocate with similar acting others, (v) accentuate intergroup differences, or (vi) act prototypically, as well as to (vii) conform to demand characteristics. As absolutely no subjects explained their behaviour by saying that the experiment or experimenter "forced" them to act in that way, the present study offers no unambiguous support for an explanation of minimal group behaviour in terms of demand characteristics.

To recap, the results of this study offer little support for generic norm theory, equity theory, intergroup accentuation theory, self-categorization theory or a demand characteristics explanation of minimal group behaviour. Social identity theory suffers badly from the fact that social categorization failed to result in widespread in-group favouritism or even identification with in-groups. It also suffers from the lack of evidence that those subjects who discriminated did so in order to obtain positive in-group distinctiveness, social

identity and self-esteem. Indeed, one discriminating subject's self-esteem seemed to diminish as a result of her actions. The behavioural interaction model perhaps receives the most support of the theories considered. Many subjects showed an interest in furthering economic self-interest and many subjects' strategies were consistent with the particular pattern of interdependence they can be expected to have perceived. The model cannot explain why social categorization resulted in several different perceptions of interdependence, however, and cannot explain why some subjects seemed more concerned with finding and using a decision-making rule than they were with furthering their economic self-interest. Overall, then, all of the theories considered suffered from the sheer variety of behaviours displayed in the present study and the number of differing explanations offered by subjects to account for their actions.

Limitations of the Tajfel matrices

There was little evidence in the present study for the claim that the Tajfel matrices do not represent subjects' behaviours. Individual contributions to pull scores may be misleading when considered in isolation, particularly those of subjects who followed random overall strategies, but such anomalies tend to cancel each other out when mean pull scores are calculated. Moreover, the individual contributions to pull scores of those subjects who did follow "meaningful" overall strategies were remarkably consistent with those overall strategies.

Nor did the present study provide any evidence that the Tajfel matrices lack content validity in not permitting the allocation of negative gains (i.e. losses, penalties, etc.). No subjects complained about such a restriction, even when engaging in in-group favouritism.

A lack in content validity was suggested, however, by some of the restrictions that subjects felt were imposed upon them by the structure of the matrices. For example, several subjects were distressed that maximum joint profit and fairness were negatively interdependent along half of matrix type 1. These subjects wished to pursue a policy of maximum joint profit, but were unwilling to do so when it entailed considerable unfairness - both in that one recipient would receive much less than the other recipient and that the same recipient would receive less points than they would have were the point of maximum fairness selected. This was why pursuing maximum joint profit was more acceptable on matrix type 2 than it was on matrix type 1: in the former case maximum joint profit meant that the recipient who received the least points still received more than they would have had the point of maximum fairness been chosen. This suggests that the pull score of MJP on FAV may be artificially depressed because of a pull of "F" (i.e. joint profit without unacceptable individual loss) on "MJP" (i.e. maximum joint profit combined with unacceptable individual loss). Similarly, the pull of FAV on F may be artificially elevated by a pull of "F" on "MJP". At the very least the Tajfel matrices employed in this study lack content validity in that they do not permit easy pursuit or measurement of a strategy attempting to maximize both joint profit and fairness.¹⁸

Another difficulty subjects had with the Tajfel matrices occurred when the subjects had decided on particular strategies to follow, but were then confronted with a matrix which either did not permit pursuit of that strategy and/or meant that all possible options satisfied the strategy equally. Matrix type 3 was the main problem here. Several

¹⁸ Turner (1983a: 352-353) claims that confounds can be disentangled by comparisons across matrix types. Even if true it is rarely if ever done and indeed would be a complicated and difficult affair using real as opposed to "ideal" scores. Anyway, it probably is not true, as I will argue later in the thesis.

subjects adopted a strategy of maximizing joint profit and then found that their preferred strategy offered no guidance as to how to make allocations on matrix type 3. Some subjects continued with their overall strategy and simply made arbitrary decisions on this particular matrix. Other subjects, however, adopted a secondary strategy of fairness-on-this-type-of-matrix and/or modified their simple MJP strategy to one of MJP+F. This is, of course, an example of the design of the Tajfel matrices acting as a confounding variable in the study of the effects of social categorization on allocation behaviour.

A similar problem occurred occasionally with changing matrix types. Subject 2, for example, effectively follows an overall strategy of maximizing fairness for his first three allocations (although he claims to be allocating randomly), all of which happened to be type 3 matrices. A new matrix type was presented for the fourth allocation and the subject suddenly adopted an overall strategy of maximizing joint profit (with a secondary strategy of fairness on matrix type 3). Again allocation behaviour is being affected by measurement of the dependent variable, rather than by the independent variable of social categorization.

Pull scores derived from Tajfel matrices do therefore seem to be valid representations of certain aspects of minimal group subjects' behaviour, but they also seem both to prevent expression of certain behaviours and to encourage other behaviours which do not strictly result from social categorization per se.

CONCLUSIONS

The main conclusion of the present study has to be that social categorization is sufficient for neither identification with in-groups nor intergroup discrimination. The former consideration means that category membership (i.e. identification as a member of a social category) is not, as is sometimes suggested within the social identity theory literature, the same thing as social identification with a social category or group (nor even necessarily promotes it). The latter consideration means that social identity theory's "sufficiency condition" is not supported, thus potentially undermining the claim that the theory provides a necessary modification of Sherif's theory of intergroup conflict (as the minimal group paradigm does not provide proof that negative interdependence of group goals is unnecessary as a precondition for intergroup conflict).

Further, if identification as an in-group member is not the same thing as identification with an in-group, and the former does not necessarily promote the latter, the question has to be asked: when do people identify with the social categories and groups of which they perceive themselves as members? Also, if (accepted, multigroup) social categorization does not necessarily lead to intergroup discrimination, what does? The behavioural interaction model claims that identification with in-groups occurs when members perceive themselves to be positively interdependent with each other, and accepts Sherif's contention that intergroup discrimination results from perceived negative interdependence of group goals. Each of these claims will be examined and rejected later in this thesis. Nevertheless, the main conclusion of this chapter stands: (recognized, multigroup) social categorization (including self-categorization as a member of one group and not of others) is a sufficient cause of neither identification with in-groups nor intergroup discrimination.

Another main conclusion is that there is far more going on in the minimal group paradigm than straightforward in-group favouritism. A whole range of strategies are pursued, sometimes by single subjects. None of the theories considered so far are currently sufficiently broad or flexible to account for all of the behaviour displayed in the present study. The behavioural interaction model perhaps came closest, but only with a post-hoc explanation which could not account either for the various perceived interdependencies stemming from social categorization or for the strategies of subjects who seemed more concerned with being able to make consistent, coherent or meaningful allocation decisions than they were with economic self-interest. Clearly any of the theories considered so far which wishes to explain minimal group behaviour will need considerable modification; particularly perhaps social identity theory, which is widely regarded as relying heavily on the results of that paradigm.

The final conclusion reached here is that although pull scores derived from allocations made on Tajfel matrices are reasonably consistent with subjects' stated intentions, the matrices themselves are: (i) somewhat restrictive in structurally inhibiting some strategies and structurally confounding others; and, (ii) potentially corrupting in the minimal group paradigm in as much as their use seems to promote behaviour which might not manifest itself as a result of social categorization alone. Characteristics of pull scores derived from allocations on Tajfel matrices - and the nature of the behaviour they measure - is the topic of the next chapter.

CHAPTER 3: STUDY 2 - THE NATURE OF INTERGROUP DISCRIMINATION

CHAPTER OVERVIEW

This chapter reports a minimal group study (cf. Tajfel et al., 1971) which employed two different measures of social behaviour. These were pull-scores, derived from "Tajfel-matrices", and scores derived from "allocation grids", inspired by McClintock's (1988) social value vectors.

It is argued that allocation grid scores have a number of advantages over pull-scores derived from Tajfel-matrices. These include: clearly being of interval scale with equal origin and unit; superior clarity, transparency, and reliability; less susceptibility to structural confounds; and, because "bads" as well as or instead of "goods" can be allocated by subjects using allocation grids, superior construct validity, and a superior ability to distinguish between different forms of discrimination across groups.

Mean sample-level in-group favouritism was employed by subjects in this study. At an individual level, however, a large proportion of subjects employed (a variety of) dominant strategies other than in-group favouritism. Contrary to social identity theory's "sufficiency condition" (see Chapter 1), this suggests that mere social categorization is insufficient to promote individuals to engage in discrimination in favour of the in-group.

Where subjects did employ dominant strategies of in-group favouritism, two distinct forms could be identified. Both involved in-group preference, but one also involved out-group derogation, whereas the other did not. This suggests that where social categorization is followed by in-group favouritism, something other than social categorization determines the form that such discrimination takes.

INTRODUCTION

In the previous chapter it was noted that several methodological criticisms have been levelled at the Tajfel matrices and the behavioural scores derived from them. It has been suggested that "pull scores" derived from the matrices may not "provide a representative description of the subjects' behaviour" (Aschenbrenner & Schaefer, 1980: 396); that the matrices may lack content validity in usually excluding the possibility of allocating negative gains to recipients (Mummendey et al., 1992); that matrices are structurally mutually inhibitory of some strategies whilst structurally mutually entailing others (Bornstein et al., 1983a: 342); and that subjects may be forced into 'second-choice' strategies by the constraints imposed by the matrix design (Bornstein et al., 1980: 342). These criticisms contribute to a long running debate about the suitability of the Tajfel matrices to measure social behaviour (see Aschenbrenner & Schaefer, 1980; Bornstein et al. 1983a,b; Branthwaite et al., 1979; Brewer, 1979; Brewer & Silver, 1978; Brown et al., 1980; Hyland, 1979; Locksley et al., 1980; Mummendey & Schreiber, 1983; Mummendey et al., 1992; Ng, 1981; Platow et al., 1990; Turner, 1980, 1983a,b). Other aspects of that debate include arguments about the level of measurement, external validity, clarity and the reliability of pull scores.

Aschenbrenner & Schaefer (1980: 395) argue that pull scores cannot be assumed to be of "interval scale with equal origin and unit". Such characteristics are imperative if pull scores obtained on different matrices are to be meaningfully compared with each other. Brown et al. (1980: 405-406) reply that there are "two possible views on this" depending on whether pull scores are considered "spatially" (i.e. as some kind of rating scale with two opposed poles) or "numerically" (i.e. as reflecting the value of the points distributed). If they are considered numerically then Brown et al. (1980: 407) admit

that pull scores of FAV on MJP and MIP+MJP on MD are not even ordinal in parts, "since not only does a given 'pull' not uniquely describe a reward outcome, but it is also possible that a 'pull' of $x + 1$ could represent less discrimination than a 'pull' of x ". If they are considered spatially, however, all of the pull scores are interval level data and pull scores derived from different matrices can be meaningfully compared.

A spatial consideration of pull scores requires an assumption that the difference between the numerical values of adjacent points on a matrix does not affect the level of measurement of pull scores derived from that matrix as long as the "progression" from one extreme of the scale to the other is observed. Thus each of the following adjacent points on a type 1 matrix are equivalent:¹ (i) 14/11, 13/13; (ii) 14000/0, 13/13; (iii) -11/-14, 13/13; (iv) 14/-11, 13/13; and (v) 14000/-14000, 13/13. Brown et al. (1980: 408) say that neither evidence nor argument has appeared to make this assumption psychologically untenable. This is an odd claim to make as Tajfel himself (one of the co-authors of the Brown et al. paper) said that "taking away from the outgroup...may well have a very different [psychological] significance from simply giving less to the outgroup than the ingroup" (Tajfel et al., 1971: 174, original emphasis), a claim now empirically supported by Mummendey et al. (1992).

Pull scores lack external validity in the sense that they are essentially relative entities - they reflect the "pull" of one distributive strategy upon another. Obtaining a high value for a pull of FAV on F in the minimal group paradigm, for example, does not in itself legitimize the standard claim that mere social categorization is sufficient for intergroup discrimination. If anything, it entails only that social categorization is sufficient for in-group favouritism when the

¹ See Appendix 1 for the matrices and derivation of pull-scores from them.

choice is between in-group favouritism and fairness. It tells us little or nothing about what would happen if the same sample were offered the choice between ingroup favouritism and maximum joint profit following mere social categorization. Pull scores are not indicative of "absolute" or "unfettered" tendencies: they are essentially relative. All that can be claimed from minimal group experiments employing the Tajfel matrices is that social categorization is sufficient for in-group favouritism rather than particular other distributive strategies which such discrimination has been measured in opposition to. Justification of the standard claim that social categorization is sufficient for intergroup discrimination (as indicated by FAV and/or its constituent parts, MIP and MD) would require that intergroup discrimination had exerted a significant pull on every other possible distribution strategy. This would require use of considerably more matrix types than shown in Appendix 1. Theoretically, a separate matrix would be needed to measure the pull of every possible distribution strategy on every other possible distribution strategy.

The reliability of pull scores is also open to debate. Each pull score is essentially a "single-shot" measure, comprised of the outcome of mathematic combination of scores on two intergroup matrices, each of which is only presented once. In the study reported in the previous chapter there was evidence that some subjects needed to make a few allocations before they worked out which strategy they were trying to pursue and how to pursue it, particularly in view of some of the difficulties imposed by the structure and sequence of the grids. Similarly, some subjects showed signs of getting bored and impatient towards the end of their allocation sequence because of the effort and attention involved in trying to maximize their favoured strategy, which sometimes resulted in them adopting a different and more simplistic strategy. The single-shot nature of pull scores means that if either of the matrices which contribute to a particular pull-score are

positioned early or late in the allocation sequence, the pull score in question may be quite seriously distorted.

Pull scores have also been charged with lacking clarity, as some particular values of pull scores do not uniquely describe reward outcomes. For example, an MJP on FAV pull score of 6 will result if a subject allocates 13/13 on an i/o presentation of matrix 1 and 7/25 on the o/i presentation of that matrix; or 7/25 on the i/o presentation and 13/13 on the o/i one; or any of five other allocation patterns. (This is the worst case scenario: the more extreme the pull score, the fewer the reward outcomes it can represent.) This is again because pull scores represent the strength of one distribution strategy *relative* to the strength of another rather than the "absolute" strengths of one or more strategies.²

In sum, the Tajfel matrices and the pulls derived from them have then been accused of not providing a representative measure of subjects' behaviour, lacking content validity, structurally inhibiting some strategies and structurally entailing others, forcing subjects to pursue second-choice strategies, providing an inadequate level of measurement, lacking external validity, lacking reliability and lacking clarity. Defenders of pull scores have strenuously rejected all of these criticisms but there seems to be a case nevertheless for trying to develop a measure of social behaviour (as indicated by distributions of goods) which is at least more transparently immune from criticisms such as these than are the Tajfel measures.

Before considering what form such a measure might take it is worth thinking about the nature of the social behaviour(s) the measure is to represent. The most important of these, with respect to social identity theory's "sufficiency condition"

² This is "absolute" in the sense of being relatively independent of the strength of other strategies, not in the sense of being constant across situations.

(see Chapter 1),³ is intergroup discrimination. The Tajfel matrices measure two forms of discrimination in favour of the in-group: maximum in-group profit and maximum difference in favour of the in-group, which together are termed in-group favouritism. The former of these involves providing in-group members with as many goods as possible regardless of the consequences for out-group members and the latter involves trying to ensure that in-group members receive on average more goods than out-group members. The first might be called an "absolute" or "independent" in-group bias in the sense that in-group benefits are pursued with little or no regard as to how the out-group fares, while the latter might be called "relative" or "interdependent" in-group bias in the sense that pursuing in-group benefits entails a restriction of out-group benefits.

Such distinctions can be illustrated with reference to Sherif's classic 1954 "bean collection" task conducted during one of his "summer camp studies" (Sherif & Sherif, 1967). Sherif had beans strewn all over a piece of land and then had members of two groups collect as many beans as possible during a limited period. The group members were then shown projections ostensibly of the number of beans each individual had collected. In fact the same amount of beans was projected for each individual, just in different spatial arrangements. Sherif found that "each group, on the average, overestimated the number of beans collected by fellow members, and made much lower estimates of the detested out-group's performance. The tendency to overestimate was much greater for the group that had won the tournament of games just concluded. The losers overestimated their own performance and underestimated that of their rivals on this task" (Sherif & Sherif, 1967: 83, emphasis added). In fact, (rounded up to the nearest bean),

³ i.e. "that the mere perception of belonging to two distinct groups - that is, social categorization per se - is sufficient to trigger intergroup discrimination favouring the in-group" (Tajfel & Turner, 1979: 38).

the winning group overestimated in-group members' performances by an average of 12 beans and out-group members' performances by an average of 5 beans, while the losing group overestimated in-group members' performances by an average of only 4 beans and were in fact accurate, on average, about the number of beans collected by out-group members (see Sherif & Sherif, 1967: 84: Figure 5.2).

It is possible, indeed almost certain, that intergroup processes were responsible for the biases shown in the bean collection task. It is nevertheless also possible to consider the biases themselves as having two aspects, absolute and relative. Winning group members showed an absolute bias in favour of their own group because they overestimated in-group members' performances by an average of 12 beans, and losing group members demonstrated a lesser absolute bias in favour of their in-group because they overestimated the performances of in-group members by "only" 4 beans. Winning group members also showed an absolute bias in favour of the out-group in that they overestimated the average performance of out-group members by 5 beans, whereas the losing group showed no absolute bias in favour of or against the out-group because they accurately reported mean out-group members' performance.⁴

The combined effect of these biases was that the winning group showed a relative bias in favour of the in-group in that they overestimated absolute in-group members' performances by an average of 7 beans more than they overestimated absolute out-group members' performances. The losing group also showed a relative bias in favour of the in-group, albeit a lesser one, in as much as they overestimated absolute in-group members'

⁴ "In favour of" and "against" refer to the truth/claim difference where in the former case the claim is "better" than the truth and in the latter case it is "worse". It does not refer to the intention behind the bias (e.g. trying to do good for the recipient) or to how the recipients of the bias feel about the claim that is made (e.g. disappointed).

performances by an average of 4 beans but they did not overestimate absolute out-group members' performances at all, on average. Thus the winning group showed more relative in-group bias than did the losing group.

The distinction between absolute and relative biases is important because although both may result from intergroup processes neither necessarily does so. With respect to absolute bias, it is quite possible for a group to over- or under-estimate its members relative to an "absolute" standard without considering out-group members at all, and/or to absolutely over- or under-estimate the performance of out-group members without reference to the in-group or its members. Consequently, while relative biases can and probably usually do result from intergroup processes, they can also occur "unintentionally" as a result of imbalanced absolute biases. If a group over-estimates an in-group performance (i.e. makes an absolute bias in favour of the in-group) but accurately estimates an out-group performance (i.e. makes no absolute bias either in favour of or against the out-group) there is a resultant relative bias without any comparative intergroup processes ever having necessarily taken place.

Returning to the distributive strategies examined in the minimal group paradigm, maximum in-group profit is perhaps best thought of as an absolute bias in favour of the in-group, while maximum difference in favour of the in-group is perhaps best thought of as a relative bias. Maximum in-group profit is an absolute bias because it is a concern with the in-group *regardless* of what happens to the out-group. Although this is intergroup in the sense that subjects cognitively and behaviourally differentiate between in-group and out-group, it is not intergroup to the extent that subjects are (theoretically) unconcerned about the overall relative bias in favour of or against the in-group. On matrix type 1, for example, maximum in-group profit will result in a relative bias in favour of the in-group, but on matrix type 2 the same

strategy will result in no relative bias. It is also structurally possible for maximum joint profit to result in a relative bias in favour of the out-group (as occurs at least temporarily when subjects pursue MIP on i/o presentations of matrix type 2).

Maximum difference in favour of the in-group is a relative bias because it is (theoretically) the difference between in-group and out-group which subjects are interested in: they are unconcerned about the extent of absolute bias in favour of or against the in-group (as long as long as the absolute bias in favour of the in-group is greater than the absolute bias in favour of the out-group). Maximum difference in favour of the in-group maximizes the absolute bias in favour of both the in-group and the out-group on matrix type 1, for example, while it minimizes absolute bias in favour of both the in-group and the out-group on matrix type 2. Various other patterns of absolute bias are compatible with a relative bias in favour of the in-group, as we shall see below.

Thus in-group favouritism as measured by the Tajfel matrices may be measuring two conceptually different types of in-group bias, comprising as it does of both maximum in-group profit and maximum difference in favour of the in-group. Both can be considered instances of intergroup discrimination in the sense that subjects have noticed a difference between the groups and are treating them differently, but whereas maximum difference in favour of the in-group is obviously a concern for the in-group *relative to the out-group*, maximum in-group profit is a concern for the in-group with *no concern for the out-group*. Tajfel matrices can be viewed as limited, therefore, in that these potentially different types of discrimination cannot easily be disentangled. In other words, the Tajfel matrices may be confounding different types of discrimination, which must be considered a serious defect in view of the importance the notion of intergroup discrimination has in both the

minimal group paradigm and within social identity theory itself.

The possibility of a measure of social behaviour which avoids many of the problems associated with the Tajfel matrices (and many of their suggested replacements) is suggested by McClintock's (1988) use of vectors to diagrammatically represent "social values" as shown below.

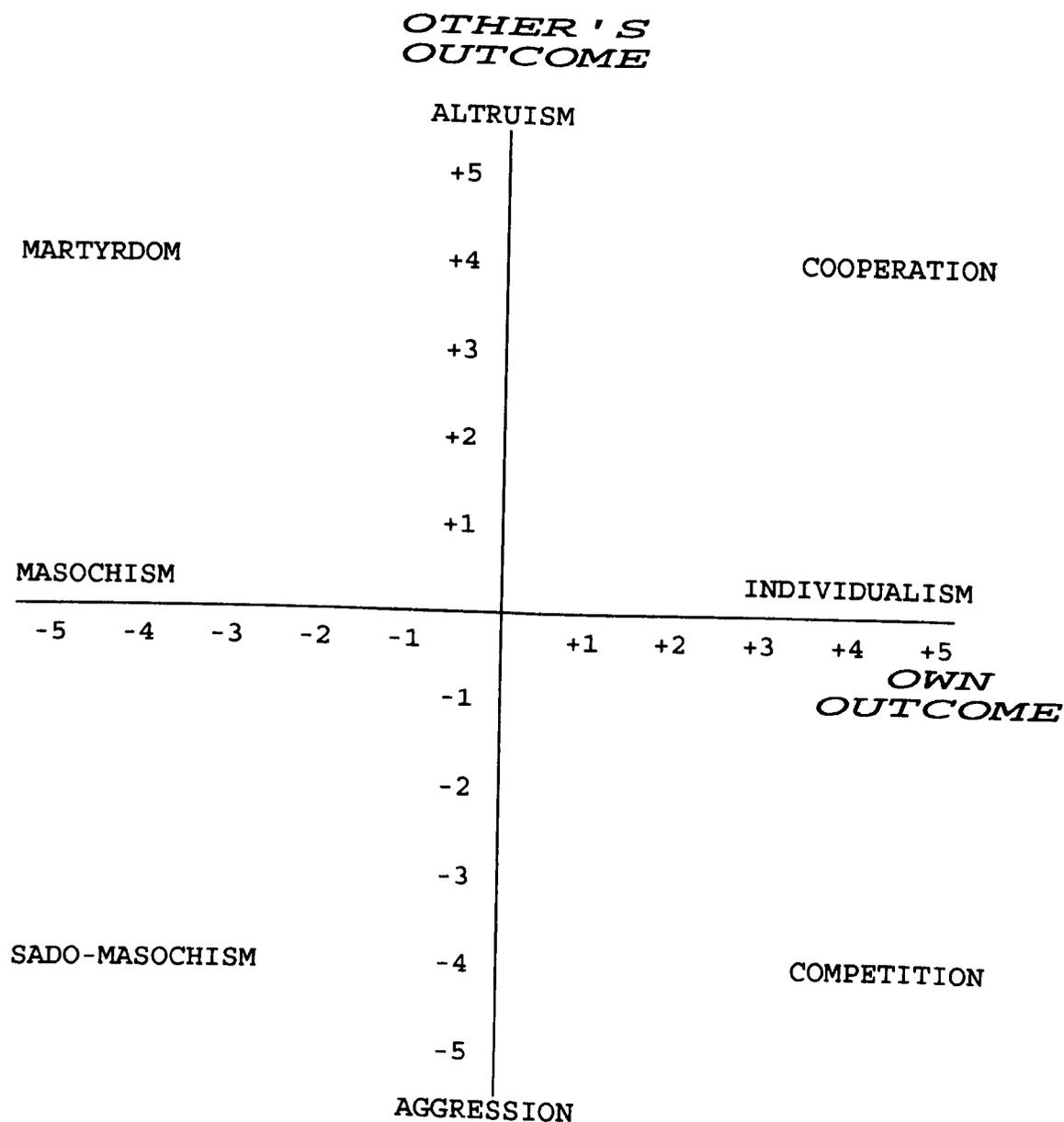


Diagram 3.1: McClintock's (1988) social value vectors

Horizontal movement to the right anywhere on the diagram represents obtaining increasing gains for the self and horizontal movement to the left anywhere on the diagram represents obtaining increasing losses (i.e. negative gains) for the self. Similarly, upward vertical movement anywhere on the diagram represents obtaining increasing gains for another and downward vertical movement anywhere on the diagram represents obtaining increasing losses for another. Crucially, movement in one plane does not entail movement in the other, although movement in both planes is of course possible simultaneously.

As can be seen on the diagram, McClintock labels eight equally spaced vectors emanating from the origin, the origin representing obtaining no gains (positive or negative) for either the self or another. Obtaining gains for the self and neither gains nor losses for another is termed individualism; obtaining gains for both the self and another is called cooperation; and so on.

McClintock's social value vectors concern individual and interindividual behaviour but can readily be adapted to represent group and intergroup behaviour, as shown in the diagram below.

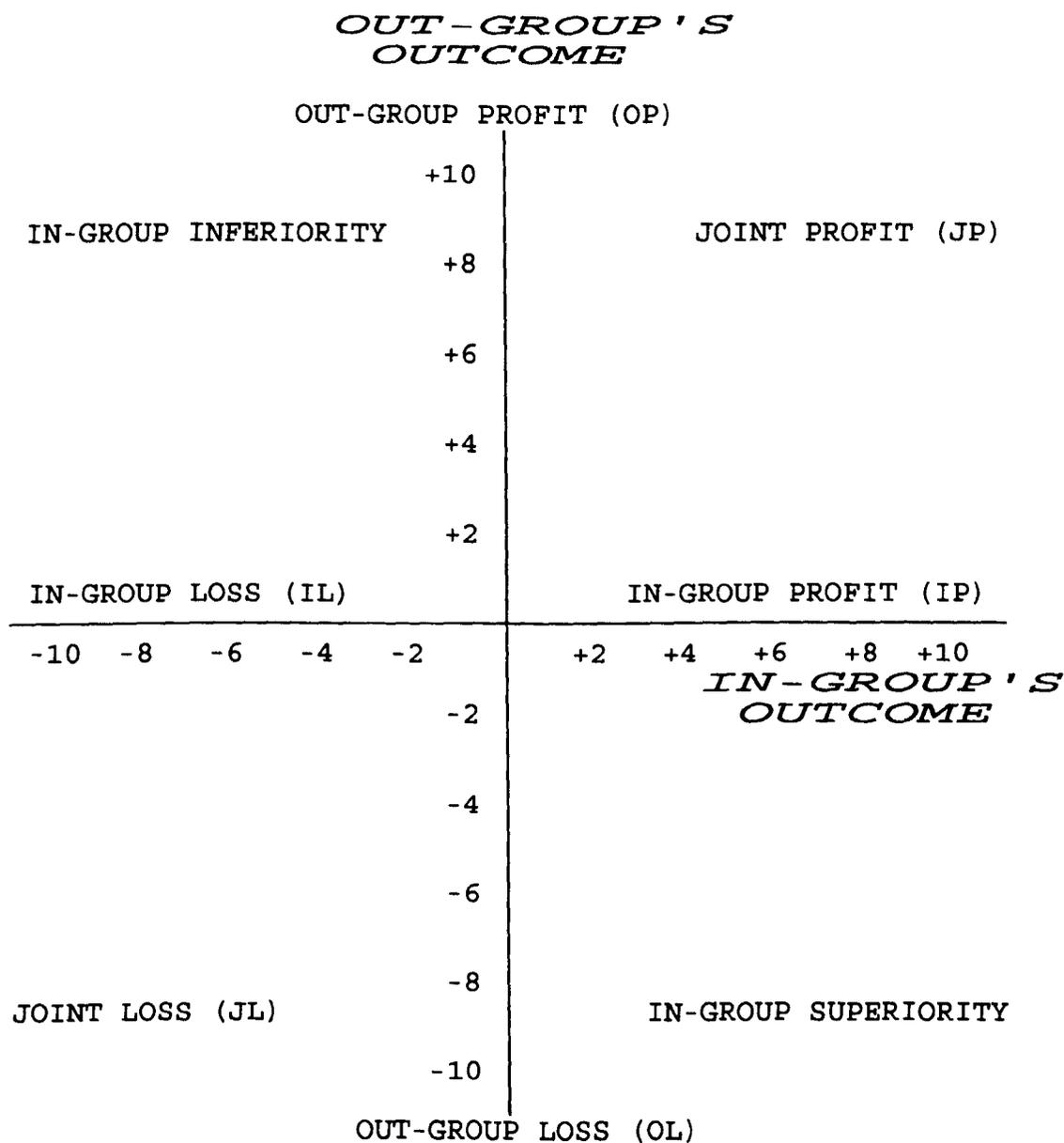


Diagram 3.2: Group and intergroup behaviour vectors

In this diagram horizontal movement to the right anywhere on the diagram represents obtaining increasing gains for the in-group or its members and horizontal movement to the left anywhere on the diagram represents obtaining increasing losses for the in-group or its members. Upward vertical movement anywhere on the diagram represents obtaining increasing gains for the out-group or its members and downward vertical movement anywhere on the diagram represents obtaining

increasing losses for the out-group or its members. And, more crucially than before, movement in one plane does not entail movement in the other, although simultaneous movement in both planes is of course possible.

The origin of the diagram represents obtaining neither gains nor losses for either in-group (members) or out-group (members).⁵ Vectors moving right and left from the origin are labelled in-group profit and in-group loss, respectively, as movement along these vectors will obtain increasing gains or losses for the in-group or its members but will obtain neither gain nor loss for the out-group or its members. Similarly, vectors moving up or down from the origin are labelled out-group profit and out-group loss, respectively, as movement along them will obtain gains or losses for the out-group or its members but will obtain neither gains nor losses for the in-group or its members.

The vector moving up from and to the right of the origin is labelled joint profit and the vector moving down from and to the left of origin is labelled joint loss. Movement along these vectors will obtain gains and losses, respectively, for both the in-group and the out-group or their members, but will not obtain either positive or negative in-group or out-group distinctiveness. Finally, the vector moving down from and to the right of origin is labelled in-group superiority and the vector moving up from and to the left of origin is labelled in-group inferiority. Movement along these vectors will clearly obtain in-group superiority over the out-group (and out-group inferiority to the in-group) and in-group

⁵ Note that deviations from the origin represent gains or losses from what the recipients "deserve to" or "should" receive. This means that the value of the origin can be set, and set differently for each recipient if need be. If, for example, one recipient scored 20 on a test and the other scored 10, these are the values that the origin would be set at, as it is deviations from these "objectively deserved" values which represents bias.

inferiority to the out-group (and out-group superiority over the in-group), respectively.

The design of a new dependent measure of intergroup behaviour follows easily and obviously from the previous diagram. All that is needed is to present subjects with a series of "allocation grids" based on the diagram with various patterns of in-group and out-group recipients identified on each axis.⁶ Subjects would then make a single mark somewhere on the diagram indicating how many gains or losses they wanted each recipient to obtain. The mean response could then be presented either diagrammatically or simply as a pair of figures reflecting mean allocations made to the in-group and to the out-group.

These figures, representing in-group profit/loss and out-group profit/loss, can, if desired, be used to calculate joint profit/loss (i.e. the two figures added together and divided by two) and in-group superiority/inferiority (i.e. in-group profit minus out-group profit and then divided by two). This provides scores for each axis of the diagram above, and the sign of the scores reveals which side of origin subjects' responses are on those axes. A fairness score is obtainable simply by subtracting the actual absolute in-group inferiority/superiority score from the maximum possible absolute in-group inferiority/superiority score.

A score for in-group favouritism, as conceptualized on the Tajfel matrices, is also obtainable, simply by taking the

⁶ With two same-group recipients the allocation is of course intra-group and the labels on each vector need to be amended. With member X as the vertical axis recipient and same-group member Y as recipient on the horizontal axis the vectors will be labelled as follows, running clockwise from the 12 o'clock position:- X-profit, joint profit, Y-profit, Y-superiority, X-loss, joint loss, X-loss, and X-superiority. Unless stated otherwise all further discussion will concern intergroup grids.

greater positive (or least negative) value of the scores of in-group profit and in-group superiority. It is not considered that such a score would usually be calculated, however, as it conflates the two types of bias considered earlier. Absolute bias in favour of the in-group is given by the in-group profit score and absolute bias against the in-group is given by the in-group loss score, just as absolute bias in favour of the out-group is given by the out-group profit score and absolute bias against the out-group is given by the out-group loss score. Relative bias in favour of the in-group (and against the out-group) is given by the in-group superiority score, while relative bias against the in-group (and in favour of the out-group) is given by the in-group inferiority score.⁷ Because absolute and relative biases in favour of the in-group might be two conceptually distinct forms of discrimination, it is thought that an in-group favouritism measure might be more confusing than it is enlightening.

The allocation grids would appear to have a number of advantages over the Tajfel matrices. First, the allocation grids yield measures which are undoubtedly of interval scale with equal origin and unit. Comparisons across studies therefore requires at most a conversion of scale. There is only one type of grid so comparisons across grid "types" does not arise.

Second, and relatedly, the main dependent measure is a single pair of scores (representing in-group and out-group gain or loss) from which a number of other scores can be calculated if required. This is in stark contrast with the Tajfel matrices which provide any number of scores according to which "pulls" are being investigated.

⁷ It is also possible to conceptualise joint profit as a "relative" bias in favour of both groups and joint loss as a "relative" bias against both groups. Although valid this conceptualization is simply too complex and potentially confusing to sensibly adopt.

Third, and again because there is only one grid type, any strategy which subjects can pursue on one grid they can pursue on all grids with the same recipient pattern, as any structural restrictions present on one grid will be present on them all. Subjects will therefore never be forced into second-choice strategies.

Fourth, because all measures can be obtained from a single presentation of a grid it is practically possible to administer several grids with the same recipient pattern, thus avoiding single-shot measures and improving the reliability of the social behaviour scores obtained.

Fifth, the primary dependent measure (i.e. the pair of scores reflecting in-group and out-group loss) has unique values for every possible position on the grid, thereby allowing greater clarity than provided by the Tajfel matrix pull scores.

Sixth, there is no need to make a "numerical assumption" as the dependent measure directly represents the allocations made to in-group and out-group, again maximizing clarity.

Seventh, the only structural restrictions placed on subjects are ones unlikely to present them with problems. Subjects cannot pursue a strategy of maximum joint profit combined with maximum unfairness, maximum in-group inferiority and minimum individual profit for one of the recipients, for example. Although pursuit of such a strategy is occasionally possible on the Tajfel matrices (i.e. on i/o presentations of matrix type 1 - see Appendix 1), it is assumed here that if someone wishes to maximize joint profit they are unlikely to also wish to actively pursue maximum unfairness, in-group inferiority and/or individual loss. It is also assumed that the eight vectors represent all strategies (simple or complex) that subjects are likely to wish to pursue.

Eighth, the grids allow negative as well as positive allocations, thus enhancing content validity. Content validity is further enhanced by the fact that subjects' scores on the grids can be generalized more easily than can subjects' pull scores derived from the Tajfel matrices. As mentioned above, pull scores, as their name suggests, are scores of the "pull" of one strategy on another when those strategies are in opposition. If a subject obtains a high F on FAV score one cannot conclude that the subject "is" fair (i.e. that will tend to act fairly in situations similar to the one they were in when the pull score was taken). All that can be concluded is that the subject will prefer fairness to in-group favouritism in those situations. It cannot be concluded that the subject will prefer fairness when it is placed in opposition to other strategies (such as maximum in-group profit), or when the subject is free to pursue any strategy. With the grid measures, however, subjects are free to follow any of a range of strategies and therefore any strategy they chose is likely to be relatively generalizable.

Finally, possible allocation of loss as well as profit allows investigation of the "sort" of in-group favouritism subjects engage in during minimal group experiments. Do subjects attempt to maximize in-group profit regardless of how the out-group fares, do they engage in absolute bias in favour of both in-group and out-group combined with relative bias in favour of the in-group, or do they obtain their relative bias in favour of the in-group in combination with absolute bias against the out-group? It seems reasonable to claim that relative bias in favour of the in-group is a far more serious affair when combined with absolute out-group derogation than it is when combined with absolute out-group enhancement only slightly less than the absolute in-group enhancement.

In the study reported in this chapter both measures of social behaviour were employed in the minimal group paradigm in order to investigate the nature of minimal group behaviour indicated

by each. Of particular interest is the nature of absolute bias(es) accompanying the relative bias of positive in-group distinctiveness.

METHOD⁸

Participants

33 first-year psychology undergraduates from Keele University volunteered to be subjects in a study in partial fulfilment of research participation requirements. The study was introduced as "A study of decision making in group situations." The sample was predominantly female (M = 3, F = 33) and ages ranged from 18 to 46 years old, with a mean of 24.3 years.

Procedure

Subjects were seen in two sessions, one of 16 and one of 17 subjects. The second session was immediately after the first. In each session it was explained that each person present would be randomly allocated as a member of one or the other of two groups and would then make a series of point allocations to pairs of other people about whom they would know nothing except their unique code number and their group membership. Subjects were told that no-one would ever be able to make an allocation to themselves, but that a record of points allocated to them by others would be kept and added to or taken away from a 400 point collection each subject started with.⁹ It was emphasized that the points would *not* be exchanged for

⁸ Farsides (1994) was based on selected results from this study.

⁹ 400 points were "given" to each subject so that early allocations of out-group loss would not be thought by them to necessarily result in out-group "debt" to the experimenter.

anything else (e.g. money), but that the experimenter would let them know each individual's and each group's total at the end of the week. It was also emphasized that making particular allocations to others would not affect subjects' own point collections.¹⁰ Both the Tajfel allocation matrices and the new allocation grids were explained and demonstrated to subjects, and any questions they had were answered.

Subjects were then categorized by the experimenter randomly pulling envelopes containing response booklets from a bag and giving one envelope to each subject. Each subject then moved to a private part of the (large) room and completed their task individually. When the task was complete the experimenter collected all the response booklets, fully debriefed the subjects and answered any questions they had. Subjects were then thanked for their participation and dismissed.

Materials

Each envelope contained a response booklet with information about the subject's unique code number and minimal group membership shown in large letters on the front cover. This information was repeated in bold in the bottom right hand corner of every page of the booklet. The first few pages of the booklets repeated in writing the main task instructions and gave further worked examples of how to use the allocation matrices and grids.

Each response booklet contained blocks of 12 allocation matrices and 12 allocation grids. The order of the blocks was counterbalanced. Three matrix types were used (see Appendix 1), each of which was presented once with each of the

¹⁰ This was necessary because a pilot study revealed that some subjects refrained from giving others points because they worried that the points allocated would come out of their own stock.

following recipient patterns:- i/o, o/i, i/i and o/o.¹¹ The single grid type was presented twelve times, three times with each of the aforementioned recipient patterns. (An i/o recipient pattern on an allocation grid is one in which in-group profit and loss occurs in the horizontal plane and out-group profit and loss occurs in the vertical plane, while the reverse is true for o/i presentations.) Grid and matrix orders were randomized within their respective blocks. An example allocation grid is shown in Appendix 6.

Following the allocation measures each booklet contained a short questionnaire with open-ended questions asking subjects to give an account of their thoughts and feelings during the study, in particular in terms of what they were trying to do, any difficulties they had and what they did about them, whether and why they changed their strategy during the study, whether they thought about how others might be behaving, and what they thought the outcome of the scoring might be.

Scoring

Pull scores and indirect in-group favouritism scores were calculated for the Tajfel matrices in the usual way (see Appendix 1 for details).

Grid scores were calculated as follows:-

IP (i.e. total in-group profit) = mean allocation to in-group members across all grids;

OP (i.e. total out-group profit) = mean allocation to out-group members across all grids;

IS (i.e. total in-group superiority) = $(IP - OP)/2$;

JP (i.e. total joint profit) = $(IP + OP)/2$;

¹¹ i = in-group member, o = out-group member.

F (i.e. total intergroup fairness) = maximum possible score - $(|IP - OP|)$; and,

FAV (i.e. total in-group favouritism) = greatest positive value of IP or IS.

In order to maximize comparability between grid and pull scores the former were all converted to range from -12 to +12, except F grid scores which were converted to range from 0-12.¹²

Analysis of the post-task questionnaire

Subjects' answers to the open-ended questions of the post-task questionnaire (which can be seen in full in Appendix 7) were subjected to a very basic quantitative content analysis by the experimenter. Legible and seemingly unambiguous answers to questions asking whether or not subjects experienced difficulties, altered their strategies or considered others' behaviour whilst making their allocations were coded either in the affirmative or the negative. For all other answers categories were generated from the data, with a new category being generated every time a subject's answer did not seem to easily and cleanly fall into a previously generated category.

¹² Minus F scores have no meaning on the grids, whereas minus scores for other strategies (except FAV, which is not usually calculated) indicate positive scores for the strategies represented by the opposite vector. E.g. -IP = +IL.

RESULTS

Order effects

Matrix/grid versus grid/matrix order effects were investigated using the MANOVA procedure of SPSS, with the following dependent measures as within-subject factors:- the six pull scores, the two indirect in-group favouritism scores, IP and OP (remembering that all other grid scores are merely transformations of IP and OP and are therefore not independent of them). No significant order effects were found and this variable was collapsed for all subsequent analyses.

Grid score reliability

IP and OP grid scores are means of multiple allocations to in-group and out-group members, respectively. In the present study these scores are the means of 12 allocations to members of each group. Of these 12 allocations 6 occurred on intergroup grids and 6 on same-group grids.

Reliability of IP and OP was calculated twice. When all allocations to in-group or out-group members were considered as individual items Cronbach's alpha was 0.86 for IP and 0.92 for OP. Separate scales were then calculated for intergroup and for same-group grid presentations. This produced alphas of 0.71 and 0.86 for IP and OP respectively on the intergroup presentations and 0.82 and 0.83 on the same-group presentations. Reliabilities were then calculated for the "total" IP and OP scales using these sub-scales as separate items. When this was done, alpha for IP was 0.82 and for OP it was 0.96.

IP and OP grid scores therefore had very acceptable internal reliability.

Sample pull and indirect in-group favouritism scores

Mean pull and indirect in-group favouritism scores are shown below.¹³

<u>Measure</u>	<u>X</u>	<u>(SD)</u>
FAV on MJP	3.485	(6.695)*
MD on MIP+MJP	2.242	(6.633)
FAV on F	3.909	(6.222)*
MJP on FAV	-0.152	(3.817)
MIP+MJP on MD	0.364	(5.225)
F on FAV	1.242	(6.586)
IndFAV1	0.636	(4.568)
IndFAV2	2.970	(6.136)*

*: $p \leq 0.005$; **: $p \leq 0.01$; †: $p \leq 0.05$, all two-tailed.

Table 3.3: Mean pull and indirect in-group favouritism scores

These results show that the pulls of FAV on F and FAV on MJP were significantly greater than zero, as was the matrix type 2 indirect in-group favouritism score.

Grid scores

In-group and out-group profit grid scores were calculated, as were other grid scores derived from these measures in order to maximize comparability with the matrix pull scores. All these grid scores are shown below.

¹³ See Appendix 1 for how these scores are derived.

<u>Measure</u>	<u>X</u>	<u>(SD)</u>
IP	4.730	(3.903)*
OP	-1.649	(5.482)
IS	3.189	(4.327)*
JP	1.541	(1.980)*
F	8.420	(3.999)*
FAV	4.905	(3.710)*

*: $p \leq 0.005$; #: $p \leq 0.01$; *: $p \leq 0.05$, all two-tailed.

Table 3.4: Mean grid scores

All grid scores except the one for out-group profit were significantly different from zero. These scores reveal that subjects showed mean absolute bias in favour of the in-group but showed no significant mean absolute bias toward the out-group. This resulted in significant mean in-group superiority (i.e. relative bias in favour of the in-group). Mean absolute bias in favour of the in-group exceeded the non-significant mean absolute bias against the out-group, so there was significant mean joint profit. Subjects' mean relative bias in favour of the in-group was not even half of what it could have been and as a result subjects' allocations were considerably more fair than they were discriminatory.¹⁴

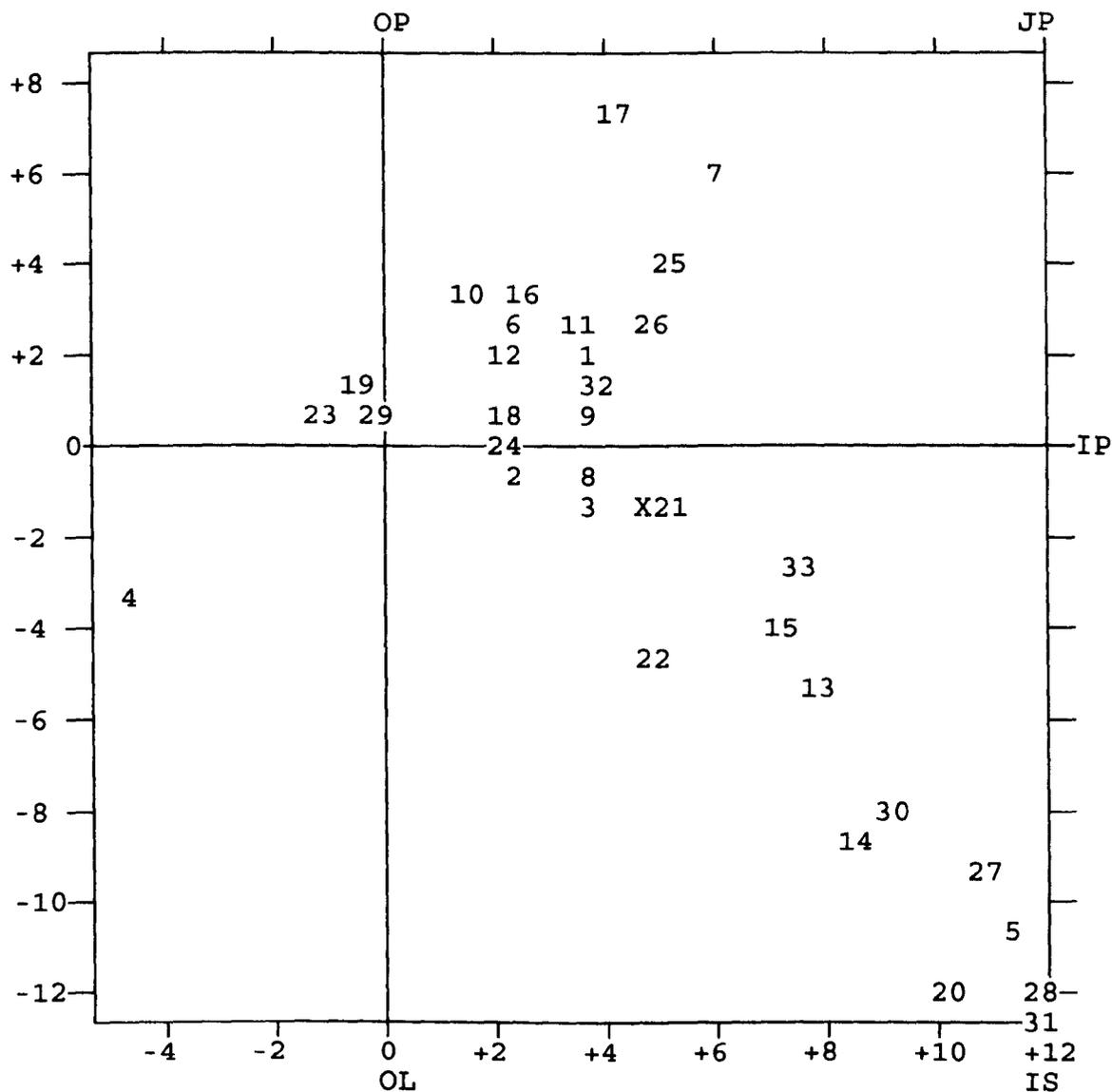
The two methods of measuring social behaviour yield similar but not identical results. Both suggest that subjects made allocations which resulted in in-group profit and in-group

¹⁴ The F score and the absolute value of the IS score do not add up to the maximum possible score (i.e. +12) because some in-group inferiority allocations were made. When the F scores and the absolute values of the IS scores of individual subjects are examined, they always add up to +12. Had the mean absolute in-group superiority/inferiority score been calculated instead of the in-group superiority score, this would also have made a combined total of +12 when added to the mean F score.

superiority, but whereas the pull scores showed that fairness exerted no significant pull on discriminatory tendencies, the grid scores indicated that subjects' allocations were more fair than they were discriminatory. The grid scores also revealed slight but significant joint profit, although such a strategy did not exert a significant pull against in-group favouritism on the Tajfel matrices.

Presenting subjects' responses diagrammatically

One of the great advantages of using grids to derive allocation scores is that subjects' responses can be presented diagrammatically. This is done below with IP/IL on the horizontal axis and OP/OL on the vertical axis. The numbers in the diagram represent subjects' respondent numbers. The "X" shows the approximate overall mean response as obtained above.



Key: IP = In-group Profit; JP = Joint Profit; OL = Out-group Loss; IS = In-group Superiority; OP = Out-group Profit; X = Overall Sample Mean IP and OP scores.

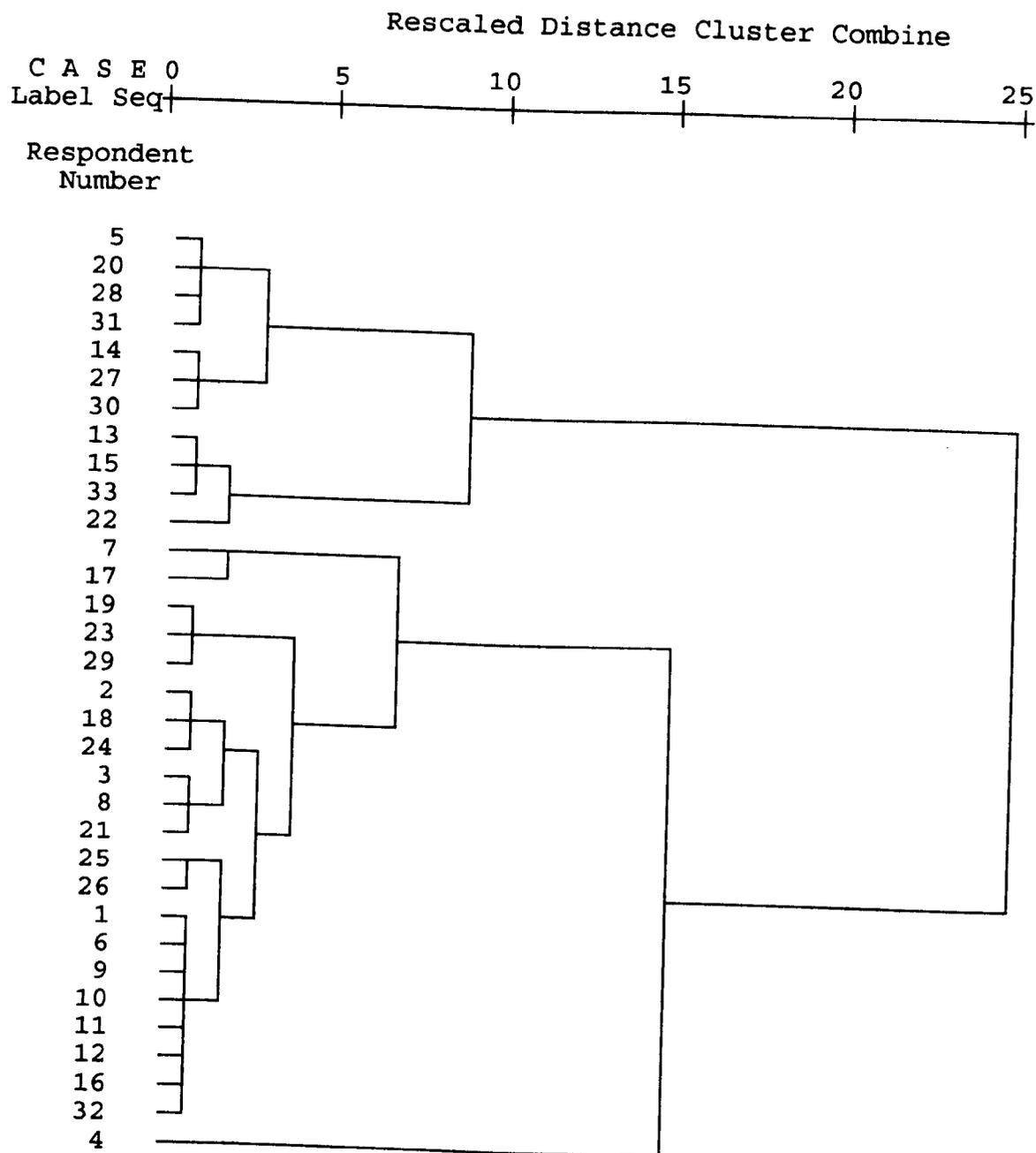
Diagram 3.5: Diagrammatic representation of subjects' allocations, by subject numbers

This diagram reveals that subjects engaged in a range of behaviours, very few of which are adequately represented by the overall mean IP and OP scores. Responses fell almost exclusively into two quadrants: either the one which obtains in-group superiority or the one which obtains joint profit. Within those quadrants certain responses seem to fall into clusters:- one collection of subjects (subjects 5, 14, 20, 27,

28, 30 and 31) appears to maximize or nearly maximize both in-group profit and out-group loss, another (subjects 13, 15, 22 and 23) appears to obtain relatively moderate in-group profit and less pronounced out-group loss, and so on.

Cluster analysis of subjects' IP and OP scores

A cluster analysis was carried out to see whether subjects could be clustered according to their IP and OP grid scores. The dendogram (using average linkage between groups) below was the result.



*Diagram 3.6: Dendrogram showing clusters by
IP and OP grid scores*

This dendrogram closely reflects the grid above.¹⁵ Behaviours fall into two main clusters: one represented by in-group superiority (subjects 5 down to 22 on the dendrogram) and the other represented by joint profit (subjects 7 down to 32 on the dendrogram). Subject 4 is in a category of their own: joint loss.

Within the first cluster there are two main sub-clusters. One (subjects 5, 14, 20, 27, 28, 30 and 31) can be characterized as "strong ethnocentrism", with subjects maximizing in-group superiority by maximizing or nearly maximizing both in-group profit and out-group loss. The other (subjects 13, 15, 22 and 33) can be thought of as "moderate ethnocentrism", with subjects obtaining both moderate in-group profit and less pronounced out-group loss. (It is possible to further sub-divide the first of these sub-clusters, with subjects 14, 27 and 30 being thought of as the strong ethnocentrics, while subjects 5, 20, 28 and 31 can be thought of as "extreme" ethnocentrics.)

The second cluster sub-divides a little less cleanly. One sub-cluster of subjects (25 down to 32 on the dendrogram) can be thought of perhaps as subjects who made slight but roughly equal joint profit allocations. A second sub-cluster of subjects (2 down to 21 on the dendrogram) are perhaps more concerned with obtaining slight but positive in-group profit whilst keeping out-group profit and loss negligible. A small number of subjects (19, 23 and 29) might be thought of as having made non-allocations, obtaining neither profit nor loss for either in-group or out-group. Two subjects (7 and 17) allocate relatively equally to in-group and out-group and obtain moderate to strong joint profit.

¹⁵ This was to be expected as cluster analysis with two variables (i.e. IP and OP) clusters items according to their distance in two-dimensional space.

Cluster analysis of subjects' pull- and indirect in-group favouritism scores

If IP and OP suggest that a variety of characterizable minimal group behaviours obtain in the present study it is important to see if the scores obtained from the Tajfel matrices suggest the same thing. Unlike the grid scores, pull-scores cannot be presented diagrammatically. They can however be entered into a cluster analysis, the resultant dendogram of which is shown below.

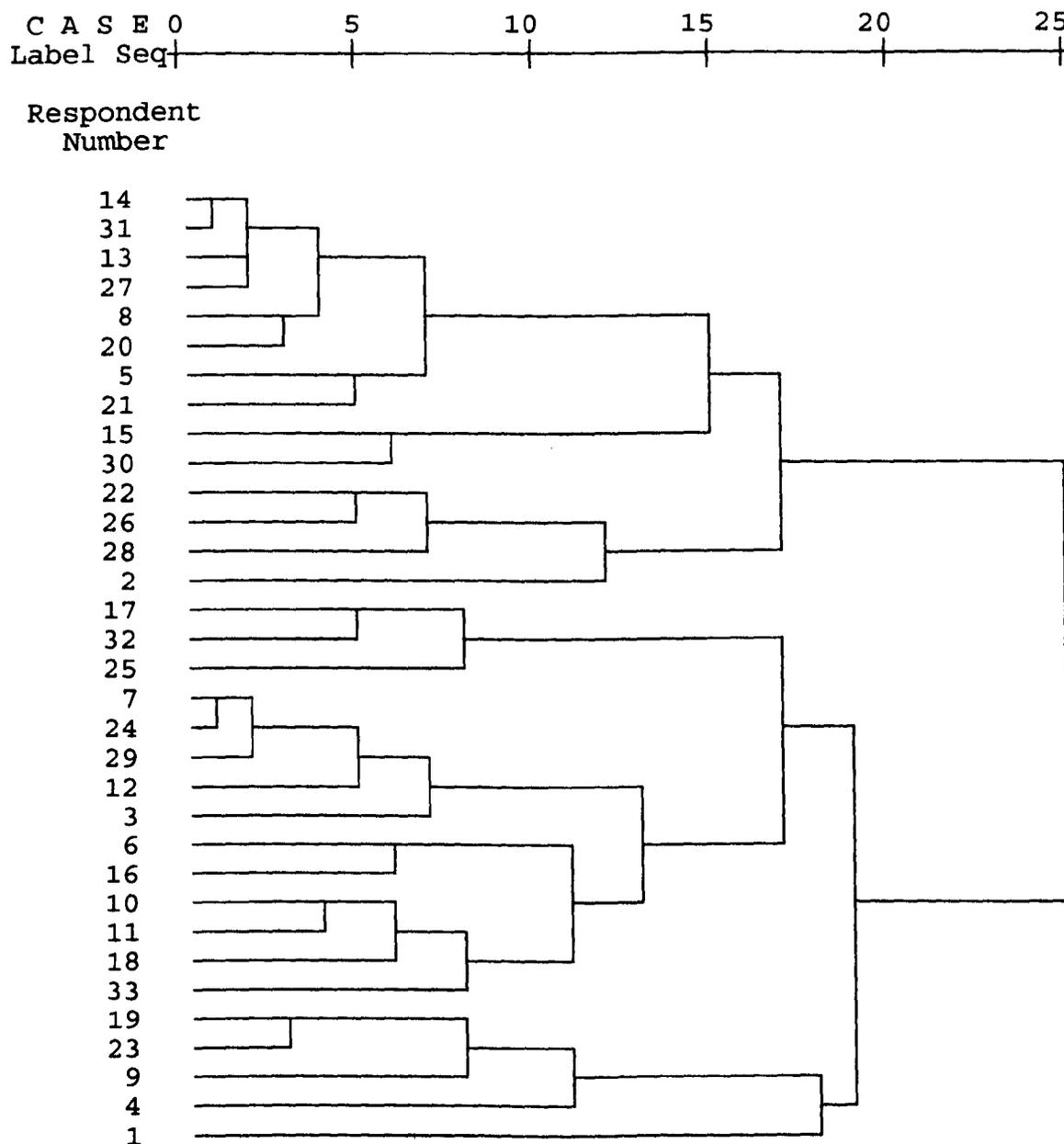


Diagram 3.7: Dendrogram showing clusters by matrix scores

Again two main clusters are derived, but it is rather hard to determine where the cut-off point for sub-clusters should be (bearing in mind that we should not be using the IP/OP diagram to help us). To interpret the dendrogram we need to calculate the mean pull- and indirect in-group favouritism scores of each set of clusters considered. When this is done for the two main clusters it is quite clear that the first cluster

(subjects 14 down to 2 on the dendogram above) can be labelled "in-group favouritism" while the second cluster can be characterized as "fair" or perhaps simply as "the absence of discrimination" (cluster 1: FAV on MJP = 9.1 (4.3), MD on MIP+MJP = 6.3 (7.0), FAV on F = 8.5 (4.4), MJP on FAV = 1.9 (3.5), MIP+MJP on MD = 1.3 (3.6), F on FAV = -2.5 (4.5), IndFAV1 = 2.4 (4.1), IndFAV2 = 7.6 (4.3); cluster 2: FAV on MJP = -0.7 (4.8), MD on MIP+MJP = -0.7 (4.6), FAV on F = 0.5 (5.1), MJP on FAV = -1.6 (3.4), MIP+MJP on MD = -0.3 (6.2), F on FAV = 4.0 (6.4), IndFAV1 = -0.6 (4.6), IndFAV2 = -0.4 (5.0)).¹⁶

Two main sub-clusters exist within the in-group favouritism cluster, one consisting of subjects 5, 8, 13, 14, 15, 20, 21, 27, 30 and 31, and the other of subjects 2, 22, 26 and 28. Three main clusters exist within the main "fairness" cluster, the first consisting of subjects 17, 25 and 32; the second of subjects 3, 6, 7, 10, 11, 12, 16, 18, 24, 29 and 33; and the third of subjects 4, 9, 19 and 23, with subject 1 not falling into any of these sub-clusters. To interpret these sub-clusters it is again necessary to examine the mean pull- and indirect in-group favouritism scores for each. These are shown below (along with the corresponding IP and OP scores).

¹⁶ Examining the IP and OP scores for these clusters supports such an interpretation. Cluster 1: IP = 7.7 (3.2), OP = -6.2 (5.0). Cluster 2: IP = 2.5 (2.8), OP = 1.7 (2.7).

Cluster	In-group Favouritism		Fairness				
	Sub-Cluster	1	2	1	2	3	4
n		10	4	3	11	4	1
FAV on MJP		11.5 (1.6)	3.3 (2.6)	-1.0 (3.5)	-0.4 (3.4)	-4.5 (4.4)	12.0 (0.0)
MD on MIP+MJP		9.9 (3.3)	-2.8 (5.0)	1.0 (1.7)	1.5 (3.0)	-6.3 (3.8)	-8.0 (0.0)
FAV on F		9.1 (4.7)	7.0 (3.8)	-0.3 (2.5)	3.5 (4.2)	-5.3 (2.1)	-6.0 (0.0)
MJP on FAV		0.5 (1.6)	5.3 (4.9)	-1.0 (3.5)	-1.5 (3.0)	-3.0 (5.4)	0.0 (0.0)
MIP+MJP on FAV		1.5 (2.8)	0.8 (5.8)	9.0 (3.0)	-2.9 (5.5)	1.5 (2.8)	-2.0 (0.0)
F on FAV		-1.7 (5.3)	-4.5 (3.4)	9.7 (2.5)	4.9 (6.1)	-3.3 (3.5)	6.0 (0.0)
IndFAV1		2.7 (4.9)	1.5 (1.3)	-3.3 (6.7)	0.6 (3.6)	-3.0 (5.1)	3.0 (0.0)
IndFAV2		7.9 (4.7)	6.8 (3.9)	5.3 (2.1)	-3.6 (3.7)	3.3 (2.6)	3.0 (0.0)
IP		8.5 (2.6)	5.9 (4.2)	4.3 (0.7)	3.0 (2.1)	-0.6 (3.7)	3.7 (0.0)
OP		-7.3 (4.1)	-3.5 (6.4)	4.3 (3.4)	1.6 (2.5)	-0.1 (2.1)	2.5 (0.0)

Table 3.8: Mean matrix and (selected) grid scores by clusters shown in illustration 3.7

From these figures it is clear that the first in-group favouritism sub-cluster has considerably stronger FAV on MJP and MD on MIP+MJP pull scores and slightly stronger FAV on F and the indirect in-group favouritism scores than the second sub-cluster. The second in-group favouritism sub-cluster seems to have a stronger MJP on FAV pull than the first, although the former's small number of subjects and relatively large standard deviation make this a tentative difference.

The first in-group favouritism sub-cluster, then, shows a stronger and a more general tendency toward in-group favouritism than the second sub-cluster. It can be seen that the IP and OP grid scores are consistent with such an interpretation.

The first "fairness" sub-cluster has stronger MIP+MJP on MD and F on FAV pull scores than do the other two sub-clusters; the second "fairness" sub-cluster seems to deviate very little from zero on any of the measures; and the third "fairness" sub-cluster obtains relatively strong negative scores on the three direct discrimination pull scores (although once again small subject numbers and relatively large standard deviations make any interpretation of this sub-cluster tentative). These "fairness" sub-clusters might be labelled then as "extreme fairness", "no strategy" and "moderate out-group favouritism" respectively. The IP and OP scores are not especially consistent with these interpretations. According to these scores the first "fairness" sub-cluster would be labelled "extreme fairness plus moderate joint profit", the second would be labelled "mild fairness, mild joint profit and possible mild in-group superiority", and the third would be labelled "extreme fairness accompanied by neither joint loss nor joint profit".

The pattern of the matrix-derived scores for subject 1 suggest an arbitrary allocation strategy. The IP and OP scores for this subject suggest instead strong fairness plus slight joint profit.

Subjects' self-reported strategies

Subjects were, it should be remembered, asked how they had made their allocations during the post-task questionnaire. Eighteen distinguishable strategies were employed. Subjects' reports were therefore re-categorized into "dominant"

strategies, according to which single strategy seemed to be the most important to them. Five dominant strategies were derived from subjects' self-reports: fairness (8 subjects), maximum in-group profit (1 subject), in-group superiority (5 subjects), strong ethnocentrism (8 subjects) and random (11 subjects).

	Dominant self-reported strategy				
	F n = 8	MIP n = 1	IS n = 5	Ethno n = 8	Random n = 11
FAV on MJP	0.8 (5.0)	12.0 (0.0)	2.2 (7.3)	11.3 (2.1)	-0.4 (4.8)
MD on MIP+MJP	-2.1 (4.8)	12.0 (0.0)	1.2 (7.3)	9.0 (5.0)	0.1 4.4)
FAV on F	-2.1 (3.3)	9.0 (0.0)	8.4 (5.0)	8.8 (5.2)	2.3 (4.7)
MJP on FAV	-1.3 (2.7)	0.0 (0.0)	0.2 (4.9)	0.8 (2.1)	-0.2 (5.2)
MIP+MJP on MD	2.1 (5.1)	0.0 (0.0)	-3.2 (6.3)	1.3 (2.8)	0.1 (6.2)
F on FAV	8.6 (4.9)	3.0 (0.0)	-1.6 (1.8)	-2.5 (5.7)	-0.3 (6.1)
IndFAV1	0.3 (2.4)	6.0 (0.0)	2.0 (3.7)	1.6 (5.2)	-0.9 (5.6)
IndFAV2	-1.1 (5.7)	12.0 (0.0)	3.2 (5.7)	8.4 (4.1)	1.1 (4.9)
IP	2.8 (2.3)	7.8 (0.0)	4.7 (1.9)	10.0 (1.7)	2.1 (2.8)
OP	3.0 (2.7)	-5.6 (0.0)	-1.3 (2.9)	-9.6 (2.8)	1.0 (2.1)
IS	-0.1 (0.9)	6.7 (0.0)	3.0 (2.3)	9.8 (2.2)	0.6 (1.2)
JP	2.9 (2.4)	1.1 (0.0)	1.7 (1.0)	0.2 (0.8)	1.5 (2.2)
F	11.4 (0.6)	5.3 (0.0)	8.8 (1.8)	2.2 (2.2)	10.9 (0.6)
FAV	2.8 (2.3)	7.8 (0.0)	4.7 (1.9)	10.2 (1.8)	2.5 (1.9)

Table 3.9: Mean matrix and grid scores by dominant self-reported strategies

Pull-, indirect in-group favouritism and grid scores were calculated for each category and are shown above (with standard deviations in brackets).

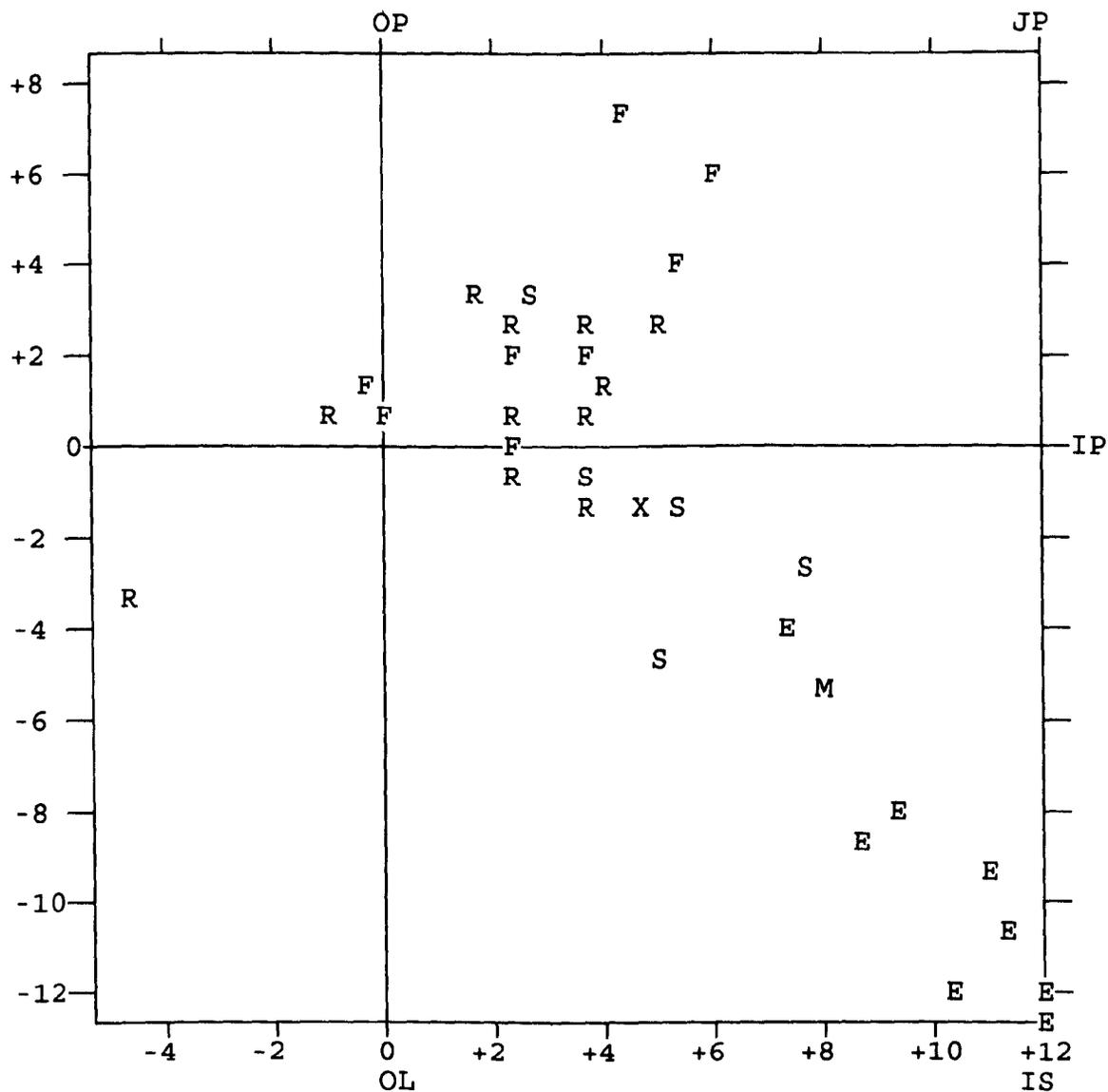
Scores derived from the Tajfel matrices did rather well at reflecting subjects' dominant self-reported strategies. Subjects whose dominant strategy was random obtained negligible scores for all pull and indirect in-group favouritism scores. Subjects whose dominant strategy was fairness obtained negligible scores for all these scores except for the pull of F on FAV, where they obtained a higher mean score than all the other categories of subjects. Both the subjects interested in in-group superiority and the subjects interested in strong ethnocentrism obtained negligible scores on the three non-discriminatory pull scores (and also on IndFAV1) and obtained high but equal scores on the FAV on F measure. The ethnocentric subjects, in addition, obtained much higher mean scores on the pulls of FAV on MJP and MD on MIP+MJP, and also on IndFAV2. The scores obtained by the single subject with the dominant self-reported strategy of maximum in-group profit suggest that he was motivated rather more by ethnocentrism than simple in-group profit (especially as the maximum MD on MIP+MJP score required the subject to chose allocations which minimized in-group profit and the maximum IndFAV2 score required the subject to minimize out-group profit with no attendant effect on in-group profit).

The grid scores of IP and OP (and the scores derived from them) distinguished well between subjects whose dominant strategy was in-group superiority and subjects whose dominant strategy was strong ethnocentrism. The latter subjects allocated more to the in-group than did the former subjects and also obtained considerable out-group loss, which the in-group superiority subjects did not do. Again the pattern of scores obtained by the "maximum in-group profit" subject suggests that they were more closely aligned to the strong ethnocentrics than to the subjects pursuing only in-group superiority.

The subjects whose dominant self-reported strategy was fairness obtained almost the maximum possible mean fairness

grid score, but so did the subjects whose dominant self-reported strategy was randomness. Indeed, the only differences between these two categories of subject are the out-group profit and joint profit scores each obtained, and these differences are slight. This suggests either that the grid scores are deficient in not being able to distinguish between strategies of fairness and of randomness or that the subjects whose dominant self-reported strategies were randomness or fairness were not actually acting as dissimilarly as those two labels would suggest.

The diagram below again shows IP and OP scores for each subject, but this time the subjects are identified by their dominant strategies instead of their respondent numbers. (F = Fairness, M = Maximum in-group profit, S = In-group Superiority, E = Ethnocentrism, and R = Random.)



Key: IP = In-group Profit; JP = Joint Profit; IP = In-group Profit; IS = In-group Superiority; OL = Out-group Loss; X = Overall Sample Mean IP and OP scores.

Diagram 3.10: Diagrammatic representation of subjects' allocations, by dominant strategies

The first thing to notice is that the random subjects almost all fall on or very close to the plane of maximum fairness. If they did in fact make random allocation decisions it is quite clear that they did so in a way which allocated equal points to in-group and out-group members. Indeed, many of them ensured that both in-group and out-group members would receive equal and positive points (i.e. fall on or near the

plane of maximum fairness within the joint profit quadrant). It is hardly surprising, then, that the grid scores of these subjects are almost indistinguishable from subjects explicitly pursuing fairness.

The second thing to note is that subjects' self-reported strategies match very closely the interpretation of the clusters obtained when IP and OP scores were cluster analyzed. The 11 subjects in the "discrimination" cluster include all 8 of the subjects with an ethnocentric dominant self-reported strategy, along with the single subject with a dominant self-reported strategy of maximum in-group profit and two of the subjects whose dominant self-reported strategy was in-group superiority. Further, all 4 of the subjects in the extreme ethnocentrism sub-cluster and all three of the subjects in the strong ethnocentrism sub-cluster had dominant self-reported strategies of ethnocentrism. The moderate ethnocentrism sub-cluster contained one subject with a dominant self-reported strategy of ethnocentrism, two subjects with a dominant self-reported strategy of in-group superiority, and the single subject with the dominant self-reported strategy of maximum in-group profit.

Only three subjects with discriminatory dominant self-reported strategies were not included in the discrimination cluster. Of these one (subject 8) claimed that although they were trying to allocate more points to their group they did experience some difficulties with the system of point allocation on the grids (see Appendix 7). Another (subject 21) said that they hoped they were allocating more points to their own group but *"then would relent and give more points to"* the other group. The third subject (subject 16) did claim to be giving more points to their own group, but actually ended up giving approximately the same amount of points to both groups.

All other subjects had dominant self-reported strategies of either fairness or randomness. All except one of these (subject 4, with a random dominant strategy) fell within the "fairness" cluster.

A discriminant function analysis revealed that 66.67% of subjects could be correctly classified into dominant self-reported strategy categories by discriminant functions based on their IP and OP scores. Most of the error came from "fairness" subjects being incorrectly classified as "random" subjects (3 from 8 cases) or vice versa (4 from 11 cases). When another discriminant function analysis was carried out with "random" subjects excluded from the sample, 86.36% of subjects were correctly classified into their dominant self-reported strategy categories. The remaining error came from one "in-group superiority" subject being incorrectly assigned as a "fairness" subject; another "in-group superiority" subject being incorrectly categorized as a "maximum in-group profit" subject; and an "ethnocentric" subject also being incorrectly classified as a "maximum in-group profit" subject. From diagrams 3.5 and 3.10 above it can be seen that these are likely to be subjects 16, 22 and 15, respectively.

The clusters obtained from analysis of the Tajfel matrix measures also fare very well in reflecting subjects' dominant self-reported strategies. All eight of the "ethnocentric" subjects fell within the "in-group favouritism" cluster, as did three "in-group superiority" subjects, the single "maximum in-group profit" subject, and two "random" subjects. Seven of the eight self-proclaimed ethnocentrics fell within the strong in-group favouritism sub-cluster, along with two subjects pursuing in-group superiority and the subject claiming to be maximizing in-group profit. The weaker in-group favouritism sub-cluster contained the final ethnocentric, one subject pursuing in-group superiority and two subjects allocating randomly.

Only two subjects claiming to pursue in-group superiority did not fall within the discrimination cluster. One was the subject who claimed to be trying to allocate more points to their own group than to the out-group but who ended up allocating equal points to both according to the grid measures (subject 16). The other (subject 33) started making allocations favouring the in-group only after starting by making random allocations, and even then *"did not wish to appear too greedy so did not award" their in-group "excessively high points"* (see Appendix 7).

All other subjects had dominant strategies of either fairness or randomness and all fell within the "fairness" cluster.

A second discriminant function analysis was carried out to see how well the subjects could be classified according to their dominant self-reported strategies by discriminant functions based on the six pull- and two indirect in-group favouritism scores derived from the Tajfel matrices. 81.82% of subjects were correctly classified, whether or not random subjects were included in the analysis. When all subjects were included error derived from two "in-group superiority" subjects being incorrectly classified, one as an "ethnocentrism" subject and another as a "random" subject; two "ethnocentrism" subjects being incorrectly classified, one as a "maximum in-group profit" subject and the other as an "in-group superiority" subject; and two "random" subjects being incorrectly classified, one as a "fairness" subject and the other as an "in-group superiority" subject. With random subjects excluded from the analysis errors came from two "in-group superiority" subjects being classified as "ethnocentric" and from two "ethnocentric" subjects being misclassified, one as a "maximum in-group profit" subject and the other as an "in-group superiority" subject. Finding out which subjects were misclassified would require a comparison of each subject's eight scores compared with the means of those scores obtained

within each cluster and sub-cluster obtained from the cluster analysis above.

Both forms of dependent measure, then, discriminated well between subjects' dominant strategies when random subjects were excluded from the analysis, but the Tajfel matrix scores did considerably better than the grid scores when all subjects were included in the analysis.

Other post-task questionnaire results

Six subjects reported that they had encountered difficulties when making their allocations. Three subjects (subjects 8, 11 and 26) complained that the allocation grids were difficult to use.¹⁷ In one case this was because a subject (11) did not want to allocate negative points and was not sure how to avoid doing so. Subject 12 said it was difficult to *justify* giving anyone a negative score, although he seemed to know how to make such allocations. Another subject (28) experienced difficulties with the task itself as she felt that she had no reason to give or take points from anyone. The other subject who experienced a problem (25) simply had difficulty keeping track of which group she had allocated most points to.

A full third of the sample (i.e. 11 subjects) reported having changed strategies at least once during their allocation task. Only one subject (11) changed strategies because of encountering difficulties when making her allocations. Two subjects (9 and 24) did so to stop themselves becoming bored. Another two subjects (13 and 33) changed to strategies biased in favour of the in-group as the task progressed and they began to think in terms of their group membership. One subject (14) merely modified her strong ethnocentrism to a weaker one to be fairer and to ensure that other individuals

¹⁷ Also, subject 22 expressed the belief that other subjects would find the grids difficult.

did not receive more points than she did herself. Two subjects (2 and 28) altered their strategies both to keep themselves interested and because of a growing awareness of in-group membership. Subject 29 changed strategy several times for different reasons. Subject 17 changed from making random allocations within the joint profit quadrant of the grids to trying to achieve fairness and group profit on the matrices. Subject 19 gave no reasons for her self-professed strategy changes.

13 subjects thought about how other subjects might be making their allocation decisions whilst making their own, but came up with a variety of conclusions. Two subjects (15 and 24) correctly identified that other subjects would use a variety of strategies. Six subjects believed others would make allocations using the same strategies that they themselves used, namely fairness and/or in-group preference (subject 1), in-group superiority (8), maximum in-group profit (13), ethnocentrism (5 and 20), or randomness with in-group preference (18). One subject who made random allocations (4) wondered whether others would allocate points to their in-group, while another randomly allocating subject (23) thought that others would pursue "group-based" strategies. A subject who pursued fairness (7) wondered whether other subjects would be ethnocentric, while another who pursued in-group superiority (33) thought that others would allocate randomly.

Almost two thirds of the sample (19 subjects) reported wondering what the overall outcome of all the allocations might be. Ten subjects (i.e. approximately a third of all subjects) thought that everybody would show in-group preference. Of these ten (5 and 27) were ethnocentric themselves, another two (8 and 22) pursued in-group superiority, one (13) claimed to be maximizing in-group profit, another (25) was fair, and four (2, 9, 18 and 32) made predominantly random allocations. The implication of these subjects' beliefs is presumably that the overall outcome would

be equal between the groups. Five subjects explicitly thought that an equal outcome was likely, two (1 and 24) who pursued fairness, two (23 and 26) who allocated randomly and one (33) whose dominant strategy was in-group superiority. Only two subjects explicitly thought or hoped that their group would achieve superiority over the out-group, one of whom (21) pursued such in-group superiority and the other of whom (14) was a self-proclaimed ethnocentric. One randomly allocating subject (3) thought that the outcome would be random and one fair subject (17) thought that the overall outcome would be one which maximized joint profit and fairness. 14 subjects failed to report any thoughts about the overall outcome of the allocation task.

DISCUSSION

Social categorization and intergroup discrimination

Both types of dependent measure used in the present study indicated that social categorization resulted in significant sample-level in-group favouritism. At the level of the individual, however, social categorization was not sufficient to trigger in-group favouritism. Subjects adopted a variety of allocation strategies including fairness and randomness as well as (two forms of) in-group favouritism. As in the previous study the main conclusion has to be that mere social categorization is not sufficient to promote individuals to engage in discrimination in favour of the in-group.

The nature of minimal intergroup discrimination

Two types of discrimination were pursued by subjects in the present study. One involved subjects obtaining a relative bias in favour of the in-group (i.e. in-group superiority) via

a moderate to strong absolute bias in favour of the in-group and little or no absolute bias against the out-group. The other involved subjects obtaining the maximum possible relative bias in favour of the in-group via maximum absolute biases in favour of the in-group and against the out-group.

The matrix measures revealed that the latter type of discrimination, termed "strong ethnocentrism", was primarily aimed at obtaining the maximum possible relative bias in favour of the in-group, with the absolute biases being merely means to that end. Subjects pursuing this strategy were willing to forego both in-group profit (i.e. on i/o presentations of matrix type 2) and out-group loss (i.e. on o/i presentations of matrix type 2) in order to maximize in-group superiority.

Subjects pursuing the other type of discrimination, termed "in-group superiority" strove to achieve the maximum possible absolute bias in favour of the in-group which could be obtained in conjunction with a relative bias in favour of the in-group. These subjects were not interested in sacrificing in-group gain or maximizing out-group loss in order to maximize in-group superiority: their desire was to maximize in-group profit once in-group superiority was assured.

An evaluation of the grid scores

As mentioned in the introduction to this chapter, grid scores have a number of advantages over the matrix scores. These stem from allowing relatively unconstrained independent allocations of positive or negative goods to in-group and out-group members using multiple presentations of a single-format grid with equal interval and origin. This results in a clear, conceptually simple yet comprehensive measure (i.e. the IP/OP pair) with good internal reliability and with good content and external validity which can be presented diagrammatically and

which can be straightforwardly compared across studies. Importantly, it also allows different "sub-types" of the same "types" of behaviour to be distinguished one from the other.

The present study reveals three weaknesses the grids may have relative to the matrices. The first is the difficulty a small minority of subjects had in using them. This is relatively easily remedied. As well as improving instruction prior to use of the grids, the grids themselves can be made easier to use. One possibility is simply making the grids smaller, with possible allocations being restricted to a range of -5 to +5. Another is to computerize the grids so that movement within the grid is automatically accompanied by a simple readout showing how many points each recipient will receive if a particular choice is made.¹⁸

The second potential weakness the grids have compared to Tajfel matrices is that the latter make it easier to compare the strength of one particular strategy relative to another. Use of the Tajfel matrices in the present study, for example, allowed us to determine that ethnocentric subjects were willing to sacrifice both in-group profit and out-group loss in order to maximize in-group superiority. As the grids do not force subjects to make such choices it would have been very difficult, if not impossible, to determine this from grid scores alone. The matrices' superiority in this respect comes from them structurally restricting the strategies which subjects can follow, while the grids' weakness comes from them allowing more or less free expression of any strategy or combination of strategies. This contradicts the earlier claim that the grids were superior to the matrices for exactly the same reasons. Whether structural restriction of strategies is a strength or a weakness is clearly determined by the nature of the task. If the aim of a study is to determine the relative strength of particular strategies then it is sensible

¹⁸ This has subsequently been done by the author, with promising results.

to use a measure suited to this, i.e. one which structurally contrasts the two strategies of interest, e.g. the appropriate Tajfel matrix. If the task is rather to determine which strategy or strategies people "freely" use in a particular social situation, it is probably preferable to use a measure which allows unfettered expression of as many likely behaviours as possible. It is argued here that the grids are superior to the matrices in this respect. No measure is perfect in all respects and use of one does not of course preclude use of another.

The third potential inferiority of the grid scores relative to the matrix scores is that the former were much less able in the present study to reliably distinguish between subjects pursuing random strategies and those pursuing fairness. This is by far the most serious weakness the grid measures have.

On the grids fairness must co-occur with either joint profit or joint loss, or with an absence of absolute biases in favour of or against either group. Truly random behaviour will (when averaged) co-occur with an absence of mean absolute biases for or against either group. Truly random behaviour can be differentiated from a strategy of maximum fairness plus no absolute bias in favour or against either group by the former's greater IP, OP and F grid score standard deviations. It is also possible to be randomly fair: to choose points at random from along the plane of maximum fairness. This will also result in maximum fairness and no absolute bias toward either group. This can be differentiated from true randomness by its smaller fairness score standard deviation, and from fairness combined with a strategy of no absolute biases by its greater IP and OP score standard deviations.

In the present study the IP, OP and F grid scores of "fair" and "random" subjects were very similar and the standard deviations of those scores were also similar. The evidence is that on average both sets of subjects were making randomly

fair allocations. That is, "fair" subjects were making their fair allocations randomly along the plane of maximum fairness and "random" subjects were making their random allocations fairly, with both showing a preference for the quadrant of maximum profit.

Subjects' accounts of their behaviour corroborate such an interpretation (see Appendix 7). All eight "fair" subjects reported employing randomness (subjects 1, 12, 19, 24, 25 and 29) and/or joint profit (subjects 7 and 17) along with their fairness, and almost half of the "random" subjects expressed a commitment to either fairness (subjects 2, 9 and 11) and/or joint profit (subjects 6 and 26) in addition to "randomness".

The reason that "random" and "fair" subjects could not be differentiated by the discriminant function analysis carried out on the grid scores was therefore that each set of subjects essentially made much the same type of response on the allocation grids, i.e. randomly fair with slight joint profit.

It is not surprising that subjects primarily interested in fairness should make relatively random allocations along the plane of maximum fairness on the allocation grids, nor that these subjects should prefer fairness to in-group favouritism on matrix type 3. Neither is it surprising that "random" subjects should make random allocations on that matrix. What needs to be explained is "random" subjects' fairness on the allocation grids, no matter how random that fairness.

A simple and likely answer is that at least some of the subjects allocating randomly were not very engaged with the task and were therefore choosing allocations which made it quick and easy to work out how many points each recipient would be awarded, i.e. equal allocations (cf. the accounts of subjects 6, 9 and 11 in Appendix 7). Such subjects' "random" allocations would therefore be "random but easy", i.e. "random but fair". Thus "fair" subjects chose randomly along the

plane of fairness because they were interested in being fair but less interested in the form that such fairness took, and random subjects made allocations in exactly the same way because it saved them the time and effort involved in making more truly random choices.

This clearly points to a serious deficiency on the part of the grids in that they make it difficult to distinguish between randomly fair subjects and "lazy" random subjects, and it also elevates the sample fairness mean. Possible remedies are the same as those mentioned above. Making the grids easier to use may make random subjects more willing to be genuinely random, particularly if the grids are computerized and accompanied with a read-out informing subjects how many points each recipient will be awarded if particular allocations are chosen. Alternatively or additionally, the grids can be used in conjunction with other measures which will more reliably discriminate between randomly fair and "lazy" random subjects. This could be Tajfel matrices, or it could be something as simple as asking the subjects how they made their allocations. Considering the main thrust of this chapter and the previous one, it would seem that there is a strong case for obtaining such accounts in minimal and near-minimal group studies anyway.

Comparing the grid measures with other measures of social behaviour

Many of the critics of Tajfel matrices have themselves developed alternative or supplementary measures of social behaviour. These fall into two broad categories.

The first (e.g. Bornstein et al. 1983; Brewer & Silver, 1978) are essentially revised allocation matrices which try and remove the structural confounds inherent in the Tajfel matrices. Whether or not they succeeded in this aim is a moot

point (see Turner, 1983a,b), but in forcing subjects to select one option from a set of allocations which determine both in-group and out-group (positive) awards, the revised measures are as vulnerable as the Tajfel matrix measures to criticisms of lacking content and external validity, reliability, clarity, an appropriate level of measurement, and full and meaningful comparability of scores across studies.

The second type of social behaviour measures give subjects freedom to make any allocations they wish within certain imposed maximums. Mummendey & Schreiber (1983) consider three such measures, only two of which need concern us here. The first, "complementary allocations",¹⁹ involves subjects freely distributing a fixed number of points between in-group and out-group (see also Ng, 1981). The second, "separate allocations", involves subjects allocating as many points as they wish (up to a maximum) to the in-group and also separately allocating as many points as they wish (up to an identical maximum) to the out-group (see also Locksley et al., 1980).²⁰

Each of these measures overcome many of the criticisms levelled at the Tajfel matrices. Additionally, all three could improve their reliability by repeated presentations of the relevant measure, and also improve their content validity by including the possibility of making negative allocations (although this would make the complementary allocation measure slightly more difficult for subjects to comprehend and use

¹⁹ The measures are actually referred to by Mummendey & Schreiber (1983) in terms "assessments" rather than "allocations" as they were interested in intergroup evaluations as opposed to more physical intergroup behaviour. The arguments explored in this chapter apply equally to each, and the terms "assessments" and "allocations" can be used as appropriate to the situation of interest.

²⁰ The third measure, "choice of dimension", involves giving subjects the option of choosing different dimensions to separately and independently award in-group and out-group up to a certain amount of points on.

than at present). Moreover, each measure has certain advantages over the grid measures introduced in this chapter.

Taking the complementary allocation measure first, this involves structural restrictions in as much as joint profit is constant and in-group and out-group receipts are negatively interdependent: giving to one group entails depriving the other of rewards it would otherwise have received. This closely represents many "real-life" situations. People or organizations are often responsible for distributing set amounts of goods (or "bads") between a fixed number of potential recipients in situations in which giving to one recipient entails others going without. That is, many "world" situations involve exactly the same sort of structural restrictions inherent in the complementary allocations measure of social behaviour. When such situations are modelled experimentally, the complementary allocations measure clearly has better external validity than measures which allow independent allocations to each potential recipient.

With minor modifications of multiple presentations and extended ranges to make negative allocations possible, the separate allocations and grid measures are almost identical. The main difference between them is that the separate allocations measure involves subjects allocating first to one group and then to the other, while the grid measure involves subjects making allocations to in-group and out-group simultaneously on intergroup grid presentations and separately on same-group grid presentations.

It might be felt that the separate allocation measure and same-group grid presentations are superior to intergroup grid presentations as the latter make intergroup processes if not inevitable then at least very difficult to avoid. The between-group implications of responses on intergroup grids are clear to (graph-literate) subjects, whereas such implications might never occur to subjects making allocations

to in-group and out-group separately.²¹ Locksley et al. (1980: 776) suggest that the Tajfel matrices "visibly present the possibility of in-group favouritism" but far more worrying is that they and the intergroup grids *probably* force subjects to make intergroup social comparisons they may not make when allocations to in-group and out-group are made separately.

In response, it might be felt that in the minimal group paradigm this argument is probably invalid. This is because even if allocations to in-group and out-group were made separately, the existence of the distinct two groups is so salient, and the allocations to the groups are so temporally and procedurally adjacent, that the allocations must be seen as so close to being simultaneous as to be functionally equivalent to simultaneous allocation.

The minimal group study reported in this study suggests that this may not be correct, however. Many subjects employed fairness within recipient-pairs but not between them. When faced with two recipients on the same grid subjects often made efforts to allocate them an identical number of points. Subjects very rarely attempted to award identical points to members of different recipient-pairs, however, separated as they were by presentations on different allocation grids. That is, inter-individual social comparison and fairness occurred during simultaneous measurement but not during separate measurement.

Whether simultaneous allocation promotes social processes that separate allocation does not is a matter for empirical investigation.²² However, until such research has been

²¹ An analogy is provided by the debate over whether exam scripts are or should be marked in comparison with contemporary scripts or in relation to an "absolute" standard.

²² This would require matched groups making allocations to in-group and out-group others using only same-group or intergroup grid measures.

conducted it would be sensible to adopt whichever means of measurement is most likely to reflect the processes under investigation. In situations which have the potential for social comparison, "built-in" simultaneous evaluation/assessment measures are probably appropriate, such as when a person is asked to evaluate each of two individuals with a view to discovering which is the most suitable for promotion. Similarly, separate allocation/assessment measures are probably most appropriate when investigating situations in which social comparison is less "inevitable", such as when the same two individuals are separately assessed in order to calculate their respective performance-related bonuses.

As well as avoiding possible promotion or exaggeration of particular social comparisons (and of any social processes stemming from such social comparisons) Mummendey & Schreiber's separate allocation measure has a distinct advantage over the grid measures: its ease of use. All subjects have to do when using such a measure is write down any number between the stated minimum and maximum which indicates how many points they wish to allocate to the allocation target of the moment. This is much easier than working out the position on a graph (i.e. grid) which corresponds to the allocations one wishes to make to each of two recipients. Because of this ease of use it is likely that few subjects would encounter problems when asked to use the separate allocation measure. It also seems possible that subjects who wished to respond randomly would be able to do so with as much ease as they could when responding fairly. There would be much less incentive than on the grids for subjects to chose maximally fair allocations simply because they were the easiest to make: any allocation is as easy as any other using the separate allocations procedure. Once the separate allocations had been determined the results could be presented diagrammatically in exactly the same way that grid scores can be. If these steps were taken the separate allocations measure would have all of the advantages of the grid measures while: (i) avoiding their possible

promotion or exaggeration of specific social comparisons; (ii) being easier to use; and, (iii) potentially discriminating more effectively between "randomly fair" and "lazily random" subjects.

Theories of minimal group behaviour

Sample-level in-group favouritism was found in the present study due to almost half of the subjects employing discriminatory dominant strategies. It is sensible therefore to consider how well these subjects' behaviour and their accounts of that behaviour support the various theories of minimal intergroup behaviour considered in the previous chapter.

There was little or no support for equity theory, intergroup accentuation theory or self-categorization theory. No subjects expressed a belief that out-group members would be more discriminatory than in-group members, nor did any claim to be engaging in in-group favouritism to counter such an anticipated bias in order to ensure an equitable outcome. No subject stated that they were trying to behaviourally accentuate intergroup differences simply for the sake of intergroup differentiation. No subject claimed to be trying to ascertain and emulate the prototypical in-group position or stated that they were trying to act as much like in-group members and as little like out-group members as possible. One subject (14) did express a commitment to both in-group superiority and individual superiority of the self from other in-group members, however, and this is consistent with self-categorization theory.

There is weak support from the present study for a demand characteristics explanation of minimal group behaviour. One subject (12) "*didn't think similar reactions every time was the answer*" but found difficulty "*justifying giving anybody a*

negative score". Another (22) "*kept wondering if*" giving more points to the in-group than the out-group "*was the right thing to do, and considered returning to a random selection process but felt that I should be 'loyal' to my own group.*" Against this, the two subjects (24 and 29) who showed some previous knowledge of intergroup discrimination studies made every effort not to conform behaviour they thought might be expected by the experimenter.

The behavioural interaction model also receives very little support from the present study. Not surprisingly, considering subjects were explicitly told that points would not be exchanged for money, no subjects reported being motivated by economic self-interest. More surprisingly, perhaps (see Turner, 1975), is that only one subject (14) seemed concerned about how many points she would personally receive as a result of participation in the study. Also, although many subjects expected participants to favour in-group members relative to out-group members, this subject was the only one who reduced in-group profit because she felt that other in-group members might not reciprocate her initial attempts at instrumental cooperation with them.

Evidence for generic norm theory is mixed. A majority of subjects reported at least considering the possibility of others discriminating in favour of their in-groups, which might be thought indicative of a generic norm of in-group favouritism. However, many of these subjects only explicitly wondered about the possibility that others might discriminate when specifically asked about others' behaviour (as opposed to doing so prior to or concurrent with actually making their own allocations), and so it is likely that few subjects' strategies could actually have been determined by such a norm. A number of subjects made comments consistent with being influenced by a generic norm of "groupness" (especially subjects 5 and 22) or by what they thought everybody else would do (especially subjects 20 and 30), but a large

proportion of these comments are just as consistent with one or more other theories of minimal group behaviour as they are with generic norm theory. A few subjects (e.g. 1, 14, 21, 30, 33) also seemed to be trying to balance competing desires to be discriminatory and fair. No subjects explained their allocations by saying they were attempting to conform to norms.

Evidence for social identity theory is also mixed. On the plus side all but 6 of the subjects (6, 10, 11, 17, 19 and 26) gave responses to the post-task questionnaire which revealed that they were aware both of the social categorization and of their group membership within it. Second, as mentioned above, subjects who discriminated seemed concerned with group rather than with individual outcomes. Third, subjects who discriminated were primarily interested in the relative group goal of obtaining in-group superiority, even if there were two distinct aims once this was achieved (i.e. maximizing that in-group superiority or maximizing in-group profit consistent with in-group superiority). On the minus side, only two subjects (14 and 21, both of whom discriminated) hoped or thought that the outcome of all allocations would be in-group superiority: many more implicitly or explicitly thought that the outcome would be roughly equal between the groups. Also, no subject claimed that using in-group favouritism raised their self-esteem: indeed three subjects (14, 21 and 30) seemed to feel rather worried about their discriminatory behaviour.

As with the study reported in the previous chapter, however, the biggest difficulty faced by all these theoretical accounts of minimal group behaviour is in accounting for the wide variety of strategies adopted by subjects following social categorization. Some subjects seemed to adopt particular strategies such as fairness or in-group preference with barely a thought. Many more seemed simply to be struggling to make some sort of sense of the task they have been asked to do.

Some subjects could not seem to make such sense, and explicitly adopted a random allocation strategy which seemed (and is) as rational as any other in the situation they found themselves in. Others hopped from one strategy to another and back again until they found one that seemed more "right" than the others. It is particularly interesting that several subjects stumbled upon and then adopted group membership-based strategies in this way. Even some of those who immediately adopted a "rational" strategy seemed to do so simply to be able to make their allocations in a consistent and structured way. Some quotes will illustrate the gist of this point:

Chose at random trying to select from all corners of the grid. Then awarded equal marks to both parties. Then one person scored high, the other low. On the actual tables [i.e. matrices] I was more inclined to take account of what group the person was in. On some of the tables I tried to give my own group higher marks. I kept changing my strategy to make it more interesting. Later in the experiment I became more conscious of the fact I was in a group and so at the end this then became a consideration. (Subject 2)

Allocated equal points on all occasions - no basis for choosing to give one person points and not the other. No reason to give or take away points, so always gave points and where possible of equal value for each task. No [did not change strategy]. I think much about it [others' strategies], although I did wonder if anyone chose to allocate points on the basis of group membership - ie, give to own group, take away from other group. Not knowing who is in which group removes any reason to decide point allocation in this way, and there is no motivation to score your own group highly. (Subject 7)

At first I made totally random choices and then I started to favour my own group and gave them a higher score (usually positive) than the other group. I kept wondering if this was the right thing to do, and considered returning to a random selection process but felt that I should be "loyal" to my own group. (Subject 22)

I found it difficult to allocate points at first when I had no reason or basis for giving or taking away. For that reason I began by giving equal points. Then I changed my strategy and decided to allocate more to people in my group (X) than to those in the other group. This was simply to make it more interesting for myself as

well as trying to give the most points to my own group whether that meant winning or not. (Subject 28)

Despite the fact that all subjects ended-up behaving in ways which could be categorized into a limited number of "dominant" strategies, the "full" strategies subjects adopted were almost unique to each individual. As Tajfel noted in the same year that the classic exposition of social identity theory (i.e. Tajfel & Turner, 1979) was published:

The construction of a social system in terms of sharply drawn social categories and the capacity to categorize and compare oneself with others in certain ways and for certain purposes are the necessary conditions for the appearance of certain forms of intergroup behaviour; they are not sufficient conditions. (Tajfel, 1979, cited in Tajfel, 1981: 48, original emphasis)

If the aim of truly social-psychological intergroup theories is to discover uniformities of intergroup behaviour, as Tajfel repeatedly claimed, then the minimal group paradigm does not seem to be the place where the fundamental principles of such theories are to be discovered. Merely finding that one shares a social category with others does not seem to be sufficient to make one behave in the same way as they do, even toward members of a specific out-group.

CONCLUSIONS

As in the last chapter the main conclusion must be that mere social categorization is not sufficient to promote individuals to be discriminatory in favour of in-groups. Significant in-group favouritism was found at sample-level, however, and this is not to be dismissed. Outside the laboratory the knowledge that one's group is (on average) being discriminated against by members of another group is likely to make one prejudiced against that out-group in turn, even if no member of the other

group discriminates against one personally, either as an individual or as a member of one's in-group. In the next chapter this distinction between individual and group levels of discrimination will be explored further.

The second main conclusion, again as in the previous chapter, is that minimally socially categorized subjects engage in a bewildering array of social behaviours, far too many for current theories of minimal group behaviour to adequately account for. An adequate theory of minimal group behaviour will need to explain both why so many allocation strategies are adopted and why these strategies can be clustered into a handful of "dominant" strategies. More importantly, the theory will need to explain and predict the emergence of more widespread and consensual intergroup behaviour, particularly intergroup discrimination of various kinds. Again this is addressed in the following chapter.

The third main conclusion is that at least two types of discrimination seem to occur within the minimal group paradigm. In line with social identity theory both seem to require that in-group superiority be attained but beyond that one form requires that the in-group superiority is maximized by maximizing both in-group enhancement and out-group derogation, whereas the other requires in-group enhancement without either out-group enhancement or significant out-group derogation. The first of these forms of discrimination links very closely to the notion of maximum difference in favour of the in-group and the second bears strong similarities to the notion of maximum in-group profit, both of which are often conflated both methodologically and conceptually (e.g. as "FAV").²³ A theory such as social identity theory, which aims to explain intergroup discrimination stemming from intergroup

²³ The second form of discrimination is not identical to maximum in-group profit, as there is a caveat to the strategy - subjects pursuing this form of discrimination desire maximum in-group profit as long as and once in-group superiority has been achieved.

comparison, really needs to distinguish between and explain the difference between these two manifestations of the search for positive social identity and self-esteem, and also needs to predict who will adopt which, and when.

Finally, it has been argued that the allocation grids employed in this study have a number of advantages over the Tajfel matrices. The two which need reiterating here are (i) the ability to differentiate between the two aforementioned "types" of relative bias in favour of the in-group, and (ii) the ease and confidence with which grid scores can be compared across studies. The latter of these is important because meaningful comparison of pull- and indirect in-group favouritism scores are so hazardous as to be extremely rare. Most reviews of matrix use restrict themselves to mentions of whether or not particular pull scores were significantly different from zero or, at very most, acknowledgement of which pull score was the greatest of those measured. This is not surprising considering the criticisms which have been levelled at the matrix measures, but it is a tremendous waste of research effort. This is not to say that the grid measures are universally superior to the matrix measures, that the matrices do not have their uses, or that the allocation grids provide the best possible measure of social behaviour. Any such claim would be manifestly false. Mummendey & Schreiber's (1983) "separate allocation/assessment" measure of social behaviour seems in particular to share all of the allocation grid measures' advantages over the Tajfel matrix measures and to additionally have a number of advantages over the allocation grid measures.²⁴ Which measure is most appropriate will, as always, be best determined by the purpose the measure is to serve.

²⁴ Particularly when amended to: (i) include repeated presentations of the measure; (ii) include the possibility of negative as well as positive gains; and, (iii) present the outcomes diagrammatically.

CHAPTER 4: STUDY 3 - DEVELOPING NORMS OF INTERGROUP BEHAVIOUR**CHAPTER OVERVIEW**

This chapter reports a study with three within-subject conditions. The first condition was similar to a standard minimal group paradigm study (*cf.* Tajfel et al., 1971). The second condition involved sub-groups of three subjects drawn from the same social categories being tape-recorded whilst collectively making a second set of allocations to in-group and out-group others. The third condition involved subjects individually making a third set of allocations to in-group and out-group others (but not to those in the same sub-groups as themselves).

In the first condition subjects employed both mean sample-level in-group favouritism and mean fairness across the groups. Mean sample-level in-group favouritism polarized from condition 1 to conditions 2 and 3. Mean sample-level fairness across the groups depolarized (to zero) from condition 1 to conditions 2 and 3.

Despite these gross sample-level results being consistent with an interpretation derived from social identity theory, it is argued that a normative interpretation is preferable.

In all three conditions a wide variety of strategies were employed, both across individuals and across sub-groups. Explanations in terms of social categorization per se are not sufficient to explain this variety of behaviour. However, an interpretation which considers the development or non-development of sub-group norms in condition 2, and also sub-group members' commitment to any norms developed, can be used to explain all condition 2 and condition 3 (individual, sub-group and sample) behaviour.

INTRODUCTION

Within the minimal group paradigm (e.g. Tajfel et al., 1971) subjects are arbitrarily or randomly allocated into one or the other of two novel social categories (i.e. "minimal groups") and asked to make decisions about how to behave toward and/or evaluate anonymous members of each group. The paradigm has a number of strict procedural criteria, the first of which is that there "should be no face-to-face interaction" between any of the subjects (Tajfel et al., 1971: 154), either within or across groups.

The results of the "standard" minimal group paradigm are well documented. On average minimal group subjects display in-group favouritism, both by allocating more "goods" (material or evaluative) to the in-group than to the out-group and by attempting to obtain in-group superiority over the out-group. What has yet to be explicitly investigated is what happens when the first procedural criterion of the minimal group paradigm is violated. Specifically, what happens when within-group face-to-face interaction is permitted? There are at least two reasons to think that this is a question of some importance.

The first is theoretical. In the "standard" form it is difficult to see how the minimal group paradigm could produce anything other than an individualistic account of "group" and "intergroup" phenomena (see also Chapter 1, pp. 36-41). As subjects are forced to make their decisions individually - with no prior or contemporaneous intragroup communication - the resultant "group" and "intergroup" behaviour cannot be anything other than the mean of *individual* behaviour. At best it can be claimed that all subjects react to the same stimulus (i.e. group membership within a multi-group structure) with more or less identical responses (i.e. in-group preference). Indeed, this is the so-called "central tenet" of social

identity (see Chapter 1). Social identity theory, in other words, offers an individualistic theory of group and intergroup behaviour: it claims that people will individually respond in predictable ways to the "social" stimulus of group membership in a multi-group situation.

The second reason why it is sensible to ask what happens when intragroup communication is allowed in the minimal group paradigm concerns the generalizability, or lack of it, of the results of experiments within that paradigm. Even if it is accepted that the minimal group paradigm reveals the intergroup behavioural consequences of "mere" social categorization, this tells us little or nothing about the intergroup behavioural consequences of "non-mere" social categorization. Most if not all non-experimental group and intergroup behaviour follows or occurs simultaneously with intragroup communication. It therefore behoves social psychologists to examine the consequences of social categorization plus intragroup communication.

These points can be illustrated with reference to Sherif's (1936) studies of group norms. Sherif's studies utilized the phenomenon known as the autokinetic effect. This is where a stationary point of light in a darkened room appears to move. There are considerable differences in how far subjects individually estimate the light moves, but all or most subjects unfamiliar with the autokinetic effect think that it does. The first stage in Sherif's studies was to establish that repeated trials led to individuals developing their own norms. That is, subjects who were asked to estimate the light's movement several times tended to base their later estimates on their own earlier ones. Put another way, all or most subjects responded to the stimulus (i.e. the stationary light) in similar ways (i.e. by estimating that it did move), but there were individual differences, both in subjects' initial responses and in the personal norms they developed over repeated trials.

The second stage of Sherif's studies involved two conditions. In the first, subjects made one estimation alone and were then repeatedly asked to estimate the light's movement in the presence of others (who had also made one solo estimation). The initial solo estimations showed the usual individual variation but on subsequent estimations made in the presence of others the estimations converged across individuals. That is, rather than subsequent estimations converging on a personal norm (i.e. on subjects' own earlier estimations), those estimations converged on a social or group norm. Importantly, the group norm that developed was not always identical to the mean of the initial individual estimations.¹

In the other second stage condition subjects were first asked to make a series of estimations in the presence of other estimating subjects (before any had made solo estimations) and were then asked to a solo estimation. In this condition initial estimations showed very little variation across subjects: subjects made very similar estimations to each other. Furthermore, these subjects' subsequent solo estimations were also very similar to the group norm that had been established and were, therefore, also very similar to each others'.

Sherif's studies of group norms reveal the individualistic perspective endemic in both the minimal group paradigm and within social identity theory. The minimal group paradigm is analogous to the situation in which subjects make an initial individual estimation of the light's movement. In the Sherif studies subjects responded in more or less the same way to an identical stimulus (i.e. the light: subjects estimated that it moved), but there were considerable individual differences (i.e. the estimations of movement varied across individuals). Similarly, in the minimal group paradigm it is claimed that

¹ Nor was it uniformly more extreme on the same side of zero than that mean, as might be expected from the group polarization phenomenon (see below).

subjects respond in more or less the same way to an identical stimulus (i.e. group membership within a multi-group situation: subjects show in-group preference), but there are considerable individual differences (i.e. the extent of in-group preference varies across individuals).² But whereas Sherif was clear that this was a (relatively) non-social, i.e. individualistic, phenomenon, social identity theory claims that it is a social one. In fact, all that is "social" is the stimulus (in the sense that subjects share category membership). Further, where Sherif thought of his solo situation as a starting point, social identity theorists have tended to treat their's as the end product.

The same points can be illustrated with reference to the group polarization paradigm (see Isenberg, 1986 and Turner, 1991 for reviews). One variant of this paradigm has three repeated-measures conditions: an individual one followed by a group one and then a second individual one. In each condition subjects are asked to respond to some stimulus, e.g. to express their attitudes on some contentious issue. Group polarization is the term used to describe the robust phenomenon that mean responses by or within a group (i.e. in condition 2) are more extreme on the same side of zero or neutral as the mean of the individual responses in condition 1. The addition of condition 3 demonstrates that the mean of individual responses in this third condition are less than or equal to those shown in condition 2, but are reliably still more extreme on the same side of zero as the mean of the individual responses in condition 1.

The group polarization phenomenon again suggests that intragroup interaction can be predicted to result in behaviour which is different from behaviour which results when

² In fact the studies reported in the previous two chapters reveal that the analogy fails to a certain extent at this point, as subjects do not respond in "more or less the same way" to group membership in an intergroup context: at least some subjects show no in-group preference.

intragroup interaction is not permitted. As with the Sherif studies, the minimal group paradigm is analogous to the first individual condition of the group polarization paradigm. Subjects are merely examined as to how they individually respond to a stimulus (i.e. group membership within a multi-group structure) and the mean of their individual responses is called "group" (and/or "intergroup") behaviour.

It might be objected that the comparisons between the minimal group paradigm and both Sherif's and the polarization paradigm are illegitimate. After all, the subjects in the latter paradigms are not explicitly members of a group whereas they are in the minimal group paradigm. There are two counters to such an objection. The first is that social identity theory has been developed in such a way as to claim that subjects in both the Sherif and the polarization paradigms are members of a group and in fact perceive themselves as such, i.e. they are a group of people confronted with a common stimulus or situation and can be expected to respond in similar ways (e.g. Abrams et al., 1990; Mackie, 1986; McGarthy et al., 1992). The second response is simply to make the as yet unsubstantiated claim that if subjects in the first conditions of the Sherif and the polarization paradigms were explicitly told that they were members of a group the differences across conditions typically found in those paradigms would still be obtained.

The thoughts expressed above provided the motivation for the experiment reported in this chapter. The experiment is essentially a combination of the minimal group and the group polarization paradigms and, as such, might be termed the minimal intergroup polarization paradigm. Condition 1 more or less followed the standard minimal group paradigm format: each subject was categorized into one or the other of two novel social categories according to explicitly random criteria and were then asked to make a series of point allocations to pairs of other subjects identified only by their unique code numbers

(known only to themselves) and by their minimal group memberships. In condition 2 the same subjects met with one or two other subjects from the same minimal in-group and were asked, as a "sub-group", to collectively decide how to make a second set of point allocations to pairs of other in-group and out-group members. Finally, in condition 3, the same subjects were asked to individually make a third set of point allocations to in-group and out-group others (not including those in their own condition 2 sub-group). The hypotheses, derived from the minimal group, Sherif and polarization paradigms, were as follows.

From the minimal group paradigm it can be predicted that:-

- H1. Sample-level in-group favouritism will be significantly greater than zero in condition 1.
- H2. Sample-level fairness across groups will be significantly greater than zero in condition 1.

From Sherif's paradigm it can be predicted that:-

- H3. Intergroup behaviour will show significantly greater variation within sub-groups in condition 1 than in condition 3.

From the group polarization paradigm it can be predicted that:-

- H4. Within sub-groups, mean in-group favouritism will be significantly more extreme in condition 3 than in condition 1, on the same side of zero.
- H5. Within sub-groups, mean fairness across groups will be significantly more extreme in condition 3 than in condition 1, on the same side of zero.

Combining hypotheses 1 and 4 suggests that:-

- H6. Sample-level in-group favouritism will be significantly greater in condition 3 than in condition 1.

Combining hypotheses 2 and 5 suggests that:-

H7. Sample-level fairness across groups will be significantly greater in condition 3 than in condition 1.

METHOD³

Subjects

24 British second-year psychology undergraduates from the University of Keele volunteered to take place in a "study of decision-making." One did not turn up. Of those who did, 17 were female and 6 were male. Ages ranged from 19 to 47 years, with a mean of 26.65 (median = 22, mode = 21).

Procedure

Over the space of one week subjects were seen in eight experimental sessions of 3 subjects each, except for one session in which only two subjects attended. After initial greetings subjects were told that they were being seen three at a time simply "for the sake of efficiency and [that] more would be seen at once if more rooms were available." In all but the first session subjects were asked if they had heard anything about the study from previous participants. All denied that they had. Subjects were then told that they would be taking part in three tasks and would be told the nature of each task immediately prior to doing it.

The first task was a "standard" minimal group paradigm (cf. Tajfel et al., 1971) with the following minor amendments: (i) social categorization was ostensibly explicitly random as

³ Farsides (1993b) was based on selected results from this study.

opposed to being based on arbitrary similarities of taste or skill (cf. Billig & Tajfel, 1973), but was in fact rigged so that all same-session subjects were in the same minimal category; (ii) in addition to (knowingly) not being able to allocate directly to the self, subjects also (unknowingly) could not allocate to others in the same session; (iii) subjects were told that they would be informed later of the allocation decisions, but were not explicitly told that they would be informed of how many points others had allocated to them personally; (iv) subjects were told that the points were "just points" and would not be changed for money or for anything else at a later time; and (v) subjects knew that they would be taking part in two further unspecified tasks upon completion of the first.

In the first condition the experimenter explained that all of the subjects across all of the sessions would be allocated at random into either "Group K" or "Group J", by the experimenter picking one envelope per subject from a large pile of such envelopes which any or all of the subjects in each session could "shuffle" if they so wished. Each envelope contained categorization information, as well a code number unique and known only to each subject, and a response booklet for the first task. In fact, the pile of available envelopes was different in each session so that all same-session subjects were allocated to the same category. Subjects were also told that all members of Group K would have code numbers between 1 and 49 and that all members of Group J would have code numbers between 51 and 99. Categorization was ostensibly random to ensure that any allocation behavioural effects obtained from the social categorization were not confounded by perceptions of within-category similarities and/or between-category differences in addition to those entailed by the social categorization itself (cf. Billig & Tajfel, 1973; Farsides, 1993a). Allocation to the same category within sessions was necessary so that same-group communication was possible in the second condition (see below). Subjects were prevented from

allocating to others within the same session to ensure that in subsequent conditions subjects could not feel that they had individually favoured or discriminated against their fellow subjects or that their fellow subjects had individually favoured or discriminated against them. Subjects were not told that they would be informed about the allocations that had been made to them personally by others in an attempt to avoid biasing subjects toward an individual-level or a group-level focus in their allocation behaviour. Subjects were told that the points were "just points" (i) in order to minimize possible strategies of economic self-interest, (ii) because funds available to the experimenter were limited, and (iii) because Turner (1975) found that the effects of social categorization were as strong if not stronger when points were allocated as when money was.

Subjects were told that the nature of the first task was to allocate points to pairs of other people who would be identified only by their group membership and by their unique code number. This would be done using "Tajfel matrices" (see Appendix 1) which were explained and demonstrated to the subjects by the experimenter. It was stressed that there was no way of knowing whether or not they would be allocating to others in the same session as this would depend on the "random" categorization. In fact subjects never allocated to others in the same session (see above). When subjects were satisfied about the nature of their task the experimenter "randomly" selected each of them an envelope (see above) which they were told not to open until they had been shown to separate rooms. Once seated in those rooms they were told to open the envelopes to discover which of the groups they had been allocated to and what their unique code number was. Instructions how to use the matrices, with examples, were repeated at the front of the response booklets. The subjects were instructed to report back to the experimenter when they had finished this first task.

When all the subjects within a session reported back to the experimenter the second condition was initiated. Subjects were led to a previously unused room where they were seated around a table. On the table was a tape-recorder and a single response booklet. The subjects were told that the second task involved them *collectively* making a second set of allocation decisions to pairs of other people identified as before. It was stressed that no-one in the room would be a recipient of any of the allocations made by the sub-group during this second task. Subjects were asked if they minded their decision-making discussions being tape-recorded as the experimenter "was particularly interested in how group decisions were made and would like to be able to listen to the tape later." No-one refused so the experimenter turned on the tape-recorder and left the room after requesting that the subjects should report back to him when this second task was complete.

Upon completion of the second task subjects were led back to their individual rooms and were asked to individually make a third set of allocations decisions. This time (as compared to the first condition) subjects were explicitly told that others within their session would not be the recipients of any of their allocations, and neither would they themselves.

When all of the subjects had completed the third stage and reported back to the experimenter they were again asked if they had heard anything about the task from previous participants. Again none admitted to having had any pre-knowledge of the experimental procedure. Subjects were then asked to promise not to discuss the procedure with anyone yet to take part. Finally, subjects were told when and how they could find out about the experiment and its results, were thanked, and were given the opportunity to make any comments about their experience within the session.

To recap: there were three conditions in the present study. The first simply followed the standard minimal group paradigm. Subjects were randomly socially categorized and asked to individually allocate points to in-group and to out-group members. In the second condition subjects were placed in a room with one or two other in-group members and each "sub-group" of subjects were asked to collectively make a second set of allocations to in-group and out-group others. In the third condition subjects were separated again and asked to individually make a final set of allocations to in-group and out-group others (excluding other sub-group members).

Materials

Eighteen "Tajfel matrices" were included in each response booklet; six of each of the three types shown in Appendix 1. For each type of matrix there were three recipient patterns: two in-group members (2 matrices), two out-group members (2 matrices) and one in-group and one out-group recipient (2 matrices, one with the in-group member as the top-row recipient and one with the out-group member as the top-row recipient). In the first condition each subject received their matrices in a different random order from each of the other subjects in their session. The booklet in the second condition contained matrices in a fourth random order. In the third condition each subject's matrices were in a random order other than the one used in their own previous individual condition. Also in the third condition the response booklet finished with a short questionnaire asking subjects their age, sex and nationality, followed by a page on which subjects were invited to make comments about the experiment they had just participated in.

RESULTS

Sample pull-scores

Mean sample pull-scores were calculated within each condition and are shown in the table below. In condition 1 - the "standard" minimal group paradigm condition - the three in-group preference pull-scores (i.e. FAV on MJP, MD on MIP+MJP and FAV on F) were all significantly above zero by one-tailed one-sample t-tests, as was the pull of F on FAV ($p < 0.005$ in all cases). Thus hypotheses H1 and H2 were supported. Neither of the other pull-scores (i.e. MIP+MJP on MD and MJP on FAV) were significantly different from zero ($p > 0.05$ in each case).

	Condition 1	Condition 2	Condition 3
FAV on MJP	3.739 (5.57)	6.522 (5.88)	6.043 (7.22)
MD on MIP+MJP	2.609 (5.20)	5.304 (6.10)	3.739 (6.59)
FAV on F	4.609 (5.53)	9.652 (3.04)	7.130 (7.13)
MJP on FAV	-0.435 (5.04)	-0.348 (1.15)	-0.043 (1.94)
MIP+MJP on MD	0.696 (5.09)	3.304 (4.19)	3.739 (4.69)
F on FAV	3.391 (5.38)	-1.565 (2.43)	0.783 (4.78)

Table 4.1: Mean pull-scores by condition

A MANOVA with the six pull-scores and the three conditions as within-subjects factors revealed a significant multivariate main effect of strategy ($p < 0.001$). Averaged across the three conditions the three in-group preference pull-scores were strongest (FAV on F = 7.130; FAV on MJP = 5.435; MD on MIP+MJP = 3.884), followed by the pull of MIP+MJP on MD (2.580), with the other pull-scores being non-significantly different to zero (F on FAV = 0.870; MJP on FAV = -0.275). There was also a significant main-effect of condition ($F_{(2,44)} = 3.93, p < 0.05$). However, averaging the six pull-scores is

a meaningless exercise and so this result is of no concern. More interestingly, there was a significant effect for the interaction of strategy x condition ($F_{(10,220)} = 5.04, p < 0.001$). Inspection of the means in the table above suggest that investigation of this interaction can be simplified in two ways. First, as the pull of MJP on FAV was consistently non-significantly different to zero in each of the three conditions, this score can be dropped from the analysis. Second, as the three in-group preference pull-scores are conceptually similar and also have similar patterns across the three conditions, a composite "in-group preference" (InPref) pull-score can be obtained simply by taking the means of those three pull-scores.

A second MANOVA with the three pull-scores of interest and the three conditions as within-subjects factors revealed a main effect of strategy ($F_{(2,44)} = 8.56, p < 0.005$), with the mean InPref score across conditions being greatest at 5.483 (with means of MIP+MJP on MD and F on FAV being 2.580 and 0.870, respectively, as reported above). The main effect of condition was insignificant for this MANOVA. The interaction, however, remained significant ($F_{(4,88)} = 10.16, p < 0.001$). The means of interest are shown in the table below.

	Condition 1	Condition 2	Condition 3
InPref	3.652 (4.59)	7.159 (4.39)	5.638 (6.01)
F on FAV	3.391 (5.53)	-1.565 (3.04)	0.783 (7.13)
MIP+MJP on MD	0.696 (5.09)	3.304 (4.19)	3.739 (4.69)

Table 4.2: Mean F on FAV and MIP+MJP on MD pull-scores, and mean InPref scores, by condition

Inspection of the mean pull-scores within each condition shows that InPref was considerably greater in condition 2 than in condition 1, with the condition 3 mean falling between those

two values. The condition 3 InPref mean was significantly greater than the condition 1 mean by a one-tailed related t-test ($t_{(22)} = -2.43, p = 0.012$), thus supporting H6: sample-level mean in-group favouritism was significantly greater in condition 3 than in condition 1. For F on FAV the condition 2 mean was considerably less than the condition 1 mean, and indeed was significantly less than zero. Here again the condition 3 mean (which was not significantly different to zero) fell between the means in the previous two conditions. The condition 3 F on FAV mean was significantly different from its condition 1 mean by a one-tailed related t-test ($T_{(22)} = 2.28, p = 0.033$), but the condition 3 mean was less than the condition 1 mean, thus failing to support hypothesis H7: sample-level mean fairness across groups was not significantly greater in condition 3 than in condition 1. MIP+MJP on MD followed a third pattern, with the condition 2 mean being higher than the condition 1 mean, while the condition 3 mean remained at its condition 2 level. The condition 3 mean of MIP+MJP on MD was significantly greater than its condition 1 mean, as indicated by a two-tailed related t-test ($t_{(22)} = -2.99, p = 0.014$).

Pull-score polarization within sub-groups

Hypotheses 4 and 5 predict that intergroup strategies will polarize within sub-groups. The table below allows these hypotheses to be examined. The numbers contained within the cells of the main body of the table represent the absolute differences between the condition 3 and condition 1 sub-group means for each strategy. The sign of each of these numbers represents whether or not the condition 3 means were more extreme than but on the same side of zero as the condition 1 means, i.e. whether polarization occurred for each strategy within each sub-group. Positive values indicate that polarization did occur and a negative ones that "depolarization" occurred (i.e. that the condition 3 mean was

less extreme than or on the other side of zero to the condition 1 mean). Row means therefore represent the mean polarization between conditions 1 and 3 for each sub-group, averaged across all six pull-scores, and column means represent the mean polarization between conditions 1 and 3 for each pull-score, averaged across all eight sub-groups. The overall table mean, shown in the lower-right cell of the table, represents the mean overall within sub-group polarization of intergroup behaviour between conditions 1 and 3, averaged across all eight sub-groups and all six pull-scores.

	FAV on MJP	MD on MIP+MJP	FAV on F	MJP on FAV	MIP+MJP on FAV	F on FAV	X
1	-2.00	-3.66	1.66	-2.66	1.00	-1.00	-1.11
2	2.67	-1.00	3.66	-0.66	-0.33	-0.33	0.67
3	5.00	4.00	5.33	-0.33	0.00	-2.67	1.89
4	-4.00	-1.34	4.00	-0.66	-5.33	-8.00	-2.56
5	0.67	5.33	-1.66	0.00	-7.33	2.34	-0.11
6	0.00	-1.00	-7.00	-1.00	2.00	3.00	-0.67
7	6.67	-1.00	5.00	-4.67	4.00	-5.00	0.83
8	4.67	-1.33	6.00	0.67	4.67	-4.00	1.78
X	1.71	0.00	2.12	-1.16	-0.17	-1.96	0.09

Table 4.3: Polarization by strategy and sub-group

Averaging across the eight sub-groups and the six pull-scores, no significant polarization (or depolarization) of intergroup behaviour occurred between conditions 1 and 3 of the present study. The condition 3 mean of intergroup behaviour (i.e. averaged across all six pull-scores) was less than a tenth of a scale-point more extreme than and in the same direction as the condition 1 mean, when averaged across the eight sub-groups.

The column totals suggest that there were differences across intergroup strategies, with the pull-scores of FAV on MJP and FAV on F showing polarization (averaged across the eight sub-groups) of 1.71 and 2.12 scale-points, respectively, and the corresponding "reverse" pulls of MJP on FAV and F on FAV showing depolarization (averaged over the sub-groups) of 1.16 and 1.96 scale-points, respectively. The mean pull-scores of MD on MIP+MJP and MIP+MJP on MD showed little difference across conditions 1 and 3 (averaged across the sub-groups). However, none of the pull-scores (averaged over sub-groups) indicated significant polarization or polarization between conditions 1 and 3.⁴ No support is found for hypotheses H4 or H5, therefore, as neither strategies of in-group favouritism nor strategies of fairness across groups polarized within sub-groups between conditions 1 and 3.

Similarly, although the row means in the table above suggest differences in intergroup behaviour (i.e. averaged across the six pull-scores) across sub-groups, with, for example, sub-groups 3 and 8 showing polarization and sub-group 4 showing depolarization, none of the sub-group's polarization (or depolarization) was significantly different from zero change across conditions 1 and 3.

Before leaving the table above, however, it is worth noting that there is strong suggestive evidence that polarization and depolarization did occur on particular strategies for individual sub-groups. Thus, for example, sub-groups 3, 7 and 8 showed considerable polarization on the strategies of FAV on MJP and FAV on F, while sub-groups 4, 7 and 8 showed substantial depolarization on the strategy of F on FAV.

⁴ By one-sample t-tests with $n = 8$ (i.e. the sub-groups) comparing the extent of polarization with zero difference between condition 1 and condition 3 means.

Variance within sub-groups across conditions 1 and 3

Hypothesis H3 predicts that subjects' intergroup behaviour will show greater variation around sub-group means in condition 1 than in condition 3. The table below shows the difference between the sub-group variances in conditions 1 and 3, with positive numbers indicating less variance in condition 3 than in condition 1 and negative numbers indicating greater variance in condition 3 than in condition 1.

	FAV on MJP	MD on MIP+MJP	FAV on F	MJP on FAV	MIP+MJP on FAV	F on FAV	X
1	3.464	1.732	-2.887	0.000	-1.732	-1.732	-0.193
2	-2.506	-0.435	-1.425	1.939	-0.143	3.204	0.106
3	6.083	3.141	5.508	0.577	0.000	7.371	3.780
4	-6.553	-4.707	-0.956	6.351	4.483	-1.732	-0.519
5	0.309	2.887	-4.989	1.000	2.044	-3.136	-0.314
6	-2.828	-9.900	-1.414	1.414	-1.414	-1.414	-2.593
7	1.155	-1.782	5.533	3.691	-4.423	5.533	1.618
8	2.646	3.215	6.557	6.602	4.509	6.083	4.935
X	0.221	-0.731	0.741	2.697	0.415	1.772	0.853

Table 4.4: Differences in sub-group behavioural variation in conditions 1 and 3 by strategy and sub-group

The overall table mean indicates that (averaged across the eight sub-groups and the six pull-scores) variance around sub-group means was only slightly (and non-significantly) less in condition 3 than in condition 1. Overall, then, no support is provided for hypothesis H3: variance of intergroup behaviour within sub-groups was not significantly greater in condition 1 than in condition 3. Column totals suggest, however, that variance reduction from condition 1 to condition 3 was greater for some intergroup strategies than for others, particularly for the strategies of MJP on FAV and F on FAV (averaged over

sub-groups). One-sample t-tests with sub-groups as the unit of analysis revealed significant within sub-group variance reduction from condition 1 to condition 3 only for MJP on FAV ($t_{(7)} = 2.96, p < 0.05$).

Similarly, row totals suggest that variance reduction across the conditions of interest was greater for some sub-groups (i.e. sub-groups 3 and 8) than for others, while at least one sub-group (sub-group 6) seemed to show more within sub-group variance in condition 3 than in condition 1 (averaged across the six pull-scores in each case). One-tailed one-sample t-tests with the six pull-scores as the unit of analysis revealed that sub-groups 3 ($t_{(5)} = 3.047, p < 0.05$) and 8 ($t_{(5)} = 6.954, p < 0.001$) each showed significant decreases in within sub-group variance from condition 1 to condition 3, while the changes in variance across these two conditions were insignificantly different from zero for all other sub-groups.

Qualitative data

Sub-group discussions in condition 2 of the present experiment were transcribed in full (see Appendix 8). In this subsection some attempt will be made to précis or summarize the decision-making processes that occurred within each sub-group.

Sub-group 1 began condition 2 with each sub-group member explaining how they had made their allocations in condition 1. Two subjects reported trying to give equal points to each recipient on each allocation (despite at least one of those subjects considering the possibility, advisability and/or the expectation of in-group preference) while the third subject reported adopting a random strategy, as "it didn't seem to make any difference who you allocate what to anyway...because it doesn't apply to any of us and it doesn't involve any money anyway". A suggestion from the random allocator to make their "group" decisions by turn-taking was accepted fairly quickly

and all subsequent allocations followed the pattern of two maximally fair decisions followed by one random or arbitrary one. Thus, no attempt was made to reach a consensual "general" allocation strategy and the only "sub-group" norm which emerged was one of turn-taking.

This interpretation is fully consistent with the available quantitative data (see especially Appendix 9, in which tables and graphs are provided of each subjects' allocations in each of the three conditions, organized by pull-scores and sub-group memberships). In conditions 1 and 3 two subjects (66 and 81) made maximally fair intergroup allocations and the third (40) made random intergroup allocations. In condition 2 intergroup allocations were maximally fair for FAV on MJP, MD on MIP+MJP, MJP on FAV and MIP+MJP on MD, but were not maximally fair for F on FAV or FAV on F. This was clearly because on intergroup allocations involving matrix types 1 and 2 the "fair" subjects made the "group" choices, while the "random" subject made the two matrix type 3 "group" choices (see Appendix 8). In condition 3 the two "fair" subjects (66 and 81) again made maximally fair intergroup allocations throughout, while the third subject's (40) condition 3 allocations are consistent with allocating randomly.

Sub-group 2 began condition 2 with each of the subjects making suggestions for particular allocations. The male subject flitted between various strategies, such as fairness ("*Even Stevens?*"), randomness ("*Haven't got a reason for it*") and in-group preference ("*I feel a close affinity to members of K*"). He was also interested in the rationale for decisions that others made or suggested. After he asks "*Why are we actually doing it like this?*" it seems as though the tape-recorder was turned off, to be turned on again to general laughter. The subsequent decision-making suggests that some attempt was made during this hiatus to reach a consensus about how to make subsequent allocations, i.e. to develop one or more operational sub-group norms. These seem to have been (i) to

maximize in-group profit and/or in-group superiority (it is not clear what the norm was for deciding between these two options when they clashed, i.e. on matrix type 2) on intergroup allocations, (ii) to give equal points on allocations with two in-group recipients, and (iii) to make allocations involving two out-group recipients arbitrarily. Trouble arises when one of the females violates the main (i.e. intergroup) norm by making a "fair" allocation. The male asks for an explanation: "*Hang on. Why didn't you just give them all to K...How about keeping in the true spirit of...fellow group members...*". Perhaps she did not want to comply with the suggested sub-group norm(s) or perhaps she was trying to cover up for mistakenly violating it (or them), but the female concerned justifies her decision by claiming an affinity for the number 13. The main sub-group norm (if it indeed exists) is violated again later, again by a female making a seemingly arbitrary choice, and the male again objects ("*I'm not being over-ruled*"). If sub-group norms were decided it seems that at least one female member of the sub-group was not committed to them by the end of condition 2, but was instead responding randomly and/or arbitrarily.

The allocations made by these sub-group members across the three conditions are again consistent with such an interpretation (see Appendix 9). Two subjects' (08 and 34) choices were much more similar to each other's in condition 3 than in condition 1 on matrix type 1 and 3 intergroup allocations. That is, these subjects responded differently to each other in condition 1 but by condition 3 were displaying maximal preference for FAV over both MJP and F (and correspondingly were showing no preference for MJP or F over FAV). The third subject (24) did not conform to these strategies in condition 3. Indeed, an allocation made by this "random" subject was responsible for the condition 2 FAV on F score not reflecting the "sub-group" norm adhered to by the other two (see Appendices 8 and 9). This sub-group's allocations over the three conditions also suggests that no

sub-group norm at all was developed for intergroup allocations on matrix type 2, where maximum in-group (and joint) profit is opposed by maximum in-group superiority, i.e. each subject returned to whatever strategy they had individually employed in condition 1.

Sub-group 3 began condition 2 by explaining their condition 1 individual strategies in some detail. The male reported trying to give each member of each recipient-pair equal points (*"you don't really know any of them so, why not, especially if it was for money...when you hope they'd do the same for you"*). Both females reported trying to obtain in-group superiority on intergroup decisions and each chose maximum joint (i.e. in-group) profit on allocations involving two in-group members, but whereas one chose "fairness" on allocations involving two out-group members the other adopted a strategy of minimum out-group profit. A set of sub-group norms quickly develops: (i) maximize in-group superiority on intergroup allocations wherever possible, (ii) on intergroup allocations where this is not possible (i.e. on o/i presentations of matrix type 3), opt for "fairness" (where "fairness" also maximizes in-group profit and minimizes out-group profit, and additionally avoids in-group inferiority), (iii) minimize joint (i.e. out-group) profit on allocations involving two out-group members, with random or arbitrary allocation permissible where this is not possible (i.e. on matrix type 3), (iv) maximize joint (i.e. in-group) profit for in-group recipient pairs except where this means that one member would receive less than they would have were the point of maximum fairness selected (i.e. on matrix type 1). Commitment to these sub-group norms (and the task) is high, with the male (previously a "fair" man) at least as keen as the females to comply to them (*"No, 7 and 1's got a bigger difference"*).

This sub-group's allocations across the three conditions is once again consistent with such an interpretation (see Appendix 9). For matrix types 1 and 3 a variety of responses

is evident in condition 1 intergroup allocations. By condition 3, however, all the sub-group members act in exactly the same way as each other and as they did in condition 2, i.e. they maximally prefer FAV over both MJP and F (and therefore neither MJP nor F exert any pull over FAV). The only deviation from this comes during intergroup allocations involving matrix type 2. Two of the subjects (06 and 21) again act identically to each other and to the "sub-group" norm (i.e. maximally preferring MD over MIP+MJP and minimally preferring MIP+MJP over MD), but the third subject (35) acts slightly differently. This subject showed no preference for either MD over MIP+MJP or for MIP+MJP over MD in condition 1 (and in fact opted for the point of maximal fairness for both condition 1 matrix type 2 intergroup allocations) and although he did move toward the sub-group norm by condition 3, he did not conform to it totally by *maximally* preferring MD over MIP+MJP. The reason for this cannot be ascertained. It is noteworthy however that he did choose the point of maximum in-group profit and in-group superiority on the o/i presentation of matrix type 2 (in line with the others and with the sub-group norm) but did not maximize in-group superiority at the cost of in-group profit on the o/i matrix presentation. Instead he chose the allocation which maximized in-group profit whilst ensuring in-group superiority (i.e. 12 points to the in-group member and 11 to the out-group member). Had he allocated any more points to the in-group member then he would not have achieved in-group superiority. Similarly, had he obtained any more in-group superiority then he would have sacrificed in-group profit. This is clearly in the "spirit" of the sub-group norm, but can perhaps be considered a somewhat more sophisticated modification of it.

Sub-group 4 began condition 2 with one of its members assuming a leadership role and allocating (lesser) tasks to the other sub-group members. A procedural protocol was arranged so that each sub-group member took it in turns to make "collective" decisions. The rationales for individual decisions were vague

and sparse to begin with, although early allocations seemed to be based on choosing the allocation on or near the scale midpoint ("we're actually visually averaging it as opposed to arithmetically averaging it"), with the slight deviations being determined by the subjects' seating positions ("we're choosing them relative to our position...with respect to the table"). After about 6 allocations the subjects began to go into some detail about how they had made their individual decisions in condition 1. The "leader" seemed to adopt fairness modified by slight in-group superiority on intergroup decisions, maximum joint (i.e. in-group) profit on in-group decisions, and fairness on out-group decisions. A second subject (who explicitly explained her individual strategy to the experimenter at the end of the experiment) allocated points so that the recipient of each pair with the higher code number got the lower points, regardless of whether the allocation was an intergroup or a same-group one (a strategy which incidentally would have resulted in in-group superiority on all intergroup allocations because of this subject's Group K membership and the relationship between category membership and individual code numbers). The third subject's individual strategy is difficult to ascertain. No explicit sub-group norms seemed to be established in condition 2, although there did seem to be increasing convergence on or compliance to the self-appointed leader's strategy ("What do you think about that one?"; "Shall we go for the one..."). There was, however, little evidence of strong commitment to this "sub-group norm".

A glance at this sub-group's allocations across the conditions (see Appendix 9) confirms the absence of any efficacious sub-group norms. In condition 3 each of the subjects chose quite different intergroup strategies, both to each other and, at least on matrix type 3, to their condition 1 strategies.

Sub-group 5 began their condition 2 session by explaining their individual condition 1 strategies in response to the

early question, "What are we going to base the decision on?". The male seemed to have made arbitrary decisions, based on "The number 7". One female had the general strategy of giving in-group members high marks and out-group members low ones. The second female seemed to have used fairness. The subjects then discovered their shared minimal group membership and quickly adopted a general sub-group norm of in-group favouritism, despite one female being a little dubious ("*This seems a bit petty doesn't it, just because somebody's not in your group giving them low marks*"). "Supplementary" norms were generated and adopted as the need arose. In the end the sub-group norms, to which there appears to have been high commitment, were: (i) maximum in-group superiority must be taken when available, (ii) when it is not (i.e. on o/i presentations of matrix type 3), choose "fairness" (which is also maximum in-group profit and minimum out-group profit, and additionally avoids in-group inferiority), (iii) maximize joint (i.e. in-group) profit on in-group allocations, where possible, (iv) where no position on in-group allocations maximizes joint (i.e. in-group profit) choose "fairness", (v) on out-group allocations minimize joint (i.e. out-group) profit, and (vi) where this is not possible choose "fairness".

Looking at this sub-group's allocations across the three conditions (see Appendix 9) it becomes clear that two of the subjects (52 and 85) were highly influenced by the sub-group norms. In the third condition these subjects' responses are identical to each other and to the condition 2 allocations (i.e. the sub-group norms), even though their condition 1 responses were quite divergent, especially on matrix type 2 intergroup allocations (i.e. the pulls of MD on MIP+MJP and MIP+MJP on MD). As in the third sub-group, however, one subject's (80) responses are more problematic. This subject's condition 3 intergroup allocations were substantially different to the allocations made by the other subjects and furthermore bore no stable relation to her condition 1 responses. It is impossible to ascertain why this is so, and

almost as difficult to guess. It can be noted, however, that after the first four of her condition 3 allocations this subject's responses all fall within one place of the position of maximum fairness, across all matrix types and recipient-patterns. Whether this reflects an explicit rejection of the sub-group norms in favour of an individually preferred one of fairness, or is simply the result of the subject losing interest in the task generally and adopting a "quick and easy" response pattern in order finish the experiment as soon as possible cannot be determined.

Sub-group 6 had only two members. In condition 2 the male admitted straight away that the whole task "*seems totally meaningless*" to him ("*This is absolutely nonsensical*"). There developed an element of turn-taking in condition 2, but the male made the majority of decisions, always arbitrarily. The female showed some inclination to opt for "*fairness*", almost by default, but became increasingly happy to divert from it and emulate the male's arbitrariness ("*I think we should have that one because it's my birthday*"). Thus the only "sub-group norm" that developed was one of randomness, and commitment to the task as a whole was low.

This sub-group's responses across the three conditions (see Appendix 9) is consistent with an interpretation that both subjects were responding more or less arbitrarily throughout the experiment.

Sub-group 7 attempted from the start of condition 2 to reach consensus for any decision that was made ("*Yes? Everyone?*"). Although they did not make their individual allocation decisions in condition 1 known to each other, they did tell the experimenter after the whole experimental session was over. One subject (58) simply strove to obtain more points for in-group members than for out-group members. A second (79) had the same aim, but modified it by trying to give more points to those with code numbers close to her own. The third

(96) allocated randomly at first but developed a preference to give more points to people with similar code numbers to his own. (It is worth remembering that adopting a strategy of giving more points to people with code numbers similar to one's own would result in in-group preference on intergroup allocations.) In condition 2 the subjects in Sub-group 7 began by taking it in turns to make suggestions which were then approved by the others before the selections were finally made ("So equal?"). Following an the intergroup allocation during which the in-group was favoured they each revealed their minimal group memberships and discovered their shared minimal group membership. Although no sub-group norms were made explicit it is clear that they developed ("We've elected this strategy that we're sticking to rigidly now", "We have. Yes. I know"). These norms were, roughly, (i) wherever possible maximize in-group superiority, (ii) on all other occasions, choose maximum "fairness". Commitment to the first norm was high, but one subject seemed unhappy that the second norm was not partisan enough ("what we should probably have gone...for isn't two 13s...It's probably 8 and the 3, or the 7 and the 1...as the two aggregate points are lower than that"). The subjects in this sub-group also told the experimenter their condition 3 strategies. The subject (96) who had employed randomness then a code number strategy in condition 1 pursued a condition 3 strategy of allocating more to the in-group than to the out-group. The subject (79) who had modified in-group bias with a code-number strategy in condition 1 abandoned that modification in condition 3, following instead the simple sub-group norm of in-group bias. And the subject (58) who had employed simple in-group bias in condition 1 adopted a strategy of minimizing out-group profit where possible in condition 3 (which has an outcome identical to one resulting from a strategy of maximizing in-group superiority on the Tajfel matrix types used in this study).

These subjects' responses over the three conditions (see Appendix 9) supports this interpretation, but perhaps not as

strongly as might be expected. On intergroup allocations involving matrix type 3 (i.e. FAV on F and F on FAV) all three subjects' responses are nearly identical to each other and to the condition 2 allocations, despite wide divergence across subjects in their equivalent condition 1 allocations. Similarly, on intergroup allocations involving matrix type 1 (FAV on MJP and MJP on FAV) two of the subjects (58 and 79) make condition 3 choices identical to each other and to the condition 2 choices, despite their condition 1 responses being different to each other and to the condition 2 choices. The third subject's (96) choices on these matrices are slightly different in as much as his condition 3 FAV on MJP score was closer to his equivalent condition 1 score than to the group norm/others' FAV on MJP scores, and his condition 3 MJP on FAV score is less than both his condition 1 MJP on FAV score and the group norm/others' MJP on FAV scores (the latter two of which are identical), although these differences are slight. As with sub-groups 3 and 5 above, however, the main interpretive problems are caused by responses on matrix type 2 (MD on MIP+MJP and MIP+MJP on MD). Here the subjects' condition 3 scores bear no discernable relation to either their condition 1 scores or to the "sub-group norm"/condition 2 scores. By "breaking down" the two relevant pull-scores into their constituent parts it becomes clear that in condition 3 the three subjects made near-identical allocations on the o/i presentation of matrix type 2. Two subjects chose 19/25 (i.e. the point of maximum in-group profit and superiority, and the same allocation as was chosen in the group condition), and the third (96) chose 18/23 (i.e. one scale-place away from this point). The difference between the subjects' pull-scores from matrix-type 2 therefore came predominantly from their different choices on the i/o presentation of this matrix. One subject (96) chose the point of maximum difference, which was the point "collectively" chosen in condition 2. Another (subject 79) chose the opposite end of the scale, i.e. the point of maximum in-group (and joint) profit. The third subject (58) chose 12/11, i.e.

the point of maximum in-group profit possible whilst simultaneously achieving in-group superiority. None of the subjects' i/o matrix 2 choices in condition 3 were the same as the choices those subjects made in condition 1. As suggested in relation to sub-group 3 above, this is consistent with an interpretation that all three subjects were attempting to follow a strategy of in-group favouritism (i.e. the sub-group norm), but were unclear as to the appropriate form that such in-group favouritism should take on i/o presentations of matrix type 2 (e.g. MD, MIP, or MIP whilst ensuring MD). It is worth mentioning that the equivalent choice in condition 2 was the one which generated the greatest amount of in-group discussion as to the appropriate allocation ("*Are we going for difference or greater number?*" - see Appendix 8), although it should additionally be noted that this was also the first allocation after the subjects' made their minimal category memberships known to each other.

Sub-group 8 began condition 2 with each of the subjects revealing their minimal group memberships, although they do not explicitly tell each other how they had made their individual allocations in condition 1. This sub-group made herculean efforts to reach consensus on each decision although a consensual strategy proved hard to come by, largely because the subjects seemed to misunderstand what each other wanted to do. With time, experience and effort, though, a main intergroup allocation norm of maximum in-group profit emerges ("*the maximum number of K, it doesn't matter if we give them a higher number, does it... 'cos whatever they get doesn't affect us so all we're interested in is maximizing our own*"; "*The whole point is to maximize...our group scores, isn't it...?*" "Yeah").

Once again this interpretation is corroborated by looking at the subjects' allocations over all three conditions (see Appendix 9). In condition 3 subjects unanimously opted for a strategy of MIP on matrix types 2 and 3, just as they had in

condition 2, despite wide variation in their previous individual (i.e. condition 1) allocations. That is, in condition 3, as in condition 2, pulls of MD on MIP+MJP and F on FAV were zero and pulls of MIP+MJP on MD and FAV on F were maximal. This time the fly in the interpretive ointment is found for matrix-type 1 allocations. Two of the subjects (14 and 48) make identical allocations to each other but not to the equivalent condition 2 choice. The reason for this difference, however, is that the "final" sub-group norm had not been established by the time the relevant type 1 matrices were completed in condition 2 (they were completed whilst the female was pushing for a "compromise" sub-group norm of fairness). These subjects' condition 3 allocations are, in fact, consistent with the "final" sub-group norm (of maximum in-group profit), even though the condition 2 allocation was not! The reason for the third subject's (48) deviation from this norm is not clear, although it may be relevant that this deviation stems from a single "fair" allocation on a i/o presentation of a type 1 matrix (the o/i presentation obtained maximum in-group profit, as did all the other intergroup allocations) and that this was the very first allocation subject 48 made during condition 3.

To summarize the qualitative data, four sub-groups developed and tended to adhere to strong primary sub-group norms, three of which were ethnocentric (sub-groups 3, 5 and 7), while the fourth was one of maximum in-group profit (sub-group 8). A fifth sub-group (2) may have developed a primary sub-group norm of in-group preference, but if so commitment to the norm was weak. A sixth sub-group (6) may have developed a single "in-group norm" of arbitrariness. The remaining two sub-groups (1 and 4) developed no sub-group norms.

It should be stressed that where sub-groups developed and adhered to norms, those norms operated at two levels. The primary norms just referred to can be thought of as general (i.e. superordinate) norms, reflecting the "overall stance"

sub-groups wanted to take in the minimal group situation as a whole (e.g. discriminatory, profiteering). In particular situations, however, several sub-groups employed a number of different specific norms, so that, for example, different norms would be activated for different recipient patterns and for different matrix types. Often these specific norms were simply subordinate (i.e. lower-level) manifestations of the sub-groups' primary (higher-level) superordinate norms: merely dictating how the general norms should be best applied in various particular circumstances. Sometimes, however, these specific norms were supplementary, in the sense that the general norm did not apply (made no sense, could not be followed) in particular circumstances, and so additional specific norms were sometimes developed to guide group behaviour in such situations. Employing these terms it is clear that it has been argued above that several sub-groups developed and adhered to general norms, but variations in these sub-group members' individual allocation decisions arose because of those members' disagreement or confusion about, or rejection of, subordinate or supplementary specific norms.

DISCUSSION

Condition 1 of the present study replicated the "typical" minimal group paradigm result of significant sample-level in-group favouritism and fairness across groups, thus supporting hypotheses H1 and H2. Sample-level in-group favouritism (as indicated by a composite measure of FAV on MJP, MD on MIP+MJP and FAV on F) polarized from condition 1 to condition 3, but sample-level fairness across groups (i.e. F on FAV) depolarized. Thus hypothesis H6 received support but hypothesis H7 did not. The pull score of MIP+MJP on MD also polarized from condition 1 to condition 3. Intergroup behaviour within sub-groups did not polarize from condition 1 to condition 3, either "overall" or when particular intergroup

strategies were considered separately. Thus no support was obtained for hypotheses H4 or H5. Similarly, variance of intergroup behaviour within sub-groups was not significantly less in condition 3 than in condition 1, so support was not found for hypothesis H3.

The first main result of the present study is clearly that intragroup communication resulted in greater sample mean in-group favouritism and less sample mean fairness across groups than intergroup behaviour without intragroup communication. Thus, while mere social categorization may be sufficient to promote intergroup discrimination, social categorization accompanied by intragroup communication appears to exacerbate that intergroup discrimination, whilst simultaneously decreasing intergroup fairness. This supports generalization of social identity theory's "sufficiency condition" beyond the minimal group paradigm: it appears that social categorization per se promotes intergroup discrimination, and the stronger or more salient that categorization, brought about here by intragroup communication, the stronger the subsequent discrimination in favour of the in-group.

The second main result obtained in the present study is that intra sub-group behaviour was not less variable in condition 3 than in condition 1. This suggests that contrary to expectations intragroup communication did not lead to the development of and adherence to in-group norms.

In direct opposition to these conclusions it is possible to argue that the results of the present study are best understood in terms of the development of in-group norms. Such an argument consists of three main claims. First, that no current theory of minimal group behaviour can account for the variety of behaviour typically found in the minimal group paradigm. Second, that social category or group membership can only serve as an explanatory (and predictive) variable for behaviour when one or more norms exist which are associated

with that group membership. Third, that the best social psychological explanation which can be given for behaviour subsequent to intragroup interaction is one which considers whether or not that behaviour conforms to in-group norms determined or subscribed to during that interaction.

Condition 1 of the present study yielded the typical minimal group paradigm result of significant sample mean in-group favouritism. As in previous studies, however, not all subjects discriminated following social categorization. Further, some subjects not only did not discriminate, they actively pursued particular non-discriminatory strategies such as fairness, randomness, arbitrariness, strategies based on recipients' code numbers, etc.. No current theory of minimal group behaviour can account for such variety of behaviour, still less predict which particular subjects will pursue which particular allocation strategies following social categorization. This is because all current theories of minimal group behaviour posit single explanations for that behaviour. Thus, any non-predicted behaviour has to be "explained away" by such theories as exceptional cases.

To take social identity theory as an example, this theory explains minimal group behaviour by hypothesizing that mere social categorization results in subjects identifying with their imposed "in-groups" which they then need to perceive as positively distinct from relevant "out-groups" in order to serve social identity and self-esteem needs. Thus, subjects are held to engage in intergroup discrimination in order to bring about such perceptions of in-group superiority. That is, mean in-group favouritism in the minimal group paradigm is claimed to result from individuals engaging in intergroup discrimination following mere social categorization in order to serve individual self-esteem needs. Social identity theory therefore has to "explain away" any instances of individuals not engaging in in-group favouritism following mere social categorization. It does this by claiming that those who do

not discriminate fail to do so either because they do not identify with their imposed in-group and/or because they serve their self-esteem needs in other ways (e.g. through salient personal identity or alternative salient social identities). This is unsatisfactory because any and all non-discriminatory behaviour following social categorization can be "explained away" in such a fashion (making social identity theory both rather weak as a predictive theory and potentially unfalsifiable - see Abrams, 1992: 62). It is also unsatisfactory because the theory can offer no explanation for non-discriminatory *strategic* behaviour following social categorization (e.g. fairness).⁵

The sheer variety of behaviours resulting from mere social categorization in the minimal group paradigm suggests that something other than the social categorization itself is needed to give a full and satisfactory explanation of minimal group behaviour. Unless and until that "something else" can be specified, minimal group behaviour must be considered to be largely determined by "individual differences" within the subject population.

The second strand of the argument being presented here is that social category or group membership can only serve as an explanatory (and predictive) variable for behaviour when one or more norms have been identified which are associated with that group membership. No such norms are associated with minimal group membership because of the strict procedural criteria applied within the minimal group paradigm. Minimal groups have no history and so new members of those groups cannot refer to previously established in-group norms. Also, because no intragroup communication is allowed in the minimal group paradigm, minimal group members cannot establish in-group norms themselves. Upon finding themselves in the minimal group paradigm subjects can, if they feel so inclined,

⁵ Similar remarks can be directed toward other posited explanations of minimal group behaviour (see below).

infer what appropriate in-group behaviour might be, perhaps on the basis of "clues" from the experimenter or the experimental situation, or perhaps on the basis of norms associated with other groups subjects are members of. Alternatively subjects can act in accordance with personal norms or personally subscribed to ideologies. Alternatively again, subjects can simply dismiss the task as pointless and adopt random or arbitrary behavioural guidelines.

Intragroup interaction makes the development of in-group norms possible. Such norms are potentially functional in that they can either provide guidelines for appropriate behaviour and/or they can validate previously adopted individual behaviour as appropriate. Each of these functions are likely to be attractive to subjects in the unfamiliar situation of the minimal group paradigm. If subjects are unclear about what appropriate behaviour might be within the minimal group paradigm then one option open to them in condition 2 of the present study is to look to others in an identical situation for guidance and/or verification. When this occurs then in-group norms may develop.

The development of in-group norms is clearly not inevitable during intragroup interaction, however. First, if people believe that they already know how to behave appropriately then they will have no need to look to others for guidance or verification (although an awareness that other group members are following alternative strategies may make such people examine whether the reasons for their own strategies are legitimate). Second, if no suggested behaviour seems any more appropriate than any other then there will be little motivation to modify behaviour already adopted. Third, if a variety of behaviours different to one's own are suggested as superior to one's own, but no reasons exist to believe that they actually are superior (e.g. when none of the other suggestions come to be adopted by others in similar circumstances), then again there will be no reason to adopt

alternative strategies in preference to one's own. In summary, if one is given no reason to suppose that one's own behaviour is less appropriate than available alternatives then one will not feel the need to turn to others for guidance or for verification of one's own behaviour, and in-group norms will not develop.

In the present study subjects in condition 2 had already made a series of individual allocation decisions. If those subjects decided upon a strategy that seemed appropriate to the situation, and were not presented with reasons to think that alternative strategies were more appropriate, then the subjects would feel no compunction to develop or adhere to in-group norms. In sub-group 1, for example, two subjects had individually decided that fairness was an appropriate strategy but the third had decided that randomness was appropriate (in the absence of any reason to adopt a more specific strategy). The fair subjects provided the random subject with no reason to believe her strategy inappropriate and the random subject provided the fair subjects with no reason to think their strategy inappropriate. (The fair subjects did however verify each other's strategy as appropriate.) Adopting a condition 2 allocation strategy of turn-taking meant that each subject could pursue her own individually preferred strategy and no in-group norm was necessary or desirable. A similar analysis can be made with respect to sub-group 4.

An implication of such reasoning is that in-group norms would probably have been more commonly adopted in the present study had subjects not made a series of individual decisions before being placed in the sub-group situation. A future study might sensibly investigate this.

When people are unsure what constitutes appropriate behaviour there are two main methods by which they can be convinced that a particular behaviour is appropriate. They can be presented with persuasive (rational or rhetorical) arguments and/or they

can be presented with evidence that others in similar circumstances are more or less consensually behaving in a particular manner. In sub-group 3, for example, the male subject had adopted a fairness strategy in condition 1 but discovered in condition 2 that each of his fellow sub-group members had adopted strategies of in-group preference. This subject clearly had no specific commitment to his previous strategy ("*you don't know any of them so why not?*") and in the face of others in a similar situation to himself behaving in a consistent manner he brought his behaviour into line with their's, and a sub-group norm of in-group preference developed. A similar analysis can be given with respect to sub-group 7.⁶

A defining characteristic of the minimal group paradigm is minimal group membership. It is to be expected that strategies which take such membership into account would be more persuasive (i.e. would seem more appropriate) than strategies which do not, particularly if subjects discover that others who have adopted group-relevant strategies are in the same minimal group as themselves. This is illustrated by the development of an in-group norm for sub-group 5. The members of this sub-group each employed different condition 1 (i.e. individual) strategies: one random, one (probably) fair, and one in-group preferential. Upon discovering in condition 2 both that different individual strategies had been employed by other in-group members and that all three subjects shared the same minimal group membership it was the group-relevant (i.e. in-group preferential) strategy which became adopted as

⁶ It can also be argued that the female member of sub-group 6 was persuaded by the male member that arbitrariness was the most appropriate behaviour in the minimal group situation and that an in-group norm of arbitrariness developed. Similarly, it can be argued that the other members of sub-group 2 (particularly the male member) failed to convince the third member that in-group favouritism was the most appropriate strategy in the situation and that this was the reason for this sub-group's failure to develop a consensual in-group norm.

the sub-group norm. Similarly, when the members of sub-group 8 discovered that they shared minimal group membership strenuous efforts were made to adopt behaviour which took account of that shared minimal group membership, in this case maximum in-group profit. Again, in sub-group 7 in-group preference became adopted as the main in-group norm, despite the fact that 2 of the three sub-group members had adopted similar non group-related allocation strategies in condition 1 (i.e. ones based on subjects' unique code numbers).

To summarize the arguments being made with respect to the development of in-group norms, it is claimed that intragroup communication will lead to the development of in-group norms to the extent that group members feel a need for guidance toward or verification of appropriate behaviour. Such a need will be inversely proportional to members' beliefs that they have already discovered appropriate behaviour. Discovering that other group members have adopted behaviour dissimilar to their own may cause group members to re-examine the appropriateness of their own behaviour, but unless reasons are provided to make them believe that such alternatives actually are more appropriate they will not modify or abandon their own preferred strategies. Strategies will tend to be deemed appropriate to the extent that they are adopted by people in similar circumstances and to the extent that they take shared salient characteristics into account. Thus, discovering that others in the same situation as oneself have adopted behaviours similar to one's own is likely to lead to the development of a group norm reflecting those similarities.

Much of this is speculative. No doubt a more thorough account can and should be developed of the conditions in which in-group norms will or will not arise and/or be adhered to. Similarly, a more rigorous account needs to be developed to explain and predict the content of in-group norms which do develop. Future research could attempt to do both of these things. The important point here is that in-group norms did

develop for some sub-groups and that the behaviour of those sub-groups (i.e. in condition 2) is best understood in terms of those norms, rather than in terms of (awareness of) sub-group or minimal group membership per se.⁷ When in-group norms did not develop, sub-group (i.e. condition 2) behaviour is best understood in terms of the sub-group members' individual (i.e. condition 1) allocation strategies and the decision-making processes employed in condition 2 (e.g. turn-taking), again rather than in terms of (awareness of) sub-group or minimal group membership per se.

Once in-group norms have been established they can be used by group members even when acting alone. Whether or not they are used will depend on many things, including the salience of in-group membership, the appropriateness of the in-group norms to the current situation, the compatibility of those norms with personally desired behaviours or outcomes, etc.. Few of these considerations are relevant in the present study, however, and by and large we should expect condition 3 behaviour to reflect in-group norms when such norms have been adopted. And, by and large, it does. Even when it does not, a reasonable explanation can usually be provided for why it does not (e.g., because of confusion as to how best to implement a primary in-group norm in particular circumstances, see the qualitative analysis in the results section above). When in-group norms were not developed by sub-groups in condition 2, the condition 3 behaviour of those sub-groups' members is best understood in terms of their individually adopted condition 1 strategies.

Because intragroup interaction does not inevitably lead to the development of or adherence to in-group norms the lack of a

⁷ An exception is provided by sub-group 8's minimal FAV on MJP score. Had this group been pursuing its strategy of maximum in-group profit this score should have been maximal. This exception is accounted for when it is realised that the relevant condition 2 allocations were made *before* the sub-group norm was established (see the qualitative analysis in the results section above).

decrease in behavioural variability within sub-groups from condition 1 to condition 3 does not negate an explanation of the present study's results in terms of in-group norms. A number of sub-groups did not develop in-group norms, at least one sub-group appeared to develop an in-group norm of arbitrariness, some primary in-group norms were sufficiently vague to allow a variety of norm-consistent behaviours in specific situations, and the development of in-group norms does not (usually) preclude the possibility of individual deviations from those norms. All of these considerations mean that the lack of a general decrease in behavioural variability within sub-groups across conditions 1 and 3 was to be expected. If a future study encouraged development of and adherence to non-random in-group norms then such a general decrease in behavioural variation within sub-groups would be expected. Also, if those norms were non-discriminatory then it would be predicted that condition 3 behaviour would be less discriminatory than condition 1 behaviour, as well as being more homogeneous.

The argument being developed here also implies that in-group norm formation will not reliably result in polarization of within sub-group behaviour. In-group norms may reflect individual behaviour already consensually adopted by in-group members, or they may preclude the use of previously adopted individual strategies. In either case the strategies considered will not polarize. Polarization will also not be expected when no in-group norm formation takes place.

Explaining subjects' condition 3 behaviour in terms of adherence or non-adherence to in-group norms is clearly a social psychological explanation. The claim is that in certain circumstances in-group members' behaviour will conform to norms associated with the in-groups they are members of. Knowledge of those norms allows (in certain circumstances) prediction of group members' behaviour. Particularly, knowledge of in-group norms (and knowledge that group members

are committed to those norms) allows explanation and prediction of uniformities of (group and) intergroup behaviour, an ability identified by Tajfel "the aim of a theory of intergroup behaviour" (Tajfel, 1981a: 46, emphasis added). Such an account is clearly non-reductionist in that it relies on considering groups as "entities in their own right" (as opposed to considering them as mere aggregates of individuals), with "group-level" norms associated with them which may or may not accord with the personal norms of in-group members.

It might reasonably be argued that the normative explanation being developed here is unacceptably post-hoc. The main claim being made in this chapter is that (awareness of) group membership can only serve as an explanatory and predictive variable when at least one norm can be identified which is associated with the group membership of interest. Such norm identification could not occur in advance in the present study, either for the minimal groups or for the sub-groups comprising of members of previously minimal groups. This was for the simple reason that such norms *did not exist* in advance of the present study. Because minimal group members are not offered the opportunity of generating or adopting common in-group norms, no satisfactory social psychological explanation of their behaviour can be offered at all. Similarly, because in the present study sub-groups had to generate their own in-group norms (as opposed to accepting previously established ones), explanation of sub-group members' behaviour *had to be* post-hoc.

Normative explanations are not necessarily post-hoc, however. Where shared norms of group behaviour can be identified it is possible to predict that group behaviour will (in some circumstances) conform to such norms. If, for example, minimal groups are comprised of people who all subscribe to a particular salient ideology (e.g. egalitarianism) then it is possible to predict that their minimal group behaviour would

conform to such ideological norms. Further, if such people were then permitted intragroup communication it can be predicted that the in-group norms would be solidified and more consistently adhered to. This suggests that the best way of predicting the behaviour of group members (as group members) is to identify (or manufacture) the norms, stereotypes, or social representations associated with in-group membership. Put like this the present argument seems perhaps a little less contentious.

It should also be acknowledged that although it was not done, it would have been perfectly possible to predict with an impressive degree of accuracy the condition 3 behaviour of subjects in the present study given information about (i) their behaviour in condition 1,⁸ (ii) whether or not in-group norms were established in condition 2, and, if so, (iii) what they were.⁹ Such prediction would not be post-hoc if done before the results of the third condition were known. Importantly, using such a method it would have been possible to predict both mean behaviour for *particular sub-groups* and the behaviour of *particular sub-group members*, as well as mean behaviour for the sample overall. This is in stark contrast to the seemingly more parsimonious and a priori explanation offered by social identity theory.

Before reaching conclusions from the present study the normative argument being developed should be more explicitly compared against other potential explanations of the present study's results, such as provided by social identity theory (Tajfel & Turner, 1979), generic norm theory (Branthwaite et al., 1979), equity theory (Ng, 1986), the behavioural

⁸ This information would be needed to predict the behaviour of subjects in sub-groups in which no in-group norms were adopted, i.e. for which no social psychological explanation could be offered.

⁹ Note that the actual behaviour in condition 2 need not be known (and might actually be misleading).

interaction model (Rabbie et al., 1989), intergroup accentuation theory (Doise & Sinclair, 1973), self-categorization theory (Turner et al., 1987), or the demand characteristics explanation (Gerard & Hoyt, 1974). To do this it is necessary to be clear about what is and what is not being argued. It is not being argued that the normative account being developed can explain minimal group behaviour. Nor are any strong claims being made about when norms will arise or be adhered to. What is being claimed is that no satisfactory social psychological explanation can be offered for minimal group behaviour, and that any genuine social psychological explanation of group behaviour has to be restricted to instances in which group norms exist or are developed.

In this thesis it has now been repeated argued, and demonstrated, that social categorization (accompanied by recognition that one is a member of one group but not of another) is not sufficient to promote intergroup discrimination. In the study reported in this chapter, as with the studies reported in the two preceding chapters, social categorization was followed by subjects employing a variety of allocation strategies, both across individuals and across allocations for the same individuals. None of the present explanations of minimal group behaviour, relying as they do on single explanatory variables, can provide satisfactory accounts of such diversity. In the present study it has also been seen that even when social categorization is accompanied by intragroup communication, a variety of individual and sub-group behaviours are still adopted by subjects. Thus, explanations of such behaviour still cannot be given in terms of social categorization per se, or in terms of social categorization plus the mere fact of intragroup communication. It cannot be feasibly argued, for example, that intragroup communication merely makes more salient the single explanatory variable each of the above accounts relies upon, e.g. social identity contingent self-esteem, generic

norms, expected in-group favouritism on the part of out-group members, greater positive interdependence with in-group than with out-group members, in-group membership in a multi-group context, in-group prototypes, or demand characteristics. Such a move would still require the "explaining away" of much if not most of the behaviour exhibited by subjects in conditions 2 and 3.¹⁰

Neither is it the case that explanations of group polarization, such as provided by persuasive arguments theory (Vinokur & Burnstein, 1978) or social comparison theory (Myers & Lamm, 1976), are sufficient to explain the results of the present study. Although each of these theories have something to say about when and why group norms will develop and be adhered to, and about what the content of such norms will be, each are explicitly explanations of group polarization, and it is claimed (and found) here that intragroup communication is not sufficient to provoke such polarization. Thus, as with the explanations provided by theories of minimal group behaviour, much if not most of the present study's results must be "explained away" by theories of group polarization.

The strongest alternative to a normative account of the present study's findings comes from Turner's (1981) referent informational influence model which (i) combines aspects of both normative and informational influence (Deutsch & Gerard, 1955), and (ii) provides the basis for Turner's development of Tajfel's social identity theory which resulted in self-categorization theory (Turner et al., 1987). Referent informational influence is held by Turner (1981: 108) to take place in three stages:-

¹⁰ Although such variables were not explicitly examined in the content analysis conducted on sub-group members' decision making processes in condition 2, a brief examination of the transcripts of those processes (see Appendix 9) reveals that concern for such variables was far from being universal. This again this severely calls into question explanatory accounts simply in terms of such variables.

1. the individual defines himself as a member of a distinct social category;
2. the individual forms or learns the stereotypic norms of that category. He or she ascertains 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. the individual assigns these norms to himself in the same way that he assigns other stereotypic characteristics of the category to himself when his category membership becomes psychologically salient. Thus his behaviour becomes more normative (conformist) as his category membership becomes salient.

The normative account being developed here has considerable similarities to referent informational influence, and so the two need not be seen, in general, to be in competition. There are differences between the two, however, only the main one of which will be dealt with here (but see also Chapter 8).

The referent informational influence model sees in-group norms as developing via in-group members conforming to in-group prototypes whose main characteristic is that they maximize or reflect a ratio of similarities between in-group members and differences between in-group members and out-group members: norms are prototypical to the extent that they "are used to define the category as different from other categories". The normative account being developed here, however, explicitly rejects that (i) in-group norms are necessarily developed and adhered to via a process of reflecting or accentuating within-group similarities and between-group differences, and (ii) that accentuating within-group similarities and between-group differences is a necessary function of in-group norms. That is, although some norms are clearly "intergroup", both in origin and focus, not all are. It is quite possible, and indeed common, for groups to develop norms which simply reflect in-group characteristics, serve in-group needs, and

prescribe in-group behaviour (broadly conceived), with no other group, conceived as a group, necessarily involved.¹¹

CONCLUSIONS

The main conclusion of this chapter is that despite the fact that the "main" results of the present study seem at first glance to support a social identity theory explanation of minimal group behaviour, and also to support the generalization of that theory beyond the minimal group paradigm, the best interpretation of the present study's results is a normative one. Although complex and largely (in the present study) post-hoc, such an explanation is genuinely social psychological in that it can potentially predict and explain uniformities (and variations) of group and intergroup behaviour at both the group and the group member levels of analysis. Explanations in terms of social categorization per se cannot do this, relying as they do on universal individualistic motivational accounts. At best such accounts can only explain (often barely representative) mean behaviour and must "explain away" all deviations from such behaviour.

A second conclusion has to be that much more work needs to be done to specify when group norms will or will not be adopted and adhered to, and what the content of those norms will be when they are. This point is returned to in Chapter 8, where it is suggested that Moscovici's social representations theory (e.g. Farr & Moscovici, 1984) and Turner's referent

¹¹ An example is offered by the primary in-group norm for sub-group 8: maximum in-group profit without regard to the consequences of this for (i) out-group profit or loss, or (ii) in-group distinctiveness from the out-group - "it doesn't matter if we give them a higher number, does it... 'cos whatever they get doesn't affect us so all we're interested in is maximizing our own".

informational influence model (e.g. Turner, 1981) will be of particular use when addressing such issues.

Central to the normative explanation developed here is the claim that group behaviour is better explained and predicted by group norms than by the individual social identity contingent self-esteem needs of in-group members (or by any other postulated universal individual motivation). Ironically, many social identity theorists are themselves now coming to doubt this motivational aspect of social identity theory (e.g. Hogg & Abrams, 1993). Nevertheless, in the next chapter it will be argued that social identity theory's self-esteem hypothesis has been widely misunderstood and has never been adequately tested. Before a normative explanation of intergroup behaviour can be properly pitted against social identity theory's self-esteem hypothesis it is critical that the clearest and most persuasive form of that hypothesis should be identified. Additionally, in a later chapter it will be argued that even the best form of the self-esteem hypothesis fails to do full justice to Tajfel's theoretical explanation of the motivation for particular forms of group and intergroup behaviour. Finally, it will be argued that the normative account being developed here is consistent with Tajfel's "intended" account of motivation within social identity theory.

CHAPTER 5: STUDY 4 - THE NATURE OF SELF-ESTEEM**CHAPTER OVERVIEW**

The study reported in this chapter employed four different measures of self-esteem. The first was a "general trait" measure (i.e. self-esteem from all aspects of self over time). The second was a "general state" measure (i.e. self-esteem from all aspects of self at a particular moment). The third was a "specific trait" measure (i.e. self-esteem derived from a subject's national group over time). The fourth was a "specific state" measure (i.e. self-esteem derived from a subject's national group at a particular moment). In addition, a "general state" in-group evaluation measure was employed (i.e. measuring evaluation of a subject's national group at a particular moment).

An argument is presented that the self-esteem hypothesis within social identity theory is best tested with a "specific (to a particular in-group) state" self-esteem measure. It was therefore hypothesized that of all the self-esteem measures, this would be the one most sensitive to changes in self-esteem as a result of manipulations intended to make subjects feel either more positive or more negative about their national group than they had moments earlier.

As predicted, the two trait self-esteem measures detected no significant change in self-esteem as a result of either the positive or the negative manipulation. Both state self-esteem measures detected a significant rise in self-esteem as a result of the positive manipulation, but only the one specific to national group membership detected a significant fall in self-esteem as a result of the negative manipulation.

It was concluded that a "specific state" self-esteem measure is most appropriate when testing the self-esteem hypothesis.

INTRODUCTION

Social identity theory says that at least some intergroup discrimination is motivated by individuals' self-esteem needs and that intergroup discrimination can sometimes satisfy such needs (Tajfel & Turner, 1979). Abrams & Hogg (1988; Hogg & Abrams, 1990) identify two corollaries of this self-esteem hypothesis:-

1. Successful intergroup discrimination enhances social identity and thus elevates self-esteem. Self-esteem is a dependent variable, a product of specific forms of intergroup behaviour.
2. Depressed or threatened self-esteem promotes intergroup discrimination because of a need for self-esteem. Self-esteem is an independent variable, a motivating force for specific forms of intergroup behaviour.

Hogg & Abrams (1990: 33)¹

Empirical tests of corollary 1 of social identity theory's self-esteem hypothesis have investigated whether or not in-group favouritism raises self-esteem. Oakes & Turner (1980) found that following social categorization subjects who engaged in in-group favouritism had higher post-task self-esteem than subjects who merely read a newspaper. Lemyre & Smith (1985) found that social categorization lowered self-esteem from its pre-categorization levels and that in-group favouritism restored self-esteem to its pre-categorization levels (but that fair allocations across groups did not, and neither did allocations solely to in-group members or solely to out-group members). Wagner et al. (1986) found no correlation between levels of in-group favouritism and subsequent increases in self-esteem relative to pre-discrimination levels. Similarly, Hogg & Sunderland (1991: 58) found that "greater intergroup discrimination was not

¹ See also Abrams & Hogg (1988: 320)

associated with higher post-test levels and/or significant increase in" self-esteem. Finally, Vickers, Abrams & Hogg (1985, cited in Hogg & Abrams, 1990: 34; 1988, cited in Abrams & Hogg, 1988: 320) found that in-group favouritism lowered self-esteem when a "local norm of cooperation was salient" (Abrams & Hogg, 1988: 320, see also Diehl, 1989).

While it is interesting and important whether or not intergroup discrimination (however conceived) per se affects self-esteem, this does not test Abrams & Hogg's formulation of corollary 1 which clearly states that it is *successful* intergroup discrimination which is predicted by social identity theory to raise self-esteem. Although Abrams & Hogg do not specify what constitutes successful intergroup discrimination, none of the studies reviewed by them investigated whether or not subjects who discriminated believed their discrimination to have been (or likely to have been) successful (e.g. in achieving or maintaining secure and legitimate positive in-group distinctiveness). If they did not, then corollary 1 of social identity theory's self-esteem hypothesis, as formulated by Abrams & Hogg, has simply not been tested by these studies.

The principal method by which corollary 2 of social identity theory's self-esteem hypothesis has been tested has been to investigate whether subjects with (or assumed to have) low self-esteem show more in-group favouritism than subjects with (or assumed to have) high self-esteem. Wagner et al. (1986) found that members of low-status groups displayed more out-group derogation than members of high-status groups. Hogg & Sunderland found that following social categorization subjects with relatively low self-esteem discriminated in favour of their in-group more than subjects with relatively high self-esteem. Crocker & Schwartz (1985) found that subjects with relatively low self-esteem engaged in greater subsequent out-group derogation than subjects with relatively high levels of self-esteem while Crocker et al. (1987) found that subjects

with relatively low self-esteem engaged in greater out-group and in-group derogation than subjects with relatively high levels of self-esteem. Subjects with relatively low self-esteem in these two studies did not, however, display more preference for the in-group over the out-group than subjects with relatively high self-esteem. Finchilescu (1986) found that members of low status groups discriminated more in favour of their in-group than members of high-status groups.² Sachdev & Bourhis (1984) found that minority group members displayed similar amounts of in-group favouritism as displayed by majority group members.³ Abrams (1982, 1983, cited in Abrams & Hogg, 1988: 320-321) found that levels of self-esteem were positively correlated with subsequent in-group favouritism. Sachdev & Bourhis (1985) found that members of equal, high and total power groups indulged in greater levels of in-group favouritism than subjects in low or no power groups. Sachdev & Bourhis (1987) similarly found that members of equal and high status groups indulged in greater levels of in-group favouritism than subjects in low status groups.⁴

² This result is confounded by the fact that some members of high and low status groups felt that they "should" really belong to the other group and had been miscategorized. Overall, subjects who thought that they were "really" members of the low-status group discriminated more than subjects who thought that they were "really" members of the high-status groups, whether or not subjects had actually been categorized into their "proper" groups. It is a moot point, however, whether subjects who think that they "should" belong to an ostensibly low-status group personally evaluate that group as low-status.

³ Simon & Brown (1987: 708) found that minority group members in their study "identified more strongly with their in-group than did subjects allegedly belonging to a nonminority", thereby providing suggestive evidence against Sachdev & Bourhis's (1984) assumption that minority group membership automatically confers less self-esteem than majority group membership.

⁴ Similarly, Long et al. (1994) found that subjects with relatively high (personal) self-esteem engaged in greater in-group favouritism than subjects with relatively low (personal) self-esteem.

As before, while it is interesting and important whether or not subjects with low self-esteem subsequently discriminate in favour of in-groups more (or less) than subjects with high self-esteem, this does not test Abrams & Hogg's formulation of the second corollary of the self-esteem hypothesis within social identity theory. That formulation clearly states that intergroup discrimination can be promoted by low or *threatened* self-esteem. None of the studies reviewed by Abrams & Hogg seem to have investigated whether or not the subjects with (or assumed to have) high self-esteem felt that their self-esteem was secure. If they did not, then corollary 2 of social identity theory's self-esteem hypothesis, as formulated by Abrams & Hogg, has simply not been tested by these studies.

A second difficulty with empirical tests of Abrams & Hogg's second corollary of social identity theory's self-esteem hypothesis concerns the operationalization of "low" or "depressed" self-esteem. In some studies (e.g. Crocker & Schwartz, 1985; Crocker et al., 1987; Long, et al., 1994) subjects have been differentiated into categories of "low" and "high" self-esteem on the basis of median or tripartite splits on self-esteem scores. That is, self-esteem scores were taken for all subjects and those with *relatively* high self-esteem have been deemed "high" self-esteem subjects and subjects with *relatively* low self-esteem have been deemed "low" self-esteem subjects, even though the mean self-esteem for "low" self-esteem subjects was above the scale mid-point. Thus "low" self-esteem subjects have not had low self-esteem in the sense of being "negative" or worse than neutral; their self-esteem was "positive". These studies have therefore tested whether levels of self-esteem are negatively correlated with subsequent levels of in-group favouritism, regardless of whether or not the prior levels of self-esteem are "low" or "depressed" in the sense of being "negative" or "lower than neutral". Social identity theory's self-esteem hypothesis is not that the lower the self-esteem the greater the subsequent in-group favouritism, however, even as identified by Abrams &

Hogg. It is that subjects with "negative" self-esteem will discriminate more than subjects with (secure) "positive" self-esteem. Studies that compare the levels of discrimination shown by subjects with varying levels of "positive" self-esteem do not therefore test Abrams & Hogg's second corollary of social identity theory's self-esteem hypothesis.⁵

Abrams & Hogg's formulations of the corollaries of social identity theory's self-esteem hypothesis are in any case inadequate, in that social identity theory does not claim either that (i) "successful intergroup discrimination ...elevates self-esteem", or that (ii) "depressed or threatened self-esteem...promotes intergroup discrimination. Social identity theory rather claims: (i) that successful intergroup discrimination is one route to making self-esteem more positive or more secure; and, (ii) that depressed or threatened self-esteem will promote intergroup discrimination *under certain conditions*. More accurate social identity theory self-esteem hypothesis corollaries would run something like this (see Chapter 1).

⁵ It should be noted that another difficulty associated with examining self-esteem in this way is that the sort of self-esteem examined is not derived from particular group memberships. As such, a purely individual link is assumed between self-esteem and in-group favouritism, rather than a link between self-esteem derived from particular group memberships and discrimination by those group members. This entails understanding social identity theory as providing an individualistic explanation of the conditions for and consequences of intergroup discrimination: which is diametrically opposite to Tajfel's explicit intentions (Billig, 1976; Long & Spears, forthcoming; Tajfel, 1972, 1979).

1. Intergroup discrimination which *achieves* or *enhances* positive in-group distinctiveness from one or more relevant out-groups on comparison dimensions of value to in-group members will make a positive *contribution* to the *magnitude* of in-group members' social identity and self-esteem, provided that the discrimination is neither accompanied by nor results in a conflict of values. Intergroup discrimination which *protects* or *maintains* positive in-group distinctiveness from one or more relevant out-groups on comparison dimensions of value to in-group members will make a positive *contribution* to the *security* of in-group members' social identity and self-esteem, provided that the discrimination is neither accompanied by nor results in a conflict of values.
2. Depressed or threatened self-esteem will promote intergroup discrimination when: (i) the in-group is salient; (ii) the discrimination will occur against one or more relevant out-groups; (iii) the discrimination occurs on dimensions of value to in-group members; (iv) the discrimination is perceived as a viable route to secure and legitimate positive in-group distinctiveness; and, (v) discrimination is perceived as compatible with or more attractive than alternative available methods of elevating or securing positive self-esteem.

These corollaries have some obvious disadvantages as compared to the ones identified by Abrams & Hogg. Primarily, they are far longer, more complex and difficult to operationalize. Nevertheless, they more accurately reflect social identity theory's self-esteem hypothesis and they have important consequences when that hypothesis is empirically tested.

Abrams & Hogg's first corollary claims that "successful intergroup discrimination...elevates self-esteem." In fact, within social identity theory intergroup discrimination is not predicted to *elevate* self-esteem when that discrimination is pursued in order to make in-group superiority more *secure*. Nor is intergroup discrimination predicted by social identity theory to necessarily *elevate* self-esteem when it is accompanied by or results in a conflict of values for in-group members, for example when intergroup discrimination is perceived as an illegitimate route to in-group superiority. Neither is intergroup discrimination predicted by social

identity theory to elevate self-esteem when the social identity of interest is not salient for in-group members, when the intergroup discrimination occurs on dimensions of no interest or value to in-group members, or when the out-group discriminated against is not perceived as a relevant comparison group for the in-group. The "reformulated" corollary 1 above makes clear that social identity theory predicts that intergroup discrimination will only elevate self-esteem when it obtains or increases the positive in-group distinctiveness of a salient in-group from a relevant out-group on dimensions of importance to the in-group, and when the positive in-group distinctiveness is not accompanied by a conflict of values. Intergroup discrimination will therefore not reliably elevate self-esteem if: (i) the in-group is not salient; (ii) the dimension on which intergroup discrimination occurs is unimportant to in-group members; (iii) the out-group is not perceived as a relevant comparison group; (iv) the intergroup discrimination is indulged in to enhance the security of in-group superiority; or, (v) the intergroup discrimination is accompanied by a conflict of values.

Abrams & Hogg's second corollary claims that depressed or threatened self-esteem promotes intergroup discrimination. In fact, social identity theory predicts that depressed or threatened self-esteem will promote intergroup discrimination *when such intergroup discrimination is perceived as an available route to elevated or more secure self-esteem.* Depressed or threatened self-esteem is not predicted to promote intergroup discrimination when in-group membership is not salient, for example. Nor will depressed or threatened self-esteem promote intergroup discrimination when alternative means to elevated or more secure self-esteem are perceived to be more attractive and/or available (e.g. by becoming a member of group that already has positive in-group distinctiveness, or by psychologically dissociating oneself from a negatively distinct group when one's personal identity is positive). Again, depressed or threatened self-esteem will not promote

intergroup discrimination when such intergroup discrimination is perceived as unlikely to bring about the desired effects on self-esteem (e.g. when the intergroup discrimination is perceived as unlikely to result in positive in-group distinctiveness without a conflict of values: in the jargon of social identity theory, when cognitive alternatives are unavailable). The "reformulated" corollary 2 makes it clear that depressed or threatened self-esteem will only promote intergroup discrimination when it is perceived as a viable and attractive method of meeting those self-esteem needs, i.e. when it will result in positive in-group distinctiveness from a relevant out-group on a dimension of value to in-group members without a conflict of values, and when such discrimination does not preclude apparently "easier", "safer" or more "reliable" methods of elevating or securing self-esteem needs.

The importance of one aspect of these reformulated corollaries can be highlighted by re-examining Sachdev & Bourhis' (1985) "test" of corollary 2. Hogg & Abrams (1990: 35) report that "Sachdev & Bourhis found that greater power...was associated with greater discrimination" and reason that "Corollary 2 of the self-esteem hypothesis might well predict the opposite: the lower the power...the lower the social identity contingent self-esteem and thus the greater the discrimination". Such reasoning completely ignores the interaction between levels of self-esteem and the availability of intergroup discrimination to meet self-esteem needs (i.e. power: cf. Ellemers et al., 1990, 1992; Ng, 1982; Reicher, 1987; Reicher & Levine, 1994). In Sachdev & Bourhis' (1985) study the subjects with (or assumed to have) the lowest self-esteem (and therefore the most to gain by *successful* intergroup discrimination) also had no power to address their self-esteem needs via in-group favouritism as such discrimination could not be successful: all it could do would be to antagonize the all-powerful out-group. That is, these subjects had no "cognitive alternatives" to their inferior position and although their

social identity contingent self-esteem was low, social identity theory predicts that they would not engage in intergroup discrimination. At the other extreme, subjects in Sachdev & Bourhis (1985) with (or assumed to have) the highest self-esteem also had total power. They therefore needed to engage in only slight in-group favouritism in order to obtain positive and secure in-group distinctiveness on the dimension of that discrimination (i.e. points) and thereby positive contributions to social identity contingent self-esteem. Social identity theory therefore predicts that such subjects would display some in-group favouritism (as they were not yet positively distinct in terms of the dimension of comparison) but also that they would display less discrimination than subjects whose guarantee of obtaining positive in-group distinctiveness was less assured (i.e. subjects with only equal or high power). And they did. In between these two extremes were subjects with low, equal or high power. Social identity theory predicts that all of these subjects would have social identity contingent self-esteem needs (as their in-groups were not yet positively distinct on the dimension of value, i.e. points) and that they would address those needs via in-group favouritism to the extent that they perceived such discrimination as likely to succeed. Thus, social identity theory predicts that the greater the non-total power (with equal self-esteem needs) the greater the discrimination, as the greater the non-total power the greater the likelihood that the discrimination would be successful (i.e. in achieving positive in-group distinctiveness on a dimension of value to the in-group without antagonizing a more powerful out-group). This is exactly what Sachdev & Bourhis found. Sachdev & Bourhis' (1985) study therefore provides support for corollary 2 of social identity theory's self-esteem hypothesis (in its reformulated form) rather than disconfirming it as argued by Abrams & Hogg.

A final problem with tests of social identity theory's self-esteem hypothesis concerns the nature of the self-esteem

measures used. Self-esteem is the result of personally evaluating the contents of one's self-concept. The self-concept can be thought of a list of self-descriptions which are salient within a given period of time. Self-descriptions, or "identities", can be "personal" (e.g. My name is Tom) or "social" (e.g. I am male). Thus, an individual's self-esteem at any given moment will be the overall result of them evaluating each salient particular identity within their self-concept. So, for example, a person who had just come first in a running-race but whose team had lost might feel good about the former (positive personal identity of oneself as a runner) but bad about the latter (negative social identity of oneself as a member of a losing team) immediately after the race, and thus might have neutral or ambivalent "overall" self-esteem at that moment. The important point being that particular social identities only make contributions to "overall" individual self-esteem. At any given time the contributions particular social identities make to an individuals' self-esteem will depend on: (i) how "much" of their self-concept is taken up by those particular social identities (i.e. as a proportion of the "overall" active self-concept); and, (ii) how "salient" the particular social identities are within the "overall" self-concept (i.e. how distinctive and/or important those identities are in relation to other salient identities).

These considerations have direct and crucial implications for measuring self-esteem when empirically testing social identity theory's self-esteem hypothesis. To see this it is useful to conceptualize self-esteem measures in terms of (i) the time period they measure self-esteem over, and (ii) the specificity of the self-esteem measured.

Self-esteem measures can be measures of "trait" or "state" self-esteem (cf. Fleming & Courtney, 1984; Heatherton & Polivy, 1991). The latter measure self-esteem at a given moment whereas the former attempt to measure relatively stable "average" or "typical" self-esteem over a period of time.

Self-esteem measures can also measure "general" or "specific" self-esteem (cf. Blascovich & Tomaka, 1991). The latter measure the self-esteem associated with a particular identity or self-attribute (e.g. oneself as a woman, oneself in terms of one's physical appearance) whereas the former attempt to measure "overall" self-esteem (i.e. the esteem resulting from evaluation of all aspects of oneself, that is, one's salient self-concept as a whole). These distinctions allow individual self-esteem measures to be categorized as being of one of four types: General Trait Self-Esteem measures (GTSE, i.e. measures of individuals' typical or average overall evaluations of themselves over time); General State Self-Esteem measures (GSSE, i.e. measures of individuals' overall evaluations of themselves at particular times); Specific Trait Self-Esteem measures (STSE, i.e. measures of individuals' typical or average evaluations of particular aspects of themselves over time); or Specific State Self-Esteem measures (SSSE, i.e. measures of individuals' evaluations of particular aspects of themselves at particular times).

Most available measures of self-esteem are "trait" measures, either of "general" self-esteem (e.g. Julian et al., 1966; Rosenberg, 1965) or of self-esteem specific to particular facets of the self (e.g. Crocker & Luhtanen, 1990; Crocker et al., 1993; Luhtanen & Crocker, 1991, 1992). "Trait" self-esteem measures are identifiable mainly by their emphasis on the stability over those measures over time (e.g. as indicated by test-retest reliability). On the face of it, such measures are of no use at all when testing social identity theory's self-esteem hypothesis. Corollary 1 (either Abrams & Hogg's "traditional" formulation or the "reformulated" version above) posits *changes* in the level of self-esteem across time (i.e. before and after certain instances of intergroup discrimination), and self-esteem measures which have as a principal strength their relative invariance across time and situations are unlikely to detect such changes.

"Trait" self-esteem measures have recently been adapted for use as "state" measures simply by asking respondents to "answer in terms of how you feel at the moment" (or such-like), either before the measure is administered or before each individual item on the measure. The validity of such adapted measures has not been thoroughly investigated but is likely to be problematic (Abrams & Hogg, 1988: 319). In large part this is simply because the wording of the items on such measures so often suggests "trait" measurement. Items on the Rosenberg (1965) scale, for example, ask respondents about their self-esteem "All in all...", "On the whole" and "At times". It is not clear whether respondents answering such questions in terms of how they are feeling "at the moment" would be giving a true "state" response or an instantaneous "trait" one.

To my knowledge no "state" self-esteem measures exist for measuring self-esteem at a given time derived from specific group-memberships. That is, there appear to be no SSSE measures which measure the self-esteem individuals derive from particular social identities at particular points in time. This is important because such measures seem on the face of it to be the most suited to investigating social identity theory's self-esteem hypothesis. Corollary 1 of that hypothesis suggests that in certain situations intergroup discrimination will raise that part of self-esteem which is determined by membership of the discriminating in-group. As particular social identities are rarely likely to fully determine "overall" self-esteem, "global" measures of self-esteem, even "state" ones, are likely to be less sensitive to predicted changes in social identity contingent self-esteem than ones which attempt to directly measure that part of state self-esteem determined by the particular social identity of interest. A person may suffer depressed self-esteem as a result of their experimental group failing to beat another experimental group at some relatively trivial task, for example, but unless that person's overall self-concept is

totally or largely determined by their experimental group membership, "general" measures of self-esteem are unlikely to detect such changes. The depressed self-esteem seems far more likely to be captured by a measure which concentrates on that part of "overall" self-esteem which is determined by the experimental group membership at particular times.

Similar ideas apply to corollary 2 of social identity theory's self-esteem hypothesis. In-group favouritism is one response to low or threatened self-esteem derived from membership of inferior or insecurely superior groups. The discrimination is used in the hope that it will achieve in-group superiority or more secure in-group superiority. In the former instance successful in-group favouritism (i.e. that which achieves valued superiority for a salient in-group from a relevant out-group without a conflict of values) is predicted to elevate that part of self-esteem which is dependent on or determined by membership of the in-group. As with corollary 1, "state" "general" measures of self-esteem are only likely to detect such change (i.e. self-esteem elevation) when the particular social identity of interest totally or predominantly occupies group members' self-concepts. The elevated self-esteem seems more likely to be captured by a "state" measure which focuses directly on that part of "overall" self-esteem which is dependent upon the particular social identity of interest.

To summarize so far, Abrams & Hogg's formulations of the corollaries of social identity theory's self-esteem hypothesis have not yet been adequately tested. For corollary 1 this is in large part because the link between in-group favouritism and subsequent self-esteem has been investigated rather than the link between *successful* intergroup discrimination and subsequent self-esteem. For corollary 2 this is in large part because the link between low self-esteem and subsequent in-group favouritism has been investigated rather than the link between *low or threatened* self-esteem and subsequent intergroup discrimination. Second, Abrams & Hogg's

formulations of the corollaries of the self-esteem hypothesis within social identity theory are inadequate. Their corollary 1 neglects to adequately specify the nature of successful intergroup discrimination and ignores the importance of in-group salience, out-group relevance, the value of the dimension upon which discrimination occurs, and the role of conflicts of value involved in some instances of intergroup discrimination. Their corollary 2 neglects the role of perceptions of cognitive alternatives (i.e. whether intergroup discrimination is likely to be successful) and also ignores the importance of in-group salience, out-group relevance, the value of the dimension upon which discrimination may occur, and the availability of alternative or additional routes to improved self-esteem. Third, no tests of social identity theory's self-esteem hypothesis seem to have employed the most appropriate type of self-esteem measure: one which measures self-esteem derived from membership of a particular social group at a particular moment in time. With respect to corollary 1 such measures appear best able to capture changes in self-esteem resulting from changed in-group circumstances and with respect to corollary 2 such measures seem best able to predict intergroup discrimination intended to improve the situation of particular in-groups at particular times (and therefore the social identity and that part of self-esteem which is of interest).

It is illuminating to re-consider one of the tests of corollary 2 in the light of the above arguments. Wagner et al. (1986: 19) found that levels of general self-esteem did not significantly correlate with subsequent discrimination across groups, but that levels of self-esteem contingent upon subjects' perceived achievement at university did. The lower subjects' university achievement specific self-esteem the greater their derogation of out-group members' "discussion ability". Thus, while low general self-esteem did not result in greater discrimination across groups than high general self-esteem, thus proving no support for corollary 2 of social

identity theory's self-esteem hypothesis, low levels of specific "university achievement" self-esteem did result in greater out-group derogation on a dimension of "discussion ability" than high levels of such specific self-esteem. If it can be assumed that subjects with high "university achievement" self-esteem were secure in their evaluations of their own achievement at university this latter result supports corollary 2. The important point for present purposes is that a measure of general self-esteem was not sufficiently sensitive to "predict" subsequent discrimination across groups on a particular dimension but a measure specific to an aspect of self-esteem related to that dimension was.⁶

In the study reported here an attempt was made to develop four "equivalent" self-esteem measures, one for each "type" identified above, i.e. GTSE, GSSE, STSE and SSSE, with the "specific" aspects of the relevant measures referring to self-esteem derived from evaluation of particular group memberships (i.e. social identities). Each of these measures were then used to try and detect changes in self-esteem resulting from manipulations intended to raise or lower subjects' evaluations of a particular in-group (national) membership. Consistent with the arguments developed above, hypothesis 1 was that the SSSE measure would be significantly more sensitive to such changes than the other three measures.

A fifth measure was developed in the present study which was intended to tap evaluation of the in-group at particular points in time, i.e. a General State In-group Evaluation (GSIE) measure. This was done to enable a manipulation check to be made as to whether or not subjects' evaluations of their in-group did in fact alter as a result of the experimental manipulation intended to bring about such a change. This manipulation check can also be considered to be a test of

⁶ If it can be assumed that achievement at university was perceived by subjects to be related to or correlated with discussion ability.

social identity theory's prediction that group members' social identity contingent self-esteem is determined by those group members' evaluations of their in-group.

A final measure employed in the present study was a measure of the extent to which subjects identified with the (national) in-group of interest. Hypothesis 2, in line with social identity theory, predicted that the greater a subject's social identity, the greater the effects of in-group evaluation on that subject's self-esteem.

METHOD

Participants

246 undergraduate psychology students from either the London School of Economics and Political Science or Royal Holloway and Bedford College filled out usable response booklets. Of these 104 were male, 141 were female, and one additional respondent did not indicate their sex. A further 10 subjects returned spoiled or incomplete booklets. Approximately 50 students either declined to take a booklet or returned completely blank booklets.

Materials

Rosenberg's (1965) self-esteem scale (RSE) was used as a template to produce four separate self-esteem questionnaires and one in-group evaluation questionnaire. The Rosenberg self-esteem measure was chosen because of its widespread use in social-psychological research (Blascovich & Tomaka, 1991) and because of the relative ease with which it could be adapted for present purposes.

The questions of the original Rosenberg scale were modified only slightly to produce the first scale, a measure of general trait self-esteem (GTSE).⁷ A measure of general state self-esteem (GSSE) was produced by simply starting each question used in the GTSE scale with the phrase "AT THE MOMENT,...". A measure of specific trait self-esteem (STSE) was produced by starting each question used in the GTSE scale with the phrase "I am a member of this group and, AS SUCH...". A measure of specific state self-esteem (SSSE) was produced by starting each question used in the GTSE scale with the phrase "I am a member of this group and AS SUCH, AT THE MOMENT...". Finally, a general state in-group evaluation (GSIE) scale was produced by starting each question used in the GTSE scale with the phrase "AT THE MOMENT" and changing the object of esteem from the self ("I") to the in-group ("this group"). All questions on all scales were to be answered on 7-point scales running from "strongly agree" to "strongly disagree". Negative and positive items were initially randomly ordered and the same order was used on each version of the scale. The order of the questions was different on the first and second presentations of each scale (see below), but the order of positive and negative questions was kept the same.

10 four-page response booklets were made up. Each started with a brief introduction assuring anonymity and confidentiality. Respondents were then asked to indicate their sex and their national group. The latter served as each subject's in-group. Also on the front page of the booklet there was a three-question social identity scale measuring the strength, valence and centrality of national group membership in the respondent's self-concept (Fraser, 1991; Hofman, 1988).

⁷ A pilot study revealed that respondents disliked and reacted negatively to certain words and phrases used in the original RSE. In addition, slight word changes were necessary to make the scale as similar as possible to the other versions of the scale used in this study.

The second page of each booklet contained one of the five esteem scales mentioned above. Instructions were slightly different for each scale. Respondents receiving the GTSE scale were told to answer each question in terms of how they "USUALLY" felt about themselves, FOR WHATEVER REASON. Respondents receiving the GSSE or the GSIE scales were instructed to answer each question in terms of how they felt about themselves "RIGHT NOW", FOR WHATEVER REASON. Respondents receiving the STSE scale were told to respond in terms of how they "USUALLY" felt about themselves IN TERMS OF YOUR MEMBERSHIP OF YOUR NATIONAL GROUP. Finally, subjects receiving the SSSE scale were told to answer in terms of how they felt about themselves "RIGHT NOW", IN TERMS OF YOUR MEMBERSHIP OF YOUR NATIONAL GROUP. These instructions were in addition to any prefixes to each question mentioned above.

Page three of each response booklet differed according to the self-esteem manipulation, positive or negative. Apart from the single word used to achieve this manipulation, all respondents received the same instructions:-

Obviously there will be both good and bad aspects, but what sort of things make you feel POSITIVE (NEGATIVE) about your national group and/or your membership of it? Please spend about two minutes answering this question (and/or giving examples) in the space below.

Page four of each booklet contained a second copy of the scale each respondent received on page 2 of their booklet. On this occasion all respondents received the same instructions, which were to answer each question as honestly as possible without looking back to previous answers or trying particularly to give answers that they might think "consistent" or "expected".

All of the pages used in each of the response booklets can be found in Appendix 10.

Procedure

Response booklets were randomly distributed in three sessions at the beginning of a practical workshop (Royal Holloway), or at the beginning (LSE) or the end of a lecture (LSE). Respondents were given 12-15 minutes to complete the booklets, which were immediately collected in. Respondents were thanked for their participation.

RESULTS

Scale scoring

Each question on each scale scored 1 for "strongly agree" increasing one point at a time up to 7 for "strongly disagree". Scores were then reversed for all "pro-trait" items. Mean scores over all ten items on a given scale were then calculated. Thus, mean scores were calculated for each scale and mean scores ranged from 1 (low mean self-esteem or in-group evaluation) to 7 (high mean self-esteem or in-group evaluation).

Calculating esteem and in-group evaluation change scores

Self-esteem scores obtained on the second presentation of each measure were subtracted from self-esteem scores obtained on the first presentations of those scales. A positive change score therefore indicates a rise in self-esteem and a negative change score indicates a fall in self-esteem as a result of the experimental manipulation. Similar change scores were calculated for in-group evaluations.

Also, a single measure was calculated to indicate the mean change in self-esteem *in the predicted direction* as a result of the experimental manipulation, regardless of whether the manipulation was positive or negative. This MPC (mean

predicted change) figure is positive if the mean self-esteem change was in the direction predicted by the manipulation and negative if it is in the direction opposite to that predicted by the manipulation.

Comparable scores were calculated for in-group evaluation.

Scale reliability

Reliability coefficients were calculated for each presentation of each scale. These ranged from a low of 0.82 (first presentation of SSSE) to a high of 0.96 (second presentation of GSIE).⁸ The reliability coefficient for the social identity score was 0.86.

Initial self-esteem and in-group evaluation scores

Initial mean self-esteem scores, i.e. before the experimental manipulation, were as follows: GTSE1 = 5.64 (SD = 1.11, n = 47); GSSE1 = 5.79 (SD = 0.92, n = 52); STSE1 = 5.65 (SD = 0.84, n = 45); SSSE1 = 5.81 (SD = 0.72, n = 48). Initial mean evaluation of the national in-group was similarly positive at 5.22 (SD = 1.17), n = 46).

Manipulation checks

Subjects listed a mean of 2.78 items which made them feel positive or negative about their national group and/or membership of it. 29 subjects (11.8%) listed no items, 45 (18.3%) listed only 1 item, 53 (21.5%) listed 2 items, and almost half of the subjects (119, 48.4%) listed two or more

⁸ In full: GTSE1=0.94, GSSE1=0.92, STSE1=0.85, SSSE1=0.82, GSIE1=0.93, GTSE2=0.95, GSSE2=0.93, STSE2=0.88, SSSE2=0.83, GSIE2=0.96.

items. Each item listed was evaluated by the experimenter on a scale from 1 - 5, where 1 = strongly in direction opposite to that implied by the manipulation and 5 = strongly in agreement with direction implied by the manipulation. The mean score was 3.99 (SD = 0.73), indicating that, on average, subjects listed items which were in accordance with the manipulation. Only 8 items (3.5%) were in a direction opposite to the manipulation, 34 (14.8%) were neutral, and 187 (81.7%) were in the direction suggested by the manipulation. Thus, the vast majority of items listed were in accordance with the manipulation (i.e. were positive for the positive manipulation and negative for the negative manipulation), with a very small minority of items being in the opposite direction to that suggested by the manipulation. There were no differences in the number of items listed for the positive or the negative manipulation and neither was there any difference across manipulations in the extent to which items listed were consistent with those manipulations.

GSIE2 - GSIE1 provides a measure of how evaluation of the in-group altered as a result of the experimental manipulation. Evaluation of the in-group became more positive as a result of the positive manipulation (+0.15, SD = 0.42) and more negative as a result of the negative manipulation (-0.05, SD = 0.34). Neither of these changes were significantly greater than zero by one-tailed paired t-tests, however, although the positive change approached significance ($t_{21} = 1.62, p = 0.06$). Mean predicted change (MPC) for in-group evaluations (i.e. positive when the manipulation was positive and negative when the manipulation was negative) was +0.10 (SD = 0.38). This was only just non-significantly different from zero at the 5% level by a one-tailed one-sample t-test ($t_{43} = 1.66, p = 0.053$). Thus, mean evaluations of the in-group were in the directions predicted by the experimental manipulations and although neither effect was strong, the effect of the positive manipulation was stronger than the effect of the negative manipulation, which was insignificant.

Self-esteem changes

The table below shows the mean changes in self-esteem resulting from the experimental manipulation by both scale type and the direction of the manipulation (with standard deviations in brackets).

	GTSE	GSSE	STSE	SSSE
+	.05 (.19)	.12 (.27)	.10 (.55)	.14 (.28)
-	.11 (.31)	.04 (.40)	.05 (.40)	-.21 (.48)
MPC	-.03 (.27)	.04 (.35)	.03 (.35)	.18 (.38)

+/- = positive or negative manipulation of in-group evaluation; MPC = mean predicted change; GTSE = General Trait Self-Esteem; GSSE = General State Self-Esteem; STSE = (in-group) Specific Trait Self-Esteem; SSSE = (in-group) Specific State Self-Esteem.

Table 5.1: Mean self-esteem changes by scale-type and manipulation

The table above suggests that all of the measures indicated varying degrees of self-esteem improvement as a result of the manipulation intended to improve in-group evaluation, but only the SSSE measure indicated a fall in self-esteem as a result of the manipulation intended to worsen in-group evaluation. A 4 (scale type) X 2 (manipulation) ANOVA with esteem change as the dependent variable revealed no main effect of scale type, a significant main effect of manipulation $F_{(1,180)} = 4.02$, $p = 0.046$, and a very nearly significant interaction effect ($F_{(3,180)} = 2.58$, $p = 0.055$). Separate 1-WAY ANOVAs for each manipulation revealed that there were no significant differences between the measures for the positive manipulation ($F_{(3,92)} = 0.31$, $p = 0.82$) but there were for the negative manipulation ($F_{(3,92)} = 2.80$, $p = 0.045$). Similarly, planned contrasts between the SSSE measure and the other three measures were not significant for the positive manipulation but were for the negative manipulation ($T_{(88)} = 2.82$, $p =$

0.006). Additionally, planned comparisons between the two "state" measures (i.e. GSSE and SSSE) versus the two "trait" measures (i.e. GTSE and STSE) were not significant for the positive manipulation but very nearly were for the negative one ($T_{(88)} = 1.98$, $p = 0.051$). Unplanned comparisons (Duncan Tests) showed no significant differences between the measures for the positive manipulation but revealed that the SSSE measure was significantly different from each of the other three measures for the negative manipulation.

A 1-WAY ANOVA also revealed an almost significant difference between the measures on MPC (mean predicted change) across the manipulations ($F_{(3,184)} = 2.51$, $p = 0.060$). Planned comparisons revealed that with respect to MPC the SSSE measure was significantly different to the other three measures ($T_{(184)} = -2.58$, $p = 0.01$) and that the two state measures were significantly different to the two trait ones ($T_{(184)} = -2.04$, $p = 0.043$). An unplanned comparison indicated only that the SSSE measure was significantly different to the GTSE one.

Finally, a series of one-tailed one-sample t-tests revealed significant changes in self-esteem in predicted directions as a result of the experimental manipulations only for GSSE and SSSE on the positive manipulation ($t_{25} = 2.23$, $p = 0.04$ and $t_{24} = 2.60$, $p = 0.02$, respectively), for SSSE on the negative manipulation ($t_{21} = -2.09$, $p = 0.02$), and for SSSE for MPC ($t_{46} = 3.17$, $p = 0.002$).

It seems, therefore, that in the present study there was very little difference between the measures with regards to detecting self-esteem changes resulting from the positive manipulation (although only the two state measures showed significant levels of self-esteem improvement resulting from that manipulation), but that the SSSE scale was more sensitive than the other measures to self-esteem deterioration resulting from the negative manipulation.

The effects of social identity

Social identity scores were used to divide subjects into categories of either relatively low social identity or relatively high social identity. The former ($n = 117$, 47.6% of all subjects) had a mean social identity score of 3.57 ($SD = 0.67$) and the latter ($n = 129$, 52.4% of all subjects) had a mean social identity score of 5.62 ($SD = 0.70$). Thus, "low" social identity students had an approximately neutral mean social identity and the "high" social identity students had a reasonably positive mean social identity.

A between-groups t-test revealed no significant differences in MPC scores for in-group evaluation (i.e. GSIE) between subjects with "low" and subjects with "high" social identity. Further, a 2 X 2 ANOVA revealed no significant main or interaction effects on in-group evaluation changes according to the experimental manipulation (positive or negative) and/or the level of subjects' social identity ("low" or "high"). Thus, the social identity levels do not seem to have influenced the degree to which the experimental manipulation affected subjects in-group evaluations.

A similar set of results were obtained with the self-esteem measures. A 4 (scale type) X 2 (Social identity level) ANOVA with MPC as the dependent variable did not result in a significant main effect of social identity or a significant interaction effect.⁹ Similarly, a 4 (scale type) X 2 (manipulation) X 2 (social identity level) ANOVA with self-esteem change as the dependent variable did not result in a significant social identity main effect or any significant

⁹ There was an almost significant main effect of scale-type ($F_{(3,180)} = 2.498$, $p = 0.06$), as expected from the results reported above.

interaction effects involving social identity levels.¹⁰ Thus, as before, the SSSE scale obtained somewhat better MPC scores than the other three scales; the positive manipulation resulted in slight self-esteem improvement while the negative manipulation resulted in slight self-esteem deterioration (averaged over scale types); and the SSSE scale was better than the other scale types at detecting self-esteem deterioration as a result of the negative manipulation. For present purposes, though, the important result is that social identity levels had no effect on self-esteem change, whatever the manipulation or the self-esteem scale used.

Non-parametric analysis

The analyses above were supplemented by a series of Kruskal-Wallis 1-WAY ANOVAs comparing self-esteem and in-group evaluation changes between scale-types for each manipulation and by a series of Wilcoxon tests comparing pre- and post-manipulation levels of self-esteem between scale-types. This supplementation was done for a number of reasons. First, the assumptions necessary to employ parametric statistics may not have been met in the present study. In particular, self-esteem tends to be negatively skewed and the measures of self-esteem employed used scales which are probably more properly thought to be ordinal rather than interval.¹¹ Second, relatively small numbers of subjects fell into each cell of some of the more complex ANOVAs conducted above. Third, non-parametric statistics have the advantage of revealing how many

¹⁰ Again, the former result was expected from a previous analysis, as were the near significant main effect of manipulation ($F_{(1,172)} = 3.581, p = 0.060$) and interaction effect of scale-type by manipulation ($F_{(3,172)} = 2.450, p = 0.065$).

¹¹ All pre-manipulation self-esteem scores were negatively skewed in the present study: GTSE = -1.756, GSSE = -0.972, STSE = -0.327, SSSE = -0.296. Pre-manipulation in-group evaluation scores were also negatively skewed, GSIE = -0.921.

subjects' self-esteem changes were in the directions predicted by the manipulation, rather than relying on differences between average self-esteem changes within conditions.¹²

The table below shows the results of the separate Wilcoxon tests comparing pre- and post-manipulation self-esteem scores for each scale-type. The table also shows Wilcoxon tests comparing pre- and post-manipulation evaluations of the in-group. The latter reveals a significant improvement in subjects' evaluation of the in-group following the positive manipulation but no significant deterioration in their in-group evaluation following the negative manipulation, relative to pre-manipulation levels. Overall, evaluation of the in-group changed in the direction predicted by the manipulations, with this change only just failing to reach significance.

A Kruskal-Wallis test revealed that there were significant differences between the measures in terms of their indication of mean predicted self-esteem change as a result of the two manipulations considered together ($\chi^2 = 9.04$, $p = 0.03$). Kruskal-Wallis tests carried out separately for each of the manipulations, however, revealed no significant differences between the scale-types in indicating self-esteem change, although the test carried out on the negative manipulation approached significance ($\chi^2 = 2.77$, $p = 0.43$ for the positive manipulation and $\chi^2 = 6.40$, $p = 0.09$ for the negative manipulation).

¹² The parametric tests were run because the tests used are relatively robust (i.e. remain valid despite minor violations of the assumptions needed to run those tests) and allowed more complex analysis than is possible with non-parametric tests on SPSS.

Scale	Manip	Changes in predicted direction	Changes in opposite direction	No change	Z	p (1-Tail)
GTSE	+	9	6	7	1.22	0.111
	-	5	15	3	1.51	0.655
	Both	14	21	10	0.52	0.303
GSSE	+	9	6	11	1.79	0.037
	-	10	11	4	0.47	0.320
	Both	19	17	15	0.93	0.177
STSE	+	13	6	4	1.05	0.148
	-	7	11	4	0.57	0.286
	Both	20	17	8	0.25	0.402
SSSE	+	17	5	3	2.37	0.009
	-	13	6	3	1.79	0.037
	Both	30	11	6	2.86	0.002
GSIE	+	11	5	2	1.86	0.032
	-	9	11	2	0.41	0.341
	Both	20	16	8	1.62	0.053

Table 5.2: Number of changes in self-esteem in predicted or unpredicted directions by scale-type and manipulation

The table above makes it clear that neither of the "trait" self-esteem measures (i.e. GTSE and STSE) detected significant changes in self-esteem as a result of either the positive or the negative manipulation, or overall. The GSSE measure shows a significant increase self-esteem following the positive manipulation, no change in self-esteem following the negative manipulation, and no change in self-esteem (in the directions predicted by the manipulations) when both manipulations are considered together. The SSSE measure shows a significant improvement in self-esteem as a result of the positive manipulation, a significant deterioration in self-esteem as a result of the negative manipulation, and a significant change

in self-esteem (in the directions predicted by the manipulations) when both manipulations are considered together. Additionally, the SSSE measure shows far more consistency of self-esteem change resulting from the experimental manipulation than any of the other measures. That is, two or three times more subjects show self-esteem change in the direction predicted by the manipulation than in the opposite direction on the SSSE measure, a proportion far higher than for any other measure. Similarly, far fewer subjects show no self-esteem change following experimental manipulations on the SSSE scale than on any of the other scales.

Repeating the above analyses separately for subjects with "low" and for subjects with "high" social identity resulted in outcomes similar to the above and to each other. That is, as with the parametric tests, no significant differences were found in subjects' changes in self-esteem as a result of the experimental manipulations according to their levels of social identity.

DISCUSSION

The results above suggest that evaluation of the in-group improved as a result of the manipulation designed to have this effect but that evaluation of the in-group did not deteriorate as a result of the manipulation designed to have that effect. With regards to self-esteem, the two "trait" measures (GTSE and STSE) indicated no change in self-esteem as a result of the experimental manipulations (positive, negative or in combination), the GSSE measure indicated self-esteem improvement as a result of the positive manipulation but no self-esteem deterioration as a result of the negative manipulation and no "overall" self-esteem change when the two manipulations were considered in combination, and the SSSE

scale detected both self-esteem improvement as a result of the positive manipulation and self-esteem deterioration as a result of the negative manipulation, and therefore also indicated significant self-esteem change in the directions suggested by the manipulations when both manipulations were considered together. The SSSE measure was also more "reliable" or "consistent" than the other measures in that more subjects' self-esteem changes were in the direction suggested by the experimental manipulation than was the case for the other scales, relative to the number of subjects who showed no self-esteem change or showed self-esteem change in the direction opposite to that suggested by the experimental manipulation. Thus, hypothesis 1 received support in that the SSSE measure was more sensitive than other measures to self-esteem changes resulting from manipulations intended to raise or lower subjects' evaluations of their (national) in-groups, particularly for the manipulation intended to worsen in-group evaluation. Hypothesis 2 received no support in that there were no differences in self-esteem change as indicated by any of the measures according to whether subjects had relatively high or relatively low levels of (national) social identity.

The problem with this first conclusion is of course that although SSSE fell as a result of the negative manipulation, GSIE did not. SSSE cannot sensibly be claimed to be sensitive to changes in evaluations of the in-group if SSSE changes when in-group evaluations do not. There are three possible explanations for the mismatch found between GSIE and SSSE as a result of the negative manipulation. The first is that in-group evaluation was indeed unaffected by the negative manipulation and that social identity contingent self-esteem was indeed adversely affected. The second is that evaluation of particular attributes of the in-group were adversely affected by the negative manipulation but that the "general" in-group evaluation measure was not sensitive enough to detect these "specific" changes. The third is that evaluation of particular attributes of the in-group were adversely affected

by the negative manipulation but that the threat to social identity and self-esteem caused by this exercise led subjects to spontaneously recall and/or generate positive in-group attributes with the net result of no change in overall in-group evaluation. Each possible explanation will be considered in turn.

The first essentially claims that each measure was equally valid and the results indicated by them can be taken at face value. The challenge then is to explain why social identity contingent self-esteem fell as a result of the negative manipulation even though evaluation of the in-group itself did not. This is inconsistent with social identity theory which says that evaluation of oneself as an in-group member is wholly determined by evaluation of the in-group itself. However, the present results suggest that situations in which negative in-group attributes are made salient to people may not cause those people to think negatively about the in-group per se, or even about all in-group members, but rather may cause them to think negatively about their own membership of their group. So, for example, English football hooliganism (antagonism to Europe, or any other negatively evaluated in-group attribute or behaviour) may not make the English feel bad about England as such, or even about English people, but it may make them feel bad about themselves as English people, perhaps because of "reflected appraisal", i.e. a suspicion that negatively evaluated in-group attributes make others think negatively and stereotypically about in-group and about in-group members, and therefore about them as particular members of the in-group. Such a possibility suggests that research is needed to investigate the exact relationship between evaluation of the in-group (e.g. the police), evaluation of in-group members (e.g. policemen and policewomen) and evaluation of the self as an in-group member (e.g. myself as a policeman).

The second possible explanation for the mismatch found between changes in SSSE and GSIE as a result of the negative manipulation essentially claims that the latter measure was not as valid as the former in the present study. This argument rests upon ideas similar to those expounded in the introduction to this chapter. The claim is that the negative manipulation did result in depressed in-group evaluations, but only for specific in-group attributes (i.e. the ones the subjects thought of and listed), and that the general measure of in-group evaluation used was not sensitive enough to capture these changes in *specific* aspects of in-group evaluation. Evaluation of one's national identity can be assumed to be dependent upon many, many evaluations of specific elements of one's nation and membership of it, and relatively small changes in relatively minor specific national attributes are unlikely to affect general or overall national evaluations. The conclusion of this argument is that just as self-esteem measures should be tailored to the temporality and the generality of the self-esteem of interest, so too should measures of other psychological states such as in-group evaluation. The present study employed a general state in-group evaluation (GSIE) measure when perhaps it should have employed a specific state in-group evaluation (SSIE) one, focusing on those specific positive or negative aspects of the in-group which were made salient to subjects.¹³ This suggests that future studies could manipulate evaluations of identified specific in-group attributes and tailor their in-group evaluation measures to evaluation of these specific attributes. Additionally or alternatively, future studies

¹³ This was not possible in the present study because subjects were asked to think of their own examples of things which made them feel good or bad about their national groups and/or membership of them. This method was chosen because of the numerous nationalities of the respondents and an inability on my part to think of positive and/or negative national attributes which would apply equally to the nationalities of all the subjects from such a heterogeneous group. It was also hoped that self-accessed attributes would have more "impact" than experimenter-generated ones.

could manipulation overall in-group evaluations (or take advantage of "naturally occurring" events which could be assumed to cause changes in overall in-group evaluations) and employ GSIE measures similar to the one employed in this study. That is, future studies could tailor the specificity or generality of their in-group evaluation measures to the generality or specificity of the in-group evaluations of interest.

Two problems remain with this second response to the GSIE/SSSE mismatch, however. First, even if it is accepted that changes did occur to specific in-group evaluations (but that the general in-group evaluation measure was not sufficiently sensitive to detect them), this still does not explain why changes *did* occur to social identity contingent self-esteem. The SSSE measure is specific to national in-group membership, not to specific aspects of that membership. That is, if the negative manipulation was not sufficiently strong to cause a change in overall in-group evaluation, why was it strong enough to cause a change in the self-esteem of in-group members which was determined by ("overall") in-group membership? Second, the GSIE measure was sufficiently sensitive to capture improvements to overall in-group evaluation resulting from the positive manipulation, although these improvements were presumably every bit as specific as those caused by the negative manipulation, especially as there were no differences in the number or the "appropriateness" (i.e. the manipulation-consistent "direction") of the items listed for each of the manipulations.

The third potential explanation for the mismatch found between changes in SSSE and GSIE as a result of the negative manipulation concerns the possibility that the negative manipulation resulted in psychological processes which were not engendered by the positive manipulation. The idea here, consistent with social identity theory's self-esteem hypothesis, is that making negative in-group attributes

salient threatened subjects' social identities and that subjects responded to this threat by spontaneously and privately making positive in-group attributes more salient. Thus the net effect on general in-group evaluation was zero. This potential solution avoids the difficulty of the previous one caused by the asymmetry between the positive and the negative manipulations. There is no reason to assume that subjects asked to make positive in-group attributes more salient would have been motivated to "counter" the positive effects this had on in-group evaluations (and on social identity contingent self-esteem) by spontaneously making negative in-group attributes more salient. It shares with that proposed solution, however, the task of explaining why the zero net change in general in-group evaluation as a result of the negative manipulation was not accompanied by a similar zero net change at the level of social identity contingent self-esteem. After all, if the overall result of the negative manipulation on in-group evaluation was zero, why should subjects feel worse about themselves as in-group members?

Here again the possibility arises that there may not be 1:1 correspondence between evaluation of the in-group, evaluation of in-group members and evaluation of the self as an in-group member. Perhaps it was the case that the negative in-group attributes were more salient than the positive ones at the level of social identity contingent self-esteem but not at the level of the in-group as a whole. At the level of the group subjects may have been able to effectively counter salient negative in-group attributes by making equivalent positive in-group attributes salient, but at the level of themselves as in-group members perhaps this was not so easily achieved. The latter is, after all, potentially far more relevant to their *self-concepts*. Evaluations of the group are just that, evaluations of a group of which each subject is only a small part. Evaluations of the *self* as an in-group member is, almost by definition, far more self-relevant. One can distance oneself from others' evaluation of an in-group ("The

English do tend toward ethnocentrism, but I'm not ethnocentric") but it is perhaps more difficult to distance oneself from others' evaluations of oneself as an in-group member ("The English do tend toward ethnocentrism and I am English").¹⁴ This is all highly speculative, of course, but it seems worthy of investigation nonetheless. Future studies could investigate if (and when) making negative in-group attributes salient results in spontaneous recall and/or generation of positive in-group attributes (and indeed if making positive in-group attributes salient ever results in spontaneous recall or generation of negative in-group attributes). It could also be investigated whether negative in-group attributes are more threatening at the level of individual group membership than at the level of in-group evaluation (and/or whether positive in-group attributes are more beneficial at the level of individual group membership than at the level of in-group evaluation).

The second main finding in the present study was that social identity did not interact with the effects of in-group evaluation manipulation on self-esteem changes. This is problematic for social identity theory which suggests that the stronger people's social identity the more their self-esteem should be affected by changes in in-group evaluations. One potential reason for the failure to support this assertion in the present study could be that the social identity scale employed was invalid. A more likely reason is simply that most subjects had relatively positive national social identities. Only 26% of subjects had a social identity of less than the scale midpoint of 4, and only 6% had social identities less than 3 on this (7-point) scale. Thus, it appears that the vast majority of subjects had some sort of positive national social identity and therefore differential

¹⁴ If this is correct it suggests that social identity contingent self-esteem (SSSE) measures may be more reliable indicators of the negative aspects of group evaluations than are more direct in-group evaluation measures.

effects on self-esteem change as a result of in-group evaluation change were minimal. Future research could investigate whether the social identity contingent self-esteem of people with "genuinely" low social identity (i.e. "negative" social identity, significantly below the scale midpoint) is differentially affected by changes in in-group evaluations relative to the social identity contingent self-esteem of people with "genuinely" high (i.e. "positive") social identity.

It is important to be explicit about the relevance of the present results to social identity theory's self-esteem hypothesis. This hypothesis has not been tested at all by the present study (although it will be in Chapter 7.) All that has been shown is that the self-esteem measure which is most sensitive to changes in in-group evaluation is the one which measures social identity contingent self-esteem. This is highly relevant to the self-esteem hypothesis, however, as to my knowledge no tests of that hypothesis have employed this most appropriate measure. Thus, previous empirical studies which have found no self-esteem change as a result of successful intergroup discrimination cannot be counted as evidence against corollary 1 (even if they had adequately conceptualized and operationalized *successful* intergroup discrimination) because the measures they have employed may not have been sensitive enough to detect such "specific" and "state" (i.e. transient) self-esteem changes.¹⁵ Similarly, previous empirical studies which have failed to find that self-esteem needs promote intergroup discrimination cannot be counted as evidence against corollary 2 (even if they had adequately taken into account needs stemming from insecure as well as negative self-esteem derived from particular group

¹⁵ This criticism is made less severe by the fact that several studies have employed GSSE measures and the GSSE measure in the present study was as good as the SSSE measure at detecting rises in self-esteem resulting from the positive manipulation. On the other hand some of the GSSE measures used have been of very dubious validity.

memberships) because the measures employed may not have been sensitive enough to detect changes in specific (i.e. social identity contingent) self-esteem. The results of the present study suggest that SSSE measures are the best suited for use in tests of the self-esteem hypothesis and because such measures have never been used, the self-esteem hypothesis cannot be considered to have been tested at all.

Before finishing it is interesting to consider the wider implications of the results of the present study. If it true that self-esteem measures can usefully be thought of in terms of their temporality and specificity, it seems equally true that many other psychometric measures can also be thought of in similar ways. Thus, any measure which attempts to measure changes in postulated psychological states should be a "state" measure, and any measure which attempts to predict behaviours (broadly understood) dependent upon particular aspects of psychological states should be "specific" measures. Perhaps it is time to wonder how many research areas besides the self-esteem hypothesis have been blighted by employing "general" and or "trait" measures in empirical investigations when "specific" and/or "state" measures would have been more appropriate (cf. Ajzen & Fishbein's (1977) "principal of compatibility" in attitude research and Haslam et al.'s (Haslam & Turner, 1992; Haslam et al., 1992) work on the "context-dependency" of social stereotyping).

CONCLUSIONS

The sole conclusion of this chapter is simply that social identity theory's self-esteem hypothesis has not yet been adequately empirically tested. This claim can be justified methodologically, theoretically and empirically. The methodological point is primarily that the empirical "tests" of the self-esteem hypothesis within social identity theory

have simply not been adequate to test even Hogg & Abrams' impoverished formulations of the corollaries of that hypothesis. In particular, tests of corollary 1 have at best investigated the link between levels of in-group favouritism and subsequent changes in self-esteem rather than investigating the link between "successful" intergroup discrimination and subsequent changes in social identity contingent self-esteem. Tests of corollary 2 have, at best, typically investigated whether people with relatively low self-esteem engage in more subsequent discrimination across groups than people with relatively high self-esteem rather than investigating whether people with negative or threatened social identity contingent self-esteem engage in more subsequent intergroup discrimination than people with securely positive social identity contingent self-esteem. The theoretical point is that Abrams & Hogg's formulations of the corollaries of social identity theory's self-esteem hypothesis do not accurately reflect that theory. In particular, Abrams & Hogg's first corollary is deficient in not elucidating what constitutes "successful" intergroup discrimination and therefore obscures the importance within social identity theory of people obtaining or protecting positive in-group distinctiveness without a conflict of values. Their second corollary is deficient in not accounting for the fact that intergroup discrimination is by no means postulated by social identity theory as the only route available to address self-esteem needs and will not be taken if it is not deemed appropriate to serve such needs, for example if "cognitive alternatives" are not available. The empirical point is that the present study provides evidence that the most appropriate measure for use in tests of the self-esteem hypothesis is an SSSE one (where state social identity contingent self-esteem is measured) and no tests of the self-esteem hypothesis have to date employed such a measure. None of this, of course, provides a shred of evidence that the self-esteem hypothesis is or is not valid or correct. It simply illustrates that it has yet to be tested. Provisional tests of each corollary are

presented in Chapter 7. The next chapter, however, again examines the role of in-group evaluations on in-group members' self-esteem, this time in the context of comparing predictions derived from social identity theory with predictions derived from Tesser's (1988) self-evaluation maintenance model.

CHAPTER 6: STUDY 5 - THE CONSEQUENCES OF SOCIAL COMPARISON:
TESSER'S SELF-EVALUATION MAINTENANCE MODEL VERSUS
TAJFEL'S SOCIAL IDENTITY THEORY

CHAPTER OVERVIEW

The study reported in this chapter compared predictions derived from Tajfel's social identity theory (Tajfel & Turner, 1979) with predictions derived from Tesser's (1988) self-evaluation maintenance model, concerning the effects of social comparison outcomes on self-esteem and related phenomena.

Social identity theory predicts that, when social identity is salient, only *between-group* social comparison outcomes on in-group valued dimensions affect self-esteem. In such circumstances, the theory predicts that in-group superiority contributes positively to self-esteem, while in-group inferiority contributes negatively to self-esteem.

The self-evaluation maintenance model predicts that, regardless of whether or not social identity is salient, only *individual* inferiority to another affects self-esteem. Individual inferiority on self-relevant dimensions contributes negatively to self-esteem, while individual inferiority on self-irrelevant dimensions contributes positively to self-esteem. Individual superiority does not affect self-esteem, regardless of the self-relevance of comparison dimensions.

Social identity theory's predictions received considerably more support from the present study than the predictions derived from the self-evaluation maintenance model. However, between-individual and between-group social comparison outcomes on relevant dimensions had an *interactive* effect on self-esteem. It was concluded that social identity theory would benefit from incorporating a between-individual element into its currently purely between-group perspective.

INTRODUCTION

Social identity theory (e.g. Tajfel, 1978a; Tajfel & Turner, 1979) claims that when people think of themselves as group members (i.e. when they think of themselves in terms of their social identities), their self-esteem is dependent upon the outcome of comparisons between the groups they are members of (i.e. their in-groups) and relevant out-groups on dimensions of importance to the in-groups. Positive in-group distinctiveness (i.e. in-group superiority over relevant out-groups on positively valued dimensions) generally makes a positive contribution to evaluations of the in-group, of all in-group members (perceived as such), and of oneself (perceived as an in-group member), while negative in-group distinctiveness generally makes a negative contribution to evaluations of in-groups, in-group members and oneself.¹ The positive contribution to the self-esteem (i.e. self-evaluation) of in-group members from positive in-group distinctiveness will be secure if that positive in-group distinctiveness is itself secure, but will be threatened if the positive in-group distinctiveness is insecure. Thus, negative in-group distinctiveness is detrimental to the level of in-group members' self-esteem, and insecure positive in-group distinctiveness is detrimental to the security of such self-esteem.

Social identity theory goes on to say that the detrimental effects of negative or insecure positive in-group distinctiveness on in-group members' self-esteem can be countered by in-group members: (i) physically or psychologically distancing themselves from the in-group (i.e.

¹ The qualifier "generally" is required because exceptions to this "general" rule occur when status differentials are perceived to be "illegitimate" and/or so "stable" as to seem "natural" (e.g. Tajfel, 1981a: 279-287, see also Chapter 1 of this thesis). These caveats will not be considered in this chapter.

employing "social mobility" or "exit"); (ii) attempting to improve the between-group status differential in favour of the in-group (i.e. employing "social competition"); and/or, (iii) decreasing the importance of the comparison dimension to the in-group (i.e. employing "social creativity"). Similarly, the beneficial effect of secure positive in-group distinctiveness on in-group members' self-esteem can be enhanced by in-group members increasing the closeness between themselves and the in-group, increasing the between-group differential in favour of the in-group, and/or increasing the relevance of the comparison dimension to the in-group.

Social identity theory, as an explicitly intergroup theory, is relatively "silent" about the effects of between-individual social comparison outcomes, except in so far as such outcomes reveal, reflect, or make salient between-group social comparison outcomes. Schiffmann & Wicklund (1992) have suggested that this silence makes the theory superfluous as, they claim, all of the effects of between-group social comparison outcomes can be reduced to and explained by a theory of the effects of between-individual social comparison outcomes, such as that provided by Tesser's self-evaluation maintenance model (Pleban & Tesser, 1981; Tesser, 1980, 1984, 1986, 1988; Tesser & Campbell, 1980, 1982, 1983; Tesser & Collins, 1988; Tesser & Paulus, 1983; Tesser & Smith, 1980; Tesser et al, 1984, 1988).

Table 6.1 below shows the effects of between-individual social comparison outcomes on self-esteem according to Tesser's self-evaluation maintenance model.^{2,3}

	Relevant Dimension: "Comparison" Process		Irrelevant Dimension: "Reflection" Process	
	Distant Other	Close Other	Distant Other	Close Other
Self Inferior	-	- -	+	++
Self Superior	0	0	0	0

Key: -- = major drop in self-esteem; - = minor drop in self-esteem; ++ = major rise in self esteem; + = minor rise in self-esteem.

Table 6.1: Tesser's self-evaluation maintenance model

Self-relevant comparisons result in a process of "comparison" and self-irrelevant comparisons result in a process of "reflection".⁴ Comparisons are self-relevant if they (i) occur on dimensions which are important to a person's self-definition and (ii) involve comparisons between oneself and others who are "roughly similar" to oneself on that dimension. A comparison dimension is important to a person's self-definition to the extent that the person strives for

² It will be assumed that Tesser's concept of self-evaluation is equivalent to Tajfel's notion of self-esteem, and the two terms are used interchangeably throughout this chapter (see Tesser & Campbell, 1983: 9).

³ From this point forth all discussion concerning effects on self-esteem will be restricted to effects on the level (as opposed to the stability) of self-esteem, unless stated to the contrary.

⁴ Quote marks will be used for Tesser's hypothesized processes of "comparison" and "reflection" in order to distinguish the former from the more general notion of social comparison employed in the literature.

competence on the dimension, describes themselves in terms of the dimension, or freely chooses to engage in tasks related to the dimension (Tesser, 1988: 183). Comparisons with others are "roughly similar" to a person on a comparison dimension if the others are neither "so much better or worse" than the person as to make comparisons "difficult" (i.e. meaningless for self-evaluation purposes).

"Reflection" is analogous to Cialdini's notion of BIRGING: Basking In the Reflected Glory of a superior other (e.g. Cialdini et al., 1976). Positive contributions to self-evaluation can be obtained by "reflection" from another who is superior to oneself on dimensions which are not self-relevant, and the closer one is to the other who is superior on self-irrelevant dimensions, the greater the positive contribution to self-evaluation. Closeness is similar to Heider's (1958) concept of unit-relatedness: "anything that tends to put two persons into a unit relationship increases closeness" (Tesser, 1986: 438). However, as comparisons with others on self-irrelevant dimensions result in a process of "reflection", and as one cannot bask in the glory of an inferior other, there are no effects on self-evaluation following social comparison outcomes which reveal individual superiority to another on self-irrelevant dimensions, regardless of how close the self and other.

"Comparison" essentially involves feeling "bad by comparison" (Tesser, 1986: 438). Inferiority to another on a self-relevant dimension will promote negative contributions to self-evaluation, and the closer one is to the other who is superior on self-relevant dimensions, the greater the resultant negative contributions to self-evaluation. However, as comparisons on self-relevant dimensions result in a process of "comparison", and as one cannot look bad in comparison to an inferior other, there are no effects on self-evaluation following social comparison outcomes which reveal individual

inferiority to another on self-relevant dimensions, regardless of how close the self and other.

Like social identity theory, Tesser's self-evaluation maintenance model is "dynamic" in that it predicts that self-evaluation is a mediating variable between comparison outcomes and "dynamic" behaviour aimed at maintaining, enhancing or reinstating positive self-evaluation. In "comparison" situations, where negative contributions to self-evaluation accrue because of inferiority to a close other who is superior on self-relevant dimensions, the model predicts that a person will attempt to prevent further "comparison" by (i) reducing closeness, (ii) decreasing or reversing the performance differential, and/or (iii) decreasing the relevance of the comparison (Tesser & Campbell, 1983: 8). The model is silent about the dynamic consequences of "reflection" situations, where positive contributions to self-evaluation accrue because of association with a close other who is superior on self-irrelevant dimensions. It is possible that there are no dynamic consequences in such situations, as people are already obtaining positive contributions to self-evaluation, or it is possible that people in "reflection" situations may wish to take dynamic action in order to maintain or enhance those positive contributions, i.e. by (i) increasing closeness, (ii) decreasing relevance, and/or (iii) increasing the performance differential so that the other is increasingly superior to oneself!

The model is also silent about the dynamic consequences of social comparison outcomes which reveal individual superiority to another.

Finally, the self-evaluation maintenance model, as an explicitly interindividual theory, is silent about the effects of between-group social comparison outcomes, except in so far as such outcomes reveal, reflect, or make salient between-individual social comparison outcomes.

There are a number of similarities between Tesser's self-evaluation maintenance model and Tajfel's social identity theory. Both assume individual needs to obtain, maintain, protect, or enhance positive self-esteem; both suggest that such self-esteem needs are affected by social comparison outcomes; and both postulate dynamic cognitive or behavioural strategies to serve self-esteem needs by altering closeness to others, relevance of comparison dimensions and/or differentials between the units of comparison.

There are also a number of stark differences between the theories. First, within the self-evaluation maintenance model "closeness" can only be zero or positive, never negative, whereas it is at least implicit within social identity theory that others can be negatively close (i.e. out-groups and their members) and that comparisons with these negatively close others will have different effects than comparisons with others who are "merely" distant (i.e. who are not in-group members but who are also not relevant out-groups or members of them - see Introduction and the next chapter). Second, the self-evaluation maintenance model claims that only upward social comparisons can affect self-evaluation while social identity theory predicts that both upward and downward social comparisons will do so. Third, social identity theory explicitly posits dynamic strategies in response to situations where positive contributions to self-esteem accrue as well as situations where negative contributions to self-esteem accrue, while the self-evaluation maintenance model explicitly deals with dynamic strategies only in the latter situations. Fourth, social identity theory conceives of situations in which people will wish to increase the relevance of social comparison dimensions, whereas the self-evaluation maintenance model never explicitly conceives of such a possibility. Fifth, the self-evaluation maintenance model conceives of situations in which the outcomes of upward social comparisons can be beneficial for self-evaluation, but social identity

theory predicts that all upward comparison outcomes result in negative, or at best no, contributions to self-esteem. Sixth, and perhaps most important, social identity theory explicitly concerns only between-group social comparisons while the self-evaluation maintenance model explicitly concerns only between-individual social comparisons.

In the light of this last point, it seems worth considering whether each of the theories above can be extended to incorporate both between-individual and between-group social comparisons.⁵

An "extended" self-evaluation maintenance model

Table 6.2 below is essentially Table 6.1 modified simply by substituting "in-group" for the self and "out-group" for the other.

	Relevant Dimension: "Comparison" Process		Irrelevant Dimension: "Reflection" Process	
	Distant Out-group	Close Out-group	Distant Out-group	Close Out-group
In-group Inferior	-	- -	+	++
In-group Superior	0	0	0	0

Key: -- = major drop in self-esteem; - = minor drop in self-esteem; ++ = major rise in self-esteem; + = minor rise in self-esteem; 0 = no change in self-esteem.

Table 6.2: A between-group version of Tesser's SEM model

⁵ For now only the "direct" (self-esteem/self-evaluation) and not the "indirect" or "dynamic" consequences of social comparison outcomes will be considered whilst developing the extended theories.

Consistent with the "original" self-evaluation maintenance model, a "between-groups" version of that model predicts that where social comparisons take place on dimensions which are relevant to in-group members a process of "comparison" takes place. When in-group inferiority is revealed by such comparisons a negative contribution will be made to in-group members' self-evaluation, and the closer the in-group to the out-group, the greater the detrimental effect on the self-evaluation of in-group members from in-group inferiority on in-group relevant dimensions. Where social comparisons reveal in-group superiority to the out-group on in-group relevant dimensions, in-group members' self-evaluation is unaffected as one cannot suffer in comparison to inferior others, regardless of the closeness of the in-group to the out-group.

Similarly, where social comparisons take place on dimensions which are irrelevant to in-group members a process of "reflection" takes place. When in-group inferiority is revealed by such comparisons a positive contribution will be made to in-group members' self-evaluation, and the closer the in-group to the out-group, the greater the positive effect on the self-evaluation of in-group members from in-group inferiority on in-group relevant dimensions. Where social comparisons reveal in-group superiority to the out-group on in-group relevant dimensions, in-group members' self-evaluation is unaffected as one cannot bask in the reflected glory of inferior others, regardless of the closeness of the in-group to the out-group.

Combining the original self-evaluation maintenance model with the between-group version of it, an "extended" self-evaluation maintenance model predicts that: social comparisons which reveal superiority of the self to another or of the in-group to an out-group will not affect the self-evaluation of individuals or in-group members, irrespective of closeness and relevance; all social comparisons which reveal inferiority of the self to another or of the in-group to an out-group on

relevant dimensions will result in negative contributions to self-evaluation; all social comparisons which reveal inferiority of the self to another or of the in-group to an out-group on irrelevant dimensions will result in positive contributions to self-evaluation; and, the closer the self to a superior other or the in-group to a superior out-group, the greater the effects of social comparison outcomes on self-evaluation.

An "extended" social identity theory

A "between-individuals" version of social identity theory would suggest that when personal identity is salient individual superiority over another on a personally relevant comparison dimension will result in positive contributions being made to self-esteem while individual inferiority to another on a personally relevant comparison dimension will result in negative contributions being made to self-esteem (see Lemaine, 1974; Lemaine et al., 1978; Tajfel, 1978b: 16-17). Individual inferiority or superiority to others on self-irrelevant dimensions will not be predicted to affect self-esteem.

Combining the original social identity theory with the between-individuals version of it, an "extended" social identity theory predicts that, when the relevant personal and/or social identities are salient and when the comparison others are relevant comparison others: individual inferiority or superiority to others on personally irrelevant dimensions and/or in-group inferiority or superiority to out-groups on in-group irrelevant dimensions do not affect self-esteem; individual superiority to others on personally relevant dimensions and/or in-group superiority to out-groups on in-group relevant dimensions result in positive contributions to self-esteem; and, individual inferiority to others on personally relevant dimensions and/or in-group inferiority to

out-groups on in-group relevant dimensions result in negative contributions to self-esteem.

The present study

The study presented here was conceived primarily in response to Schiffmann & Wicklund's (1992) argument that social identity theory is superfluous as a social psychological theory as the effects of all between-group social comparison outcomes can be reduced to and fully explained in terms of the effects of between-individual social comparison outcomes.

In the study subjects were randomly categorized as members (Q, R, S or T) of one or the other of two groups (Q and R in Group M, and S and T in Group N) and were given both individual and group false feedback on two bogus performance dimensions of differing relevance to subjects: cognitive dexterity (CD, high relevance) and creative empathy (CE, no relevance).

In order to maximally test Schiffmann & Wicklund's claim that interindividual theories of social comparisons such as Tesser's self-evaluation maintenance model make intergroup theories of social comparisons such as social identity theory redundant, strenuous attempts were made to maximize the salience of subjects' social identities and of the intergroup social comparisons, and to minimize the salience of subjects' personal identities and of the interpersonal social comparisons. If Schiffmann & Wicklund's claim is correct, this should present no difficulties for the self-evaluation maintenance model. But if the claim is incorrect, and social identity theory is correct in claiming that between-group social comparison outcomes can affect self-esteem above and beyond the effects of between-individual social comparison outcomes, it should.

Table 6.3 below shows that false feedback resulted in Group M being positively distinct on the dimension of high relevance (CD) but negatively distinct on the dimension of no relevance (CE), while the opposite was true for Group N. Additionally, while for the sake of conceptual and methodological simplicity the false feedback ensured that individual-level comparisons between members of different groups corresponded to or were consistent with group-level differentials (i.e. when one was positive or negative the other was too), one individual group member outperformed the other within-groups. This resulted in different patterns of social comparison outcomes for each subject.

Subject Q, for example, was a member of a positively distinct group on the relevant dimension, a member of a negatively distinct group on the irrelevant dimension, outperformed their team-mate on the relevant dimension, and underperformed their team-mate on the irrelevant dimension. The other member of Group M, however, subject R, was similarly the member of a positively distinct group on the relevant dimension and of a negatively distinct group on the irrelevant dimension, but underperformed their team-mate on the relevant dimension while outperforming them on the irrelevant dimension.

Similarly, subjects S and T were members of a negatively distinct group on the relevant dimension and of a positively distinct group on the irrelevant dimension, but whereas subject T outperformed their team-mate on the former dimension and underperformed them on the irrelevant dimension, for subject S these between-individual team-mate comparisons had the opposite pattern of outcomes.

	Group M		Group N	
	Subject Q	Subject R	Subject S	Subject T
Team CD Score	67		38	
Team Member CD Score	37	30	15	23
Team CE Score	35		75	
Team Member CE Score	11	24	42	33

Key: CD = Cognitive dexterity, the high relevance comparison dimension; and CE = Creative Empathy, the irrelevant one.

Table 6.3: False feedback employed in the present study

Self-esteem change predictions from the self-evaluation maintenance model

For between-individual comparisons the original self-evaluation maintenance model predicts a three-way interaction between relevance, closeness and individual distinctiveness on self-esteem. For both relevant and irrelevant comparison dimensions positive individual distinctiveness (PID, i.e. individual superiority over the comparison other) will have no effect on self-esteem, regardless of closeness and relevance. Negative individual distinctiveness (NID, i.e. individual inferiority to the comparison other) will have a detrimental effect on self-esteem on the high relevance comparison dimension but a beneficial effect on self-esteem on the irrelevant comparison dimension. Further, the greater the closeness to the comparison other, the more extreme the respective effects on self-esteem of negative individual distinctiveness.

The original self-evaluation maintenance model, as a purely individualistic theory, makes no predictions about the effects of between-group social comparisons on the self-esteem of group members, except in so far as such comparisons reveal, reflect or make salient between-individual differences. As in the present study between-group differences are always consistent with between-individual differences across groups, the model suggests that in-group inferiority to the out-group reveals individual inferiority to both out-group members, while in-group superiority to the out-group reveals individual superiority to both out-group members. Thus, in-group inferiority on the relevant dimension will have a detrimental effect on in-group members' self-esteem; in-group inferiority on the irrelevant dimension will have a beneficial effect on in-group members' self-esteem; and, in-group superiority will have no effect on in-group members' self-esteem, regardless of comparison dimension relevance.

Assuming that the out-group in the present study is a close one, (as the only out-group present), the between-groups version of the self-evaluation maintenance model makes the same between-group predictions as the original model: in-group inferiority on the relevant dimension will make a negative contribution to in-group members' self-esteem; in-group inferiority on the irrelevant dimension will make a positive contribution to in-group members' self-esteem; and, in-group superiority will not affect in-group members' self-esteem, irrespective of dimension relevance.

The between-groups version of the self-evaluation maintenance model makes different predictions to those made by the original model concerning the effects of intragroup between-individual social comparisons outcomes, however. According to the between-groups version of the theory between-individual social comparisons within groups will not affect self-evaluation at all. Rather, self-evaluation will be determined by the between-group social comparison outcomes, even when

comparisons are made within groups (remembering the high salience of social identity and between-group social comparisons and the low salience of personal identities and between-individual social comparisons in the present study). Thus, in-group inferiority to the out-group on the relevant dimension will result in negative contributions to self-evaluation, regardless of whether individuals compare themselves with in-group or out-group others and, in the former case, regardless of whether the individuals are inferior or superior to in-group others. Similarly, in-group inferiority to the out-group on an irrelevant dimension will result in positive contributions to self-evaluation, again regardless of whether individuals compare themselves with in-group or out-group others and, in the former case, regardless of whether individuals are inferior or superior to in-group others. Finally, in-group superiority to the out-group on irrelevant dimensions will result in no change to self-evaluation, regardless of whether individuals compare themselves with in-group or out-group others and, in the former case, regardless of whether the individuals are inferior or superior to in-group others.

The extended version of the self-evaluation maintenance model combines the above two versions and predicts effects for both between-individual and between-group social comparison outcomes on self-evaluation. In the present study, where between-group differences are consistent with between-individual differences across groups, between-group predictions are simple, whether they occur at the genuinely between-group level or at the between-individual level across groups. Where the in-group is inferior to the out-group on a relevant dimension ("comparison" situation), individuals' self-evaluation will suffer, both because of the in-group inferiority on a relevant dimension and because of individuals' inferiority to both out-group members on that dimension. Similarly, where the in-group is inferior to the out-group on an irrelevant dimension ("reflection" situation),

individuals' self-evaluation will benefit, both because of the in-group's closeness to an irrelevantly superior out-group, and because of the individuals' closeness to both irrelevantly superior out-group members. Where the in-group is superior to the out-group, and each in-group member is superior to each out-group member, self-evaluation will not be affected by between-group comparisons, either at the genuinely between-group level or at the between-individual level across groups.

Intragroup between-individual comparisons are more complex according to the extended self-evaluation maintenance model. Where the in-group is inferior to the out-group and individuals are inferior to other in-group members on a relevant comparison dimension the in-group inferiority will contribute negatively to self-evaluations, both because of the in-group's inferiority on a relevant dimension per se, according to the between-groups version of the self-evaluation maintenance model, and because of the individuals' inferiority on a relevant dimension to both out-group members, according to the original version of that model. Also, the individuals' inferiority on a relevant dimension to in-group others will make a further negative contribution to self-evaluation, again according to the original model. These effects will be additive, so that individual and in-group inferiority on a relevant dimension combine to result in a substantially negative contribution to self-evaluation.

Where the in-group is inferior to the out-group and individuals are inferior to other in-group members on an irrelevant comparison dimension the "in-group inferiority" will contribute positively to self-evaluations, both because of the closeness of the in-group to the irrelevantly superior out-group, according to the between-groups version of the self-evaluation maintenance model, and because of the individuals' (lesser) closeness to both irrelevantly superior out-group members, according to the original version of that

model.⁶ Also, the individuals' (greater) closeness to the irrelevantly superior in-group others will make a further positive contribution to self-evaluation, again according to the original model. These effects will be additive, so that individual and in-group inferiority on irrelevant dimensions combine to result in substantially positive contributions to self-evaluation.

Where the in-group is superior to the out-group and individuals are inferior to other in-group members on a relevant comparison dimension the in-group superiority will not affect self-evaluation and thus self-evaluation will be wholly determined by the intragroup between-individual social comparison. This is negative, according to the original self-evaluation maintenance model. Thus, according to the extended version of the self-evaluation maintenance model, negative contributions to self-evaluation when individuals are inferior to in-group others on a relevant comparison dimension will be less marked when the in-group is superior to the out-group on that dimension than when the in-group is inferior to the out-group on that dimension.

Where the in-group is superior to the out-group and individuals are inferior to other in-group members on an irrelevant comparison dimension the in-group superiority will not affect self-evaluation and thus self-evaluation will be wholly determined by the intragroup between-individual social comparison. This is positive, according to the original self-evaluation maintenance model. Thus, according to the extended version of the self-evaluation maintenance model, positive contributions to self-evaluation when individuals are inferior to in-group others on an irrelevant comparison dimension will be less marked when the in-group is superior to the out-group

⁶ "In-group inferiority" is in quote marks in this sentence as it is not the inferiority per se which affects self-evaluation, it is the individuals' or the in-group's closeness to the irrelevantly superior other.

on that dimension than when the in-group is inferior to the out-group on that dimension.

Finally, where the in-group is superior to the out-group and individuals are superior to other in-group members there will be no affect on self-evaluation, regardless of the relevance of the comparison dimension, as superiority does not affect self-evaluation according to any version of the self-evaluation maintenance model.

Thus, for between-individual social comparisons within groups, the extended version of the self-evaluation maintenance model predicts a three-way interaction between in-group distinctiveness, individual distinctiveness and comparison relevance on self-evaluation. Where the in-group is superior to the out-group and individuals are superior to other in-group members, self-evaluation will be unaffected, regardless of comparison relevance. Where the in-group is inferior to the out-group on a relevant comparison dimension there will be a detrimental affect on self-evaluation when individuals are inferior to other in-group members on such a dimension and a lesser detrimental affect on self-evaluation when individuals are superior to other in-group members on such a dimension. Where the in-group is inferior to the out-group on an irrelevant comparison dimension there will be a positive affect on self-evaluation when individuals are inferior to other in-group members on such a dimension and a lesser positive affect on self-evaluation when individuals are superior to other in-group members on such a dimension. Where the in-group is superior to the out-group effects on self-evaluation will be determined solely by between-individual social comparison outcomes. Thus, negative contributions to self-evaluation will result from individual inferiority to in-group others on relevant dimensions (but not as great as where the in-group is also inferior to the out-group on these dimensions); positive contributions to self-evaluation will result from individual inferiority to in-group others on

irrelevant dimensions (but, again, not as great as when the in-group is inferior to the out-group on these dimensions); and, as already mentioned, self-evaluation will be unaffected by individual superiority to in-group others, regardless of comparison dimension relevance.

Self-esteem change predictions from social identity theory

For between-group social comparisons social identity theory predicts an interaction between dimension relevance and group distinctiveness on the self-esteem of group members. On the irrelevant dimension neither positive nor negative in-group distinctiveness will have much effect on the self-esteem of group members. On the dimension of high relevance, however, positive in-group distinctiveness will have a beneficial effect on the self-esteem of in-group members and negative in-group distinctiveness will have a detrimental effect on the self-esteem of in-group members.

As social identity theory is an explicitly intergroup theory, between-individual social comparisons within groups are not predicted to affect self-esteem at all, except in as much as they reveal, reflect or make salient between-group differences. Thus, self-esteem will be determined by between-group social comparison outcomes, even when comparisons are made within groups (remembering the high salience of social identity and between-group social comparisons and the low salience of personal identities and between-individual social comparisons in the present study). Nevertheless, it can be predicted that between-group differences are considerably less salient during intragroup between-individual social comparisons than during between-individual social comparisons across groups, especially as in the present study between-group differences are consistent with between-individual differences across groups. Thus, in-group superiority on the relevant dimension will have a beneficial effect on self-

esteem, regardless of individual inferiority or superiority to in-group others, but the beneficial effect on self-esteem will be greater following across-group individual-level social comparisons (which make the in-group superiority more salient) than following intragroup between-individual ones; in-group inferiority on the relevant dimension will have a detrimental effect on self-esteem, regardless of individual inferiority or superiority to in-group others, but the detrimental effect on self-esteem will be greater following across-group individual-level social comparisons (which make the in-group inferiority more salient) than following intragroup between-individual ones; and, in-group inferiority or superiority on irrelevant dimensions will leave self-esteem unaffected, regardless of individual inferiority or superiority to in-group others or whether between-individual social comparisons occur within or across groups.

For between-individual social comparison outcomes, therefore, social identity theory predicts a three-way interaction between dimension relevance, closeness and group-distinctiveness on the self-esteem of in-group members. In-group superiority on the relevant dimension will have a beneficial effect on self-esteem, but this will be greater when between-individual comparisons occur across groups than when they occur within groups; in-group inferiority on the relevant dimension will have a detrimental effect on self-esteem, but again this will be greater when between-individual comparisons occur across groups than when they occur within groups; and in-group inferiority or superiority on the irrelevant dimension will not affect self-esteem, regardless of whether between-individual comparisons occur within or across groups.

The between-individuals version of social identity theory makes slightly different between-individual predictions, in that individual differences are important in their own right, rather than and instead of "indirectly" in terms of

reflecting, revealing or making salient between-group differences. Thus, individual superiority to another on the relevant dimension will make a beneficial contribution to self-esteem and individual inferiority to another on the relevant dimension will make a detrimental contribution to self-esteem, and the closer the other to the self the greater the effect on self-esteem. Individual inferiority or superiority to another on the irrelevant dimension will not affect self-esteem, regardless of how close the comparison other.

The extended version of social identity theory combines the above two versions and predicts effects for both between-individual and between-group social comparison outcomes on self-esteem. In the present study, where between-group differences are consistent with between-individual differences across groups, between-group predictions are simple, whether they occur at the genuinely between-group level or at the between-individual level across groups. Where the in-group is inferior to the out-group on a relevant dimension, individuals' self-esteem will suffer, both because of the in-group inferiority on a relevant dimension and because of individuals' inferiority to both out-group members on that dimension. Similarly, where the in-group is superior to the out-group on a relevant dimension, individuals' self-esteem will benefit, both because of the in-group superiority on a relevant dimension and because of individuals' superiority to both out-group members on that dimension. Where the in-group is inferior or superior to the out-group on an irrelevant dimension, self-esteem will not be affected.

For between-individual social comparisons within groups the effects on self-esteem are again rather more complicated. Where the in-group is inferior to the out-group on a relevant dimension and individuals are inferior to other in-group members on the same dimension, the in-group inferiority will contribute negatively to self-esteem, both because of the in-

group's inferiority on a relevant dimension per se, according to the original version of social identity theory, and because of the individuals' inferiority on a relevant dimension to both out-group members, according to the between-individuals version of that model. Also, the individuals' inferiority on a relevant dimension to in-group others will make a further negative contribution to self-esteem, according to the between-individuals version of social identity theory. These effects will be additive, so that individual and in-group inferiority on a relevant dimension combine to result in a substantially negative contribution to self-esteem.

Where the in-group is inferior to the out-group on a relevant dimension and individuals are superior to other in-group members on the same dimension, the in-group inferiority will contribute negatively to self-esteem, both because of the in-group's inferiority on a relevant dimension per se, according to the original version of social identity theory, and because of the individuals' inferiority on a relevant dimension to both out-group members, according to the between-individuals version of that model. However, here the individuals' superiority on a relevant dimension to in-group others will make a positive contribution to self-esteem, according to the between-individuals version of social identity theory. These effects will be additive, so that in-group inferiority but individual superiority to other in-group members on a relevant dimension combine to result in either no affect on or a slight negative contribution to self-esteem.

Where the in-group is superior to the out-group on a relevant dimension and individuals are superior to other in-group members on the same dimension, the in-group superiority will contribute positively to self-esteem, both because of the in-group's superiority on a relevant dimension per se, according to the original version of social identity theory, and because of the individuals' superiority on a relevant dimension to both out-group members, according to the between-individuals

version of that model. Also, the individuals' superiority on a relevant dimension to in-group others will make a further positive contribution to self-esteem, according to the between-individuals version of social identity theory. These effects will be additive, so that individual and in-group superiority on a relevant dimension combine to result in a substantially positive contribution to self-esteem.

Where the in-group is superior to the out-group on a relevant dimension and individuals are inferior to other in-group members on the same dimension, the in-group superiority will contribute positively to self-esteem, both because of the in-group's superiority on a relevant dimension per se, according to the original version of social identity theory, and because of the individuals' superiority on a relevant dimension to both out-group members, according to the between-individuals version of that model. However, here the individuals' inferiority on a relevant dimension to in-group others will make a negative contribution to self-esteem, according to the between-individuals version of social identity theory. These effects will be additive, so that in-group superiority but individual inferiority to other in-group members on a relevant dimension combine to result in either no affect on or a slight positive contribution to self-esteem.

Finally, when social comparisons take place on irrelevant dimensions there will be no effect of intragroup between-individual social comparison outcomes on self-esteem, regardless of in-group or individual distinctiveness from in-group others.

Thus, for between-individual social comparisons within groups, the extended version of social identity theory predicts a three-way interaction between in-group distinctiveness, individual distinctiveness and comparison relevance on self-esteem. Where social comparisons take place on irrelevant dimensions self-esteem will be unaffected, regardless of in-

group or individual distinctiveness from other in-group members. Where the in-group is superior to the out-group on a relevant dimension self-esteem will benefit when individuals are also superior to in-group others on this dimension but will benefit much less, if at all, when individuals are inferior to in-group others on this dimension. Where the in-group is inferior to the out-group on a relevant comparison dimension self-esteem will suffer when individuals are also inferior to in-group others on this dimension but will suffer much less, if at all, when individuals are superior to in-group others on this dimension.

Dynamic predictions from the self-evaluation maintenance model⁷

As the self-evaluation maintenance model hypothesizes that the effect on self-evaluation of comparisons with a superior close other on a relevant dimension will be detrimental while the effect on self-evaluation of comparisons with an inferior close other on a relevant dimension will be neutral, the model predicts that subjects will not wish to aid the performance of close others on a relevant dimension, especially if there is an indication that the close others' performances on that dimension are likely to be superior to their own.

Similarly, as the model hypothesizes that the effect on self-evaluation of comparisons with a close other on a relevant dimension will be strongly detrimental (if the other is superior) or neutral (if the other is inferior), while the effect on self-evaluation of comparisons with a distant other on a relevant dimension will be relatively less detrimental (if the other is superior) or neutral (if the other is inferior), the model predicts that subjects will distance

⁷ For "dynamic" predictions, only the "original" versions of the self-evaluation maintenance model and social identity theory will be considered.

themselves from close others performing on a relevant dimension, especially if the others' performances on that dimension are superior to their own.

Finally, as the model hypothesizes that the effect on self-evaluation of comparisons with a close other on a relevant dimension will be detrimental (if the other is superior) or neutral (if the other is inferior), while the effect on self-evaluation of comparisons with a close other on an irrelevant dimension will be beneficial (if the other is superior) or neutral (if the other is inferior), the model predicts that subjects will reduce the relevance of a comparison dimension when asked to compare themselves with close others on that dimension, especially when the others' performances are superior to their own.

Dynamic predictions from social identity theory

As social identity theory hypothesizes that subjects with salient social identities derive their self-esteem from group-level social comparison outcomes on relevant comparison dimensions, the theory predicts that subjects will wish to improve the performance of a close (i.e. in-group) other on relevant performance dimensions, especially if there is an indication that unless this is done the in-group may be outperformed by the out-group.

Similarly, as social identity theory hypothesizes that subjects with salient social identities derive positive self-esteem from membership of in-groups which are positively distinct on relevant comparison dimensions but negative self-esteem from membership of in-groups which are negatively distinct on such comparison dimensions, the theory predicts that subjects will increase their closeness to other in-group members when the in-group is positively distinct on a relevant comparison dimension but will decrease their closeness to

other in-group members when the in-group is negatively distinct on such a comparison dimension.

Finally, as social identity theory hypothesizes that positive in-group distinctiveness on a relevant dimension has a beneficial effect on the self-esteem of subjects with salient social identities and negative in-group distinctiveness on a relevant dimension has a detrimental effect on the self-esteem of subjects with salient social identities, but neither positive nor negative in-group distinctiveness on an irrelevant dimension have much effect on the self-esteem of in-group members, the theory predicts that subjects will increase the relevance of comparison dimensions on which the in-group is positively distinct but will reduce the relevance of comparison dimensions on which the in-group is negatively distinct.

METHOD

Subjects

Over the space of four days 60 students from Keele University took part in a study which they were told was part of a programme to develop two psychometric measures for possible later adoption by the British Civil Service for selection/recruitment purposes. The majority of subjects were first year psychology students who volunteered to participate in the study in partial fulfilment of "subject time" requirements applying during the first year of their degree programme. A small minority of subjects were student friends of the above who volunteered at the last minute to take part in the study because of non-attendance by others who had previously volunteered but did not turn up (as four subjects were required in each session - see below).

Procedure

Subjects were seen in fifteen sessions of four subjects each. Upon arrival each subject was asked to sit "wherever they liked". In fact only four chairs were available, each at a separate individual desk. These chairs were arranged so that two desks were next to each other, but separated, and faced another pair of desks similarly organized.⁸ On each desk was a piece of paper clearly marked with a large "identifying letter": Q, R, S or T. Subjects were requested to write their own letter at the top of each piece of paper they used in order to be able to compare individual's responses across measures whilst assuring anonymity. Subjects were then asked to complete Rosenberg's (1965) self-esteem scale modified for use as a "state" measure (i.e. by inserting the phrase "at the moment" at the beginning of each item and removing all "trait-like" phrases such as "In general". See first self-esteem measure in Appendix 10).

These were then collected and the experimenter explained the ostensible purpose of the study. Subjects were "reminded" that the study was one part of a programme of research aimed at developing two measures of cognitive ability for possible later use by the British Civil Service in their selection procedures. Subjects were informed that one measure had provisionally been termed "cognitive dexterity" and had already been discovered to "reliably distinguish between people with high or low verbal reasoning, spatial awareness, problem solving abilities, initiative, and even social skills". The other measure, provisionally termed "creative empathy", had, however, failed to live up to expectations and did not reliably distinguish between people in terms of any practical abilities whatsoever. Subjects were told that the present study was the final one in the programme and, although

⁸ As far as the experimenter could tell subjects who were friends with other subjects sat next to them or sat opposite them with approximately equal frequency.

important for validation purposes, was not expected to disconfirm these previous findings. The importance of the validation procedure was emphasized, however, and subjects were requested to complete their tasks diligently and with full honesty.

Subjects were told that many of their tasks required the use of semantic differential scales and these scales were explained and demonstrated with examples. Ostensibly merely to ensure that subjects understood how to use semantic differential scales, and also to introduce to subjects some concepts which would be important during the main part of the study, subjects were asked to complete "scale practice sheets" which used semantic differential scales to measure their self-esteem, their feelings of closeness to the person sitting next to them (who, unbeknownst to the subjects were to be in-group members during the main study), and the relevance of the two measures being developed (i.e. cognitive dexterity and creative empathy). In fact, these measures served as pre-manipulation measures of each of these constructs. The "scale practice sheet" can be seen in Appendix 11.

The experimenter then read out the following prepared text:-

For the next part of the study I need to split you into two teams or groups. You two, Q and R, will be in Group M and you two, S and T, will be in Group N. Please move your desk so that you are close to the other person in your group. In a moment each team will be competing with the other, with the possibility of one of the teams winning an £18 prize. Before then, though, please work within your teams for one minute and come up with as many words as you can beginning with your Group letter (i.e. M or N).

When subjects had completed this short task they were then asked to work together within their groups for another minute to generate as many words as they could starting with the Group letter of the other group present (i.e. the out-group). Both the tasks (which were not scored or analysed) and the anticipated intergroup competition for desired scarce

resources were introduced to engender within-group cohesion and salient social identities related to subjects' particular experimental in-group membership.

Next, the experimenter read out the following:-

I am now going to give each of you tests of "Cognitive Dexterity" and "Creative Empathy". Although I want you to work individually, your aim should be to get as high a score as possible for your Group, with the overall goal of your Group beating the other Group by as much as possible on each test. As an added incentive, the team that beats their opponents by the largest amount this week on the cognitive dexterity test will win £18. The tests are very difficult and you will have only 10 minutes to complete both tests. If you are having trouble on a question, don't panic: leave it and come back to it later if you have time. The more questions you answer, the higher your score is likely to be. You must spend a minimum of 2 minutes on each task. You may divide the remaining 6 minutes between each task as you see fit, but I remind you that only the cognitive dexterity task has the possibility of a cash prize associated with it. Begin with the cognitive dexterity task. At the end of two minutes I will instruct you to turn to the creative empathy task. After a further two minutes I will tell you that you may continue with either or both tasks as you see fit.

This manipulation was intended to again strengthen within-group cohesion and social identity. More importantly, it was also intended to impress upon subjects the inter-group nature of the competition inherent in their task. That is, it was hoped that this manipulation would cause subjects to be more interested in the performance of their group relative to their out-group (in order to obtain more positive in-group distinctiveness from their out-group on the cognitive dexterity dimension than any other in-group obtained from their out-group across the experimental sessions) than in the performance of any one individual, other than in terms of how individual performances contributed to group performances. Finally, this manipulation was intended to enhance the dimension relevance manipulation, so that subjects perceived cognitive dexterity performance as relevant but creative empathy performance as irrelevant.

Subjects were then asked to individually complete bogus cognitive dexterity and creative empathy tasks (which were not scored or analysed). The cognitive dexterity task involved subjects completing a number of categorization tasks. Item 1 on this task, for example, informed subjects that "Any letter between f and p is in Group 1 and all other letters are in Group 2. Which groups are the following letters in (1 or 2)?: x, l, q, f, o". The creative empathy task involved subjects looking at a series of pictures (taken from an unidentified projective test) and either selecting or generating a word which they thought was most popular when "a number of established poets were asked to select a word that best communicates the "feeling" of that picture". Item 1 on this task, for example, referred subjects to a picture of a seated boy looking at a violin on the table in front of him and asked them to select one of the following words: "melancholy, gulf, ambitious, hungry". Each of the tasks (except the pictures used, which can be obtained from the author on request) can be seen in Appendix 12.

When the subjects had spent ten minutes on the two tasks their response booklets were collected in and the experimenter said:-

While I score these, please work within your teams to find as many words of three or more letters as you can within the phrase "Cognitive Dexterity".

Subjects did this filler task (which was not scored or analysed) for five minutes whilst the experimenter made a show of ostensibly marking each of the subjects' response booklets and writing down scores on a blank piece of paper. The experimenter used an overhead projector to project a blank table onto the screen, and proceeded to write the false feedback scores into that table. These scores can be seen in Table 6.3 above. Once the "raw scores" had been entered into the table the experimenter made a show of calculating the difference in team cognitive dexterity scores and stressed that Group M's superiority over Group N on this score was the

"best seen yet" and would be "difficult to beat" in later sessions. It was also pointed out that Group N beat Group M on the creative empathy score, but it was reiterated that this result had little consequence. The false feedback scores were projected onto the screen for the remainder of the experimental session.

The subjects were then asked to fill out a second "state" version of Rosenberg's (1965) self-esteem scale. Once these scales had been completed and collected the experimenter said:

Finally, this is the part of the study that is most important in developing the measures mentioned earlier. Please answer the following questions as honestly as you can.

The subjects were then each given response booklets containing the principal post-manipulation measures. This booklet was headed "Cognitive dexterity/creative empathy study" and began with a short section asking subjects to indicate their individual and team membership letters, and whether their team "won" or "lost" the cognitive dexterity and the creative empathy tasks just completed. There then followed four more sections concerning subjects' self-esteem, their willingness to have aided the performance of their team-mate in the tasks just completed had they been able to do so, how close they felt to their team-mate, and how relevant they thought the two comparison dimensions. Each section contained a number of items. Before each item subjects were asked to consider a particular team or other-individual cognitive dexterity or creative empathy score. In the self-esteem section, for example, subjects were asked to indicate their self-esteem on single semantic differential scales immediately after considering: the out-group's cognitive dexterity score; each out-group member's cognitive dexterity score; their team-mate's cognitive dexterity score; and each creative empathy score for each of these comparison others. It was assumed that being asked to focus on particular other individual or out-group scores would encourage social comparisons between

those other individuals and the self or between the out-group and the in-group for each salient dimension. The final page of the booklet also asked subjects to indicate their perceptions concerning the purpose of the present study and invited them to make any comments they wished to about the study. The response booklet used in this part of the experiment can be seen in Appendix 13.

When this final procedure was completed subjects were asked to refrain from telling subjects yet to take part in the study anything about it and were asked if they themselves had heard anything about the study prior to their own session. All promised the former and denied the latter. Subjects were then informed that a notice would be pinned up in the Psychology Department the following week indicating the results of the study and who had won the team cognitive dexterity prize. In fact the notice revealed the nature of the false feedback; explained that the study was attempting to compare Tajfel's social identity theory and Tesser's self-evaluation maintenance theory with regards to the predictions each made about the effects of various social comparison outcomes; and that two participating teams had been selected at random to each receive £18 (i.e. £9 per selected subject). The members of these selected teams were personally approached by the experimenter, given their money and quizzed once again about whether or not they had heard anything about the study prior to participating in it, and about how they had interpreted what the study was all about. All four of these subjects reiterated that they had heard nothing about the study prior to completion, insisted that they had made every effort to complete all tasks as instructed, and claimed to have accepted the experimenter's cover story concerning the ostensible development of cognitive dexterity and creative empathy measures.

RESULTS

Pre-manipulation measures

Self-esteem scores derived from the modified Rosenberg measure administered immediately upon subjects' arrival revealed slightly positive mean self-esteem ($\bar{X} = 4.84$, $SD = 0.97$, on a seven-point scale ranging from 1 = very negative self-esteem to 7 = very positive self-esteem) and no significant differences in that self-esteem according to the subjects' "identifying letters" ($F_{(3,56)} = 0.68$, $p = 0.57$) or according to subjects' (as yet unknown to them) group memberships (planned contrast, $T = -1.28$, $p = 0.21$).

Similarly, the single pre-manipulation semantic differential self-esteem scale indicated moderate positive mean self-esteem ($\bar{X} = 6.93$, $SD = 1.67$, on a nine-point scale ranging from 1 = bad self-esteem to 9 = good self-esteem) and no significant differences in that self-esteem according to identifying letters ($F_{(3,56)} = 1.01$, $p = 0.40$) or group membership (planned contrast $T = -1.40$, $p = 0.17$). This measure was significantly and reasonably strongly positively correlated with the first presentation of the modified Rosenberg self-esteem measure ($r = 0.68$, $p < 0.001$, one-tailed).

Prior to the experimental manipulation subjects felt neither close to nor distant from the person sitting next to them (i.e. their as yet unrevealed fellow in-group member, $\bar{X} = 5.07$, $SD = 2.11$, on a nine-point scale ranging from 1 = very distant to 9 = very close), with no significant differences in closeness according to identification letter ($F_{(3,56)} = 0.62$, $p = 0.61$) or group membership (planned contrast $T = 1.09$, $p = 0.28$).

Manipulation check

Strenuous efforts were taken to make cognitive dexterity a highly relevant dimension for subjects and to make creative empathy an irrelevant one. However, prior to the main experimental manipulation (i.e. the false feedback) subjects felt that both were relevant dimensions: mean cognitive dexterity (CD) relevance = 6.70 (SD = 1.53), and mean creative empathy (CE) relevance = 6.12 (SD = 1.59) (both significantly above the mid-point on nine-point scales ranging from 1 = irrelevant to 9 = relevant). A 2 X 4 ANOVA using the MANOVA procedure of SPSS,⁹ with the relevance of the CD and the CE measures as a within-subject factor and subjects' identification letters as a between-subject factor, revealed no main effect of identification letter and no significant interaction, but a main effect of relevance was found ($F_{(1,56)} = 10.2, p = 0.002$). Thus, in line with expectations, subjects evaluated cognitive dexterity performance as more relevant than creative empathy performance but the difference, although statistically significant, was slight, and it is perhaps best to say that subjects perceived both of the performance dimensions as relevant, even though they thought cognitive dexterity performance slightly more relevant than creative empathy performance. In as much as subjects did not perceive creative empathy performance to be irrelevant, the relevance manipulation failed. In all analyses to follow, therefore, both dimensions were considered as relevant and "collapsed".¹⁰

⁹ All statistical analysis was carried out using SPSS. Within this package all ANOVAs including within-subject factors have to be calculated using the MANOVA procedure. All ANOVAs including within-subjects factors reported below were therefore calculated using this procedure but for the sake of exposition this will not be made explicit each time.

¹⁰ The author nevertheless repeated each analysis with each dimension considered separately, and again with relevance included as a separate factor. In every instance, each of these non-reported additional analyses resulted in a similar pattern to those reported (i.e. with relevance collapsed).

Thus, only predictions pertaining to the effects of social comparison outcomes on relevant dimensions can be examined.

Scoring

Change scores as a result of the experimental manipulation were calculated by subtracting subjects' pre-manipulation scores from their post-manipulation scores for self-esteem, closeness, and relevance, in each case using the appropriate semantic differential scores. Change in self-esteem as a result of focusing on the out-group's cognitive dexterity score, for example, (which was assumed to encourage a between-groups comparison on this dimension), was calculated by subtracting pre-manipulation self-esteem from the self-esteem indicated immediately following this focus. These change scores could range from -8 to +8, with minus scores indicating a decrease and positive scores indicating an increase in whatever measure was being examined.

No pre-manipulation performance scores were obtained and thus no change in such scores as a result of the experimental manipulation were calculable. Thus, for performance, there are simply post-manipulation scores.

Two measures were obtained for each comparison with out-group members, one for each out-group member. Where this occurred single "comparison with a distant (i.e. out-group) other" measures were calculated by simply taking the mean average of each of these equivalent measures.

Self-esteem results: the self-evaluation maintenance model

With respect to between-individual social comparisons on relevant dimensions, the original self-evaluation maintenance model predicts that comparison with an inferior other will not

affect self-esteem but that comparison with a superior other will be detrimental to self-esteem, and that the closer the relevantly superior comparison other to the self, the greater the detrimental effect on self-esteem.

To test this a 2 X 2 ANOVA was calculated, with closeness of the comparison other to the self and individual superiority or inferiority to the comparison other as two within-subject factors. This revealed a significant main effect of individual distinctiveness ($F_{(1,59)} = 28.82, p < 0.001$), no significant main effect of closeness ($F_{(1,59)} = 0.04, p = 0.85$), and a significant interaction ($F_{(1,59)} = 15.67, p < 0.001$). The table of means below suggests that comparison with a superior other on a relevant dimension was detrimental to self-esteem, and that the more distant the relevantly superior comparison other to the self, the greater the detrimental effect on self-esteem. Further, comparison with an inferior other on a relevant dimension was *beneficial* to self-esteem, and the more distant the relevantly inferior other to the self, the greater the beneficial effect on self-esteem.

	Relevant Dimension: "Comparison" Process	
	Distant Other	Close Other
Self inferior	-0.617* (1.88)	-0.267 (1.89)
Self superior	0.692* (1.49)	0.317 (1.62)

Key: * = Significantly different from zero by one sample t-test (two-tailed)

Table 6.4: Change in self-esteem as a result of upward or downward comparison with a close or distant other

A series of two-tailed one sample t-tests revealed that only the effects on self-esteem of comparisons with a distant (i.e.

out-group) other were significantly greater than zero. That is, the fall in self-esteem as a result of comparison with a relevantly superior close (i.e. in-group) other was not significant ($t_{(59)} = -1.09, p = 0.28$),¹¹ and nor was the rise in self-esteem as a result of comparison with a relevantly inferior close (i.e. in-group) other ($t_{(59)} = 1.51, p = 0.136$). Self-esteem change was only significantly greater than zero for the fall in self-esteem following comparison with a relevantly superior distant (i.e. out-group) other ($t_{(59)} = -2.54, p = 0.014$) and for the rise in self-esteem following comparison with a relevantly inferior distant (i.e. out-group) other ($t_{(59)} = 3.59, p = 0.001$).

These results are almost wholly inconsistent with the "comparison" process within the original self-evaluation maintenance model. Comparison with a relevantly superior close other did not result in a significant fall in self-esteem; comparison with a relevantly superior distant other *did* result in a significant fall in self-esteem; and comparison with a relevantly inferior distant other resulted in a significant rise in self-esteem. Although the model can attempt to explain the greater fall in self-esteem following upward individual-level social comparisons on relevant dimensions across groups than following such comparisons within-groups in terms of the across-group performance differences being greater than the within-group performance differences,¹² the model cannot explain why comparison with a superior close other on a relevant dimension was not sufficient to cause a significant drop in self-esteem, or, especially, why comparison with a relevantly inferior distant other was sufficient to cause a significant increase in self-esteem.

¹¹ All t-test probabilities are two-tailed except where specified.

¹² I am grateful to Mark Schaller for pointing this possibility out.

With respect to between-group social comparisons on relevant dimensions, the original self-evaluation maintenance model predicts that comparison between the in-group and an inferior out-group will not affect self-esteem but that comparison between the in-group and a superior out-group will be detrimental to self-esteem (as the in-group inferiority reveals or makes salient individual inferiority to both out-group members). Table 6.5 below shows that in support of the model in-group inferiority on the relevant dimension did indeed cause a significant drop in self-esteem ($\bar{X} = -0.617$, $SD = 1.89$, $t_{(59)} = -2.53$, $p = 0.014$), but against the model in-group superiority on this dimension significantly raised self-esteem ($\bar{X} = 0.633$, $SD = 1.34$, $t_{(59)} = 3.66$, $p = 0.001$).

	Relevant Dimension: "Comparison" Process	
	In-group superior	In-group inferior
Self-esteem change	0.633* (1.34)	-0.617* (1.89)

key: * - Significantly different from zero by one sample t-test (two-tailed)

Table 6.5: Self-esteem change as a result of in-group superiority or inferiority on relevant dimensions

Several of the above results are relevant to the extended version of the self-evaluation maintenance model. Consistent with the extended model: (i) comparisons between in-groups and relevantly superior out-groups were significantly detrimental to self-esteem; and, (ii) comparisons between individuals and superior out-group members, which could only occur in conjunction with out-groups being superior to in-groups, resulted in detrimental effects on self-esteem. Inconsistent with the extended model, however: (i) comparisons between in-groups and relevantly inferior out-groups did not leave self-

esteem unaffected - they were significantly beneficial to self-esteem; and, (ii) comparisons between individuals and inferior out-group members, which could only occur in conjunction with out-groups being inferior to in-groups, also did not leave self-esteem unaffected - they too were significantly beneficial to self-esteem.

The extended version of the self-evaluation maintenance model makes different predictions to the original version of that model concerning the effects on self-esteem of intragroup between-individual social comparisons on relevant dimensions, however. In such situations the extended model predicts a two-way interaction between in-group distinctiveness and individual distinctiveness: where the in-group is inferior to the out-group there will be a detrimental effect on self-esteem when individuals are inferior to other in-group members and a lesser detrimental effect on self-esteem when individuals are superior to other in-group members; where the in-group is superior to the out-group there will be a detrimental effect on self-esteem when individuals are inferior to in-group others (but not as great as where the in-group is also inferior to the out-group on these dimensions), but no effect on self-esteem when individuals are superior to in-group others.

These predictions were tested by a 2 X 2 ANOVA with in-group distinctiveness as a within-subjects factor and "consistency of individual distinctiveness" as a between-subjects factor, with self-esteem change as the dependent variable. The latter factor refers to whether or not subjects' individual distinctiveness is consistent with their group distinctiveness (i.e. positive when the latter is positive or negative when the latter is negative) or inconsistent with it (i.e. negative when the latter is positive or positive when the latter is negative). This resulted in a significant main effect of group distinctiveness ($F_{(1,58)} = 18.45, p < 0.001$), no significant main effect of "consistency" ($F_{(1,58)} = 0.000, p =$

0.968), and a significant interaction ($F_{(1,58)} = 10.23, p = 0.002$). It can be seen from the table of means below that the extended version of the self-evaluation maintenance model is supported in so far as there is a detrimental effect on self-esteem when the in-group is inferior to the out-group on relevant dimensions, and that this detrimental effect is stronger when individuals are inferior to in-group others than when they are superior to them. Against the extended model, however, there was no change (instead of a fall) in self-esteem when the in-group was superior to the out-group and individuals were inferior to in-group others, and there was a beneficial (rather than no) effect on self-esteem when the in-group was superior to the out-group and individuals were superior to in-group others.

	In-group Inferior	In-group Superior	Mean
Self Inferior	-0.667* (2.29)	0.133 (1.31)	-0.267 (1.89)
Self Superior	-0.067 (1.64)	0.700* (1.54)	0.317 (1.62)
Mean	-0.367 (2.00)	0.417* (1.44)	

Key: * = Significantly different from zero by one sample t-test (two-tailed); and † = difference from zero approached significance by one sample t-test (two-tailed, $p = 0.061$).

Table 6.6: Changes in self-esteem as a result of in-group inferiority or superiority to the out-group and individual inferiority or superiority to in-group others

Self-esteem change results: social identity theory

Social identity theory predicts that in-group superiority on a relevant dimension will raise self-esteem and in-group

inferiority on such a dimension will lower it. As was seen in Table 6.5 above, these predictions were fully supported: for between-group social comparisons on relevant dimensions, in-group superiority significantly raised self-esteem ($\bar{X} = 0.633$, $SD = 1.34$, $t_{(59)} = 3.66$, $p = 0.001$) and in-group inferiority significantly lowered it ($\bar{X} = -0.617$, $SD = 1.89$, $t_{(59)} = -2.53$, $p = 0.014$).

With respect to between-individual social comparisons on relevant dimensions, social identity theory predicts that in-group superiority will be beneficial to self-esteem and in-group inferiority will be detrimental to self-esteem, regardless of individual inferiority or superiority to the comparison other, except in so far as such individual distinctiveness reveals, reflects or makes salient between-group distinctiveness. These effects will be stronger for between-individual comparisons with distant (i.e. out-group) others than for between-individual comparisons with close (i.e. in-group) others, therefore, as between-group comparisons are more salient in the former than in the latter situation.

To test this a 2 X 2 ANOVA was calculated, with closeness of the comparison other to the self and in-group superiority or inferiority to the out-group as two within-subject factors. This revealed a significant main effect of group distinctiveness ($F_{(1,59)} = 40.36$, $p < 0.001$), no significant main effect of closeness ($F_{(1,59)} = 0.04$, $p = 0.85$), and a significant interaction ($F_{(1,59)} = 6.89$, $p = 0.011$). The table of means below suggests that in-group inferiority was detrimental to self-esteem, but more so when the self was compared with an out-group other than when the self was compared with an in-group other. In-group superiority was beneficial to self-esteem, and more so when the self was compared with an out-group other than when the self was compared with an in-group other.

	Out-group other	In-group other
In-group Inferior	-0.617* (1.88)	-0.367 (2.00)
In-group Superior	0.692* (1.49)	0.417* (1.44)

Key: * = Significantly different from zero by one sample t-test (two-tailed)

Table 6.7: Changes in self-esteem as a result of in-group inferiority or superiority following comparisons with in-group or out-group others

A series of two-tailed one sample t-tests revealed that in-group inferiority or superiority significantly altered self-esteem from pre-manipulation levels except when in-group inferiority was accompanied by a comparison between the self and an in-group other. That is, comparison with an out-group other significantly lowered self-esteem when the in-group was inferior to the out-group ($t_{(59)} = -2.54$, $p = 0.014$) and significantly raised it when the in-group was superior to the out-group ($t_{(59)} = 3.59$, $p = 0.001$). Comparison with an in-group member significantly raised self-esteem when the in-group was superior to the out-group ($t_{(59)} = 2.24$, $p = 0.029$) but did not significantly lower it when the in-group was inferior to the out-group ($t_{(59)} = -1.42$, $p = 0.161$).

These results are almost wholly consistent with the predictions derived from social identity theory. Subjects' self-esteem increased as a result of in-group superiority and decreased as a result of in-group inferiority, especially when the self was compared with an out-group member rather than an in-group member (with the sole exception that in-group inferiority did not significantly lower self-esteem when individuals compared themselves with in-group others).

Several of the above results are relevant to the extended version of social identity theory. Consistent with the extended theory: (i) comparisons between the in-group and inferior out-groups resulted in significantly beneficial effects on self-esteem; (ii) comparisons between the in-group and superior out-groups resulted in significantly detrimental effects on self-esteem; (iii) comparisons between individuals and inferior out-group others, which could only occur when out-groups were also inferior to in-groups, resulted in significantly beneficial effects on self-esteem; and, (iv) comparisons between individuals and superior out-group others, which could only occur when out-groups were also superior to in-groups, resulted in significantly detrimental effects on self-esteem. None of the results so far are inconsistent with the extended version of social identity theory.

The extended version of social identity theory makes different predictions to the original version of that theory concerning the effects on self-esteem of intragroup between-individual social comparisons on relevant dimensions, however. In such situations the extended version of social identity theory predicts a two way interaction between in-group inferiority or superiority to the out-group and individual inferiority or superiority to in-group others: where the in-group is superior to the out-group self-esteem will benefit when individuals are also superior to in-group others but will benefit much less, if at all, when individuals are inferior to in-group others; and, where the in-group is inferior to the out-group self-esteem will suffer when individuals are also inferior to in-group others but will suffer much less, if at all, when individuals are superior to in-group others.

This is exactly the pattern of results shown in Table 6.6 above. Where the in-group is superior to the out-group on relevant dimensions there is a significant rise in self-esteem when individuals are also superior to in-group others ($\bar{X} = 0.700$, $SD = 1.54$, $t_{(29)} = 2.50$, $p = 0.009$, one-tailed), and a

slight but non-significant change in self-esteem when individuals are inferior to in-group others ($\bar{X} = 0.133$, $SD = 1.31$, $t_{(29)} = 0.56$, $p = 0.580$, two-tailed). And, when the in-group is inferior to the out-group on relevant dimensions there is a near significant fall in self-esteem when individuals are also inferior to in-group others ($\bar{X} = -0.667$, $SD = 2.29$, $t_{(29)} = -1.59$, $p = 0.061$, one-tailed), and a slight but clearly non-significant fall in self-esteem when individuals are superior to in-group others ($\bar{X} = -0.067$, $SD = 1.64$, $t_{(29)} = -0.22$, $p = 0.825$, two-tailed).

Performance results

The self-evaluation maintenance model predicts that subjects will be unwilling to aid the performance of close others on relevant dimensions, particularly when the other's performance is likely to be superior to the self's on those dimensions. Social identity theory, on the other hand, predicts that subjects will want to help the performance of in-group others on relevant performance dimensions, especially if the out-group seems likely to outperform the in-group on the dimension unless this is done.

These predictions were tested by post-manipulation measures. Subjects were asked to consider their team-mate's performance during the previously completed tasks and were then asked if they would have aided their team-mate's performance during those tasks had they legitimately been able to do so. They responded that they almost certainly would have ($\bar{X} = 7.35$, $SD = 1.87$, on a nine-point scale ranging from 1 = certainly not to 9 = certainly, with this mean being significantly above the scale mid-point, $t_{(59)} = 9.75$, $p < 0.001$). Further, subjects were equally willing to aid their team-mate's performance whether that performance was inferior ($\bar{X} = 7.40$, $SD = 2.01$) or superior ($\bar{X} = 7.30$, $SD = 2.15$) to their own ($T_{(59)} = -0.42$, $p = 0.68$), although they were more willing to help when the in-

group's performance on the previous tasks had been inferior ($\bar{X} = 7.67$, $SD = 1.87$) rather than superior ($\bar{X} = 7.03$, $SD = 2.23$) to the out-group's performance on those tasks ($t_{(59)} = 2.84$, $p = 0.006$). Also, when subjects were asked if they would have helped their team-mate's performance after being asked to consider the out-group's performance during the previous tasks (rather than their team-mate's), they still indicated that they would have ($\bar{X} = 7.21$, $SD = 1.93$, significantly above the scale mid-point, $t_{(59)} = 8.87$, $p < 0.001$), although they showed more willing when the in-group's performance had been inferior ($\bar{X} = 7.52$, $SD = 1.99$) than when it had been superior ($\bar{X} = 6.90$, $SD = 2.41$) to the out-group's ($t_{(59)} = 2.25$, $p = 0.028$).

These results are wholly inconsistent with the self-evaluation maintenance model and wholly consistent with social identity theory. As predicted by social identity theory but in opposition to the predictions derived from the self-evaluation maintenance model, subjects were very willing to aid the performance of a close (i.e. in-group) other on a relevant dimension, regardless of whether their team-mate's performance on the recently completed tasks had been superior or inferior to their own. Further, again in line with social identity theory, this willingness to help was stronger when the in-group's performance on the completed tasks was inferior to the out-group's than when it was superior, regardless of whether subjects were invited to compare their own performance with their team-mate's or whether they were invited to compare in-group and out-group performances directly.

Closeness results

The self-evaluation maintenance model predicts that subjects would wish to distance themselves from close others performing on relevant comparison dimensions, especially if the close other is likely to outperform the self on those dimensions. Social identity theory, on the other hand, predicts that

subjects would wish to increase the closeness to in-group others when the in-group achieved superiority over the out-group on relevant dimensions, but would wish to distance themselves from in-group others when the in-group "achieved" inferiority to the out-group on such dimensions.

When subjects were asked to consider their team-mate's performances on the recently completed tasks and were then asked to indicate how close or distant they felt to them, they indicated that they felt closer to them than they had prior to the false feedback manipulation ($\bar{X} = +1.042$, $SD = 1.86$, $t_{(59)} = 4.35$, $p < 0.001$), regardless of whether their team-mate's performances had been superior ($\bar{X} = +0.983$, $SD = 1.94$) or inferior ($\bar{X} = +1.100$, $SD = 1.85$) to their own ($t_{(59)} = -1.15$, $p = 0.253$). There was a significant difference between increased closeness on this measure according to whether the in-group's performance had been inferior ($\bar{X} = +0.900$, $SD = 1.82$) or superior ($\bar{X} = 1.183$, $SD = 1.961$) to the out-group's ($t_{(59)} = -2.97$, $p = 0.004$). This difference was repeated when subjects were asked to indicate their closeness to in-group others after being asked to consider the out-group's performances on the recently completed tasks (rather than their team-mate's). Again, subjects reported greater increased closeness to in-group others when the in-group's performance had been superior ($\bar{X} = 1.183$, $SD = 1.86$) rather than inferior ($\bar{X} = 0.967$, $SD = 1.90$) to the out-group's ($t_{(59)} = -2.09$, $p = 0.041$).

These results are again inconsistent with the self-evaluation maintenance model in that subjects reported increased closeness to in-group (i.e. close) others performing on a relevant dimension, even when the other's performance on such dimensions was better than their own. Consistent with social identity theory, subjects reported significantly greater increased closeness to in-group others when the in-group was superior to the out-group on relevant dimensions than when it was inferior, but contrary to social identity theory subjects

still reported significant increased closeness to in-group others in the latter situation.

Relevance results

The self-evaluation maintenance model predicts that subjects will attempt to reduce the relevance of dimensions on which they are compared to close others, particularly if the close others are likely to be superior to the self on such dimensions. Social identity theory, on the other hand, predicts that subjects will attempt to increase the relevance of dimensions on which the in-group obtains superiority over the out-group and reduce the relevance of dimensions on which the in-group is inferior to the out-group.

When subjects were asked to indicate the relevance of the performance dimensions after being invited to consider their team-mate's performance on those dimensions they significantly reduced the relevance of the dimensions compared to pre-manipulation levels ($\bar{X} = -0.592$, $SD = 1.382$, $t_{(59)} = -3.32$, $p = 0.002$). There was, however, no effect at all of individual inferiority ($\bar{X} = -0.592$, $SD = 1.69$) or superiority ($\bar{X} = -0.592$, $SD = 1.78$) to the in-group other ($t_{(59)} = 0.00$, $p = 1.000$). There was an effect of group distinctiveness on this measure, with subjects reducing the relevance of the performance dimensions less when the in-group achieved superiority over the out-group ($\bar{X} = -0.275$, $SD = 1.61$) than when it "achieved" inferiority ($\bar{X} = -0.908$, $SD = 1.79$) from the out-group ($t_{(59)} = -2.46$, $p = 0.017$). Additionally, relevance was significantly lower than pre-condition levels when the in-group was inferior to the out-group, but was not significantly lower than pre-manipulation levels when the in-group was superior to the out-group. That is, the significant difference in relevance change as a result of in-group distinctiveness came from significant relevance reduction when

the in-group was inferior to the out-group, but no significant change in relevance when the in-group was superior.

When subjects were asked to indicate the relevance of the performance dimensions after being asked to directly consider out-group performance on those dimensions (as opposed to their team-mate's), there was no significant difference in relevance reduction when the in-group achieved superiority over the out-group ($\bar{X} = -0.342$, $SD = 1.61$) compared to when it "achieved" inferiority to the out-group ($\bar{X} = -0.675$, $SD = 1.86$, $t_{(33)} = -1.17$, $p = 0.245$). However, relevance was significantly less than pre-manipulation levels when the in-group was inferior to the out-group, but was not significantly less than pre-manipulation levels when the in-group was superior. That is, although there was no significant difference in relevance change as a result of in-group superiority to the out-group, as found above there was a significant reduction in relevance as a result of in-group inferiority to the out-group but no significant change in relevance as a result of in-group superiority.

These results provide mixed support for the self-evaluation maintenance model. As predicted by the model, subjects did reduce the relevance of dimensions upon which they compared themselves with close others. They did not do so more when the close other was superior rather than inferior to them, however, thus failing to support that aspect of the model. Social identity theory also receives mixed support. In line with the theory subjects significantly reduced the relevance of performance dimensions upon which the in-group was inferior to the out-group. However, they did not increase the relevance of performance dimensions upon which the in-group was superior to the out-group, as the theory would suggest: there was no significant change in relevance following in-group superiority to the out-group.

DISCUSSION

The results of the present study do nothing to support Schiffmann & Wicklund's contention that theories of the effects of between-individual social comparison outcomes such as Tesser's self-evaluation maintenance model make Tajfel's social identity theory redundant. Both the original and the extended versions of the self-evaluation maintenance model failed to receive much support from the present findings. In particular those theories failed to predict: (i) that downward social comparisons would raise self-esteem; and, (ii) that between-group social comparisons outcomes would affect self-esteem above and beyond the effects of between-individual social comparison outcomes, even for between-individual social comparisons within groups.¹³

The "original" version of social identity theory, on the other hand, received considerable support from the present study, particularly in regard to the two main weaknesses of the self-evaluation maintenance model just mentioned in the paragraph above. The only results obtained which were inconsistent with predictions derived from the original version of social identity theory, were: (i) that there was no greater increased closeness when the in-group was superior to the out-group on relevant dimensions than when it was inferior on such dimensions; and, (ii) subjects did not increase the relevance of comparison dimensions on which the in-group was superior to

¹³ The self-evaluation maintenance models are also caused problems by comparisons with superior close others on relevant dimensions not significantly affecting self-esteem; subjects showing significant willingness to aid the performance of close (i.e. in-group) others on relevant comparison dimensions, regardless of whether or not those others were already superior to the subjects concerned; subjects reporting increased closeness to close (i.e. in-group) others performing on relevant comparison dimensions, again regardless of whether or not those close others outperformed the subjects concerned; and the lack of significantly greater relevance reduction by subjects for performance dimensions on which they were outperformed by close (i.e. in-group) others than for dimensions on which they outperformed such close others.

the out-group. These are very minor difficulties, however, especially when weighed against the vast amount of support for the theory.¹⁴

The present results do support the argument that the original version of social identity theory is however "limited" in not adequately taking into account the effects of between-individual social comparison outcomes. The self-esteem results of the present study supporting the original version of social identity theory are clearly also fully supportive of the proposed extended version of social identity theory. Where social identities were salient across-group social comparisons which revealed in-group superiority to the out-group on relevant comparison dimensions raised self-esteem and those which revealed in-group inferiority to the out-group on relevant comparison dimensions lowered self-esteem, whether the across-group comparisons occurred at a genuine between-group level or at the level of individual in-group members being superior or inferior to individual out-group members. In addition, the effects on self-esteem of intragroup between-individual social comparisons were affected by between-group social comparison outcomes.

However, in line with the extended version of social identity theory but not the original version of that theory, the effects on self-esteem of intragroup between-individual social comparisons were also affected by the between-individual social comparison outcomes as well. That is, although personal identities were made as non-salient as possible, the

¹⁴ In addition, with respect to the first point Turner et al. (1984) suggest that both group success and group failure can increase in-group cohesiveness if group members identify with the group and to some extent feel responsible for the group outcome, both of which were likely in the present study. With respect to the second point it can be noted that subjects did not reduce the relevance of dimensions on which the in-group was superior to the out-group: relevance "merely" did not increase from its already relatively high pre-manipulation level.

effects of between-individual social comparison outcomes within groups on self-esteem were determined by an interaction of in-group and individual distinctiveness. Where intragroup between-individual social comparisons revealed that individuals were superior to in-group others on relevant dimensions a significant rise in self-esteem only came about when the in-group was also superior to the out-group on such dimensions, and where intragroup between-individual social comparisons revealed that individuals were inferior to in-group others on relevant dimensions a significant fall in self-esteem only came about when the in-group was also inferior to the out-group on such dimensions. Or, to put the same point the other way around, where the in-group was superior to the out-group on relevant dimensions a significant rise in self-esteem only came about following intragroup between-individual social comparisons when individuals were also superior to in-group others on such dimensions, and where the in-group was inferior to the out-group on relevant dimensions a significant fall in self-esteem only came about following intragroup between-individual social comparisons when individuals were also inferior to in-group others on such dimensions. Thus, as predicted by the extended version of social identity theory, both individual and in-group distinctiveness affect self-esteem, even for intragroup between-individual social comparisons.

This is an important result not only because it suggests a needed modification of the original version of social identity theory, but also because it counts against important aspects of two major contemporary developments of social identity theory, namely Turner's self-categorization theory (Turner et al., 1987) and Brewer's (1991, 1993) optimal distinctiveness model.

Self-categorization theory argues that personal and social identities are mutually inhibitory or functionally antagonistic (Turner et al., 1987: 49, but see also Turner,

1988: 115): when one is salient it reduces the salience of the other. The finding here, though, that the effects on self-esteem of intragroup between-individual social comparisons were determined by an interaction of between-group and between-individual comparison information, strongly suggests that both personal and social identities were salient at the same time, rather than one or the other squeezing the other out, or some sort of flipping between the two.

The optimal distinctiveness model argues that when social identities are salient superior performance by in-group others on relevant dimensions should raise self-esteem, perhaps especially when in-group superiority on such dimensions has been achieved (Brewer, 1993: 9-10; Brewer & Weber, 1994: 268). The present results strongly indicate, however, that the superior performance of in-group others on relevant dimensions lowers self-esteem, even when social identities are very salient and in-group superiority over the out-group is achieved.

The present study, although strongly supportive of the extended version of social identity theory argued for here, does not of course fully test that theory. In particular, the design of the present experiment meant that all across-group comparisons involved "consistent" group-level and individual-level distinctiveness. That is, if the in-group was superior or inferior to the out-group, individual in-group members were also superior or inferior to individual out-group members. The extended version of social identity theory suggests that had in-group distinctiveness been fully crossed with individual distinctiveness (across-groups), then in-group superiority accompanied by individual superiority of in-group members over particular out-group others (as in the present study) would have resulted in greater self-esteem rises than in-group superiority accompanied by individual inferiority of in-group members to particular out-group others. Similarly, in-group inferiority to the out-group would be predicted to

result in greater self-esteem deterioration when accompanied by individual inferiority of in-group members to out-group others than when accompanied by individual superiority to them.¹⁵ Such predictions were not tested in the present study, however.

Another limitation in testing the extended version of social identity theory concerns the lack of dynamic predictions derived from that theory in the present study. Social identity theory makes dynamic predictions in the light of self-esteem needs derived from in-group distinctiveness, but the present study makes it clear that self-esteem can also be affected by individual distinctiveness. If we were to extrapolate social identity theory's dynamic hypotheses to the individual level it could be predicted that following comparisons with relevantly superior others subjects would wish to improve the performance differential in their own favour, would wish to distance themselves from the relevantly superior comparison others, and/or would wish to reduce the relevance of the comparison dimension. These predictions are, of course, identical to those of the self-evaluation maintenance model. Further, we could also predict that following comparisons with relevantly inferior others subjects would wish to improve or protect the performance differential in their own favour, increase their closeness to the relevantly inferior comparison others,¹⁶ and/or increase the

¹⁵ Such predictions would be particularly appropriate when neither between-group and social identity nor between-individual comparison outcomes and personal identity were deliberately made more relevant or important than the other (e.g. when there were prizes for both group and individual performances).

¹⁶ In this instance increased closeness should probably be thought of in terms of increased willingness to be compared with the comparison others, rather than in terms of identifying with them.

relevance of the comparison dimension. These predictions are inconsistent with the self-evaluation maintenance model.¹⁷

The extended version of social identity theory can thus make predictions both at the exclusively between-group and at the exclusively between-individual level. The present study makes it clear, however, that at least some (if not many) situations are neither exclusively between-group nor exclusively between-individual, and that in such situations both in-group and individual distinctiveness can affect self-esteem. Dynamic predictions in such situations are complicated because one has to consider the effects of any dynamic strategy upon both individual and in-group distinctiveness.

Generally speaking, we can predict that subjects will employ the combination of dynamic strategies which seems most likely to optimize the mix of positive in-group and positive individual distinctiveness. When only in-group distinctiveness is important this will result in the predictions derived from the original version of social identity theory, and when only individual distinctiveness is important this will result in the predictions derived from the between-individuals version of social identity theory. When both in-group and individual distinctiveness are important, however, we need to consider both the individual and the in-group differential implications of any dynamic strategy adopted.

In such situations it can be predicted that subjects will wish to aid the performance of close (i.e. in-group) others on relevant dimensions if and only if doing so seems likely to be more beneficial at the level of in-group distinctiveness than it seems likely to be detrimental at the level of individual distinctiveness. Similarly, subjects will wish to avoid

¹⁷ Both of these sets of hypotheses, however, are very similar to those formulated by Major et al. (1991) and deserve further empirical investigation.

aiding (or will wish to hamper) the performance of in-group others on relevant dimensions if and only if doing this seems likely to be more beneficial at the level of individual distinctiveness than it seems likely to be detrimental at the level of in-group distinctiveness.

To illustrate, a person will be willing to aid the performance of in-group others on relevant dimensions when doing so obtains in-group superiority for a presently inferior in-group whilst leaving their individual superiority over in-group others intact. Such a person will not be willing to aid the performance of in-group others on relevant dimensions, however, if doing so will not obtain in-group superiority for a presently inferior in-group, but will result in individual inferiority to in-group others that the person is presently superior to. Similar "performance" predictions could be formulated for all other possible permutations of situations of secure or insecure, positive or negative, and individual or in-group distinctiveness.

In situations in which both individual and in-group distinctiveness are important, closeness predictions are even more complicated than performance ones. This is because individual-level comparisons with in-group others have both individual and between-group distinctiveness implications. Thus, there will be pressures to increase closeness to superior in-group others because of the contribution those others are making to in-group superiority over out-groups (and therefore to positive evaluations of the in-group and to social identity contingent self-esteem), but there will also be pressures to increase distance from those superior in-group others because their superiority suggests both personal (i.e. individual) failure and failure as an in-group member (i.e. with regard to one's own contribution to in-group superiority over the out-group). Similarly, there will be pressures to increase distance from inferior in-group others because of their relative poor contribution to in-group superiority over

out-groups (and therefore to positive evaluations of the in-group and to social identity contingent self-esteem), but there will also be pressures to increase closeness to those inferior in-group others because their inferiority suggests both personal (i.e. individual) success and success as an in-group member (i.e. with regard to one's own contribution to in-group superiority). And, in addition to all these considerations, there will also be pressures to increase closeness to in-group others (and to the in-group) when the in-group is superior to out-groups and pressures to increase distance from in-group others (and from the in-group) when the in-group is inferior to out-groups,¹⁸ regardless of the individual distinctiveness between the self and those in-group others. With all this in mind, it does not seem unduly defeatist to say that at present the closeness predictions from the extended version of social identity theory are difficult to specify and more research is needed in this area.

Relevance predictions are similar in form to performance ones. In situations in which both individual and in-group distinctiveness are important subjects will increase or decrease the relevance of comparison dimensions in such a way as to optimize the combination of positive individual and in-group distinctiveness on relevant comparison dimensions. That is, subjects will increase the relevance of dimensions if and only if the benefits of doing so in terms of positive individual and/or in-group distinctiveness on those dimensions outweigh the cost of doing so in terms of negative individual and/or in-group distinctiveness on those dimensions.

To illustrate, if a person is both superior to in-group others and a member of a superior in-group on given comparison dimensions, they will increase the relevance of those dimensions. Similarly, if a person is inferior to in-group others and a member of an inferior in-group on given

¹⁸ But see Turner et al. (1984), and above.

comparison dimensions, they will decrease the relevance of those dimensions. If a person is inferior to in-group others but is a member of a superior in-group on given comparison dimensions, or if they are superior to in-group others but are members of an inferior in-group on given comparison dimensions (when both individual and in-group distinctiveness are equally important), whether they chose to increase or decrease the relevance of the dimensions will depend on the relative contributions of the two forms of distinctiveness on their self-esteem. That is, if the superiority contributes more positively to their self-esteem than the inferiority detracts from it, they will increase the relevance of the dimensions. If, on the other hand, the inferiority contributes more negatively to their self-esteem than the superiority contributes favourably to it, they will decrease the relevance of the dimensions.

To reiterate, the dynamic predictions from the extended social identity theory will depend upon the relative importance of individual and in-group distinctiveness and will be formulated in terms of the general rule that subjects will employ the combination of dynamic strategies which seems most likely to maximize the positive contributions to their self-esteem derived from positive in-group and/or individual distinctiveness whilst minimizing the negative contributions to their self-esteem derived from negative in-group and/or individual distinctiveness.

There are also a number of methodological limitations associated with the present study. The first concerns the repeated use of "single-shot" measures of self-esteem, performance, closeness and relevance. Such single-shot measures are obviously of dubious reliability, and the repeated use of them must also threaten their validity. It could be objected that rather than actually measuring subjects' self-esteem, etc., the measures used invited subjects to estimate how they "should" feel in response to

particular social comparisons they were invited to make. If such an objection is accepted then the present study is still important in that at the very least it suggests that subjects' "lay theories" are more in line with social identity theory (or theories) than they are in line with the self-evaluation maintenance model (or models). That is, subjects felt that in-group and individual superiority "would" or "should" raise self-esteem and that in-group and individual inferiority "would" or "should" lower self-esteem, even if such changes in self-esteem did not actually occur. The methodological weakness remains, however, and future research could profitably employ more rigorous measures to test hypotheses similar to those addressed in the present study one or a few at a time.

A related problem with the present study is that no attempt was made to control for the various "dynamic" strategies for self-esteem maintenance, protection or enhancement. That is, faced with self-esteem needs subjects in the present study could attempt to meet those needs by either changing their closeness to in-group others or by changing the relevance they attached to the comparison dimensions.¹⁹ This opens up the possibility of explaining failures to support predictions with regard to one "dynamic" variable by claiming that use of another "dynamic" variable obviated subjects' needs to use the former one. It might be claimed, for example, that subjects did not decrease closeness to superior in-group others because they met self-esteem needs by reducing the relevance of the comparison dimensions instead, which negated the need to meet self-esteem needs by reducing closeness. Such ad hoc theorizing is not particularly convincing, particularly as both closeness and relevance were still significantly above their scale mid-points following the false feedback

¹⁹ Subjects could not easily use the "performance" option to meet self-esteem needs, as they were given no opportunity to actually affect the performance of others. They were only asked if they "would have" done so.

manipulation (i.e. subjects still felt close to in-group others and felt the two dimensions relevant). Nevertheless, future research could usefully compare the dynamic predictions of the self-evaluation maintenance model and of social identity theory separately.

Before concluding this chapter it should be made explicit that the purpose of the present study was not to falsify or demonstrate the invalidity of the self-evaluation maintenance model. The purpose was rather more modest in trying to demonstrate that contrary to Schiffmann & Wicklund's claim, when social identities are salient between-group social comparison outcomes have effects not reducible to the effects of between-individual social comparison outcomes. This was found to be the case, and so Schiffmann & Wicklund's claim that theories such as Tesser's self-evaluation maintenance model make Tajfel's social identity theory redundant must be rejected. Such a finding, though, clearly says little or nothing about the validity or the invalidity of Tesser's self-evaluation maintenance model in situations in which only personal identities are salient, i.e. in the domain in which the model is held by Tesser to apply. Nevertheless: (i) the current study's results do present some difficulties for Tesser's model;²⁰ and, (ii) the proposed extended version of social identity theory does make differing predictions to (each version of) the self-evaluation maintenance model where only personal identities are salient, where only social identities are salient, and where both personal and social identities are salient. Clearly a wide-ranging and systematic programme of research would be needed to fully evaluate the relative merits of each theory.

²⁰ The finding that downward social comparisons on relevant dimensions can raise self-esteem is, for example, clearly inconsistent with the self-evaluation maintenance model.

CONCLUSIONS

The most obvious conclusion to be drawn from the present chapter is that there is no foundation for a claim that Tajfel's social identity theory is redundant when an individualistic theory such as Tesser's self-evaluation maintenance model is adopted. In the present study, where subjects social identities were salient, Tesser's model received almost no support whilst Tajfel's theory was strongly supported. In addition, Tajfel's theory predicted the effects of between-group social comparison outcomes, which Tesser's individualistic model is unable to do.

Second, when the between-group differential hypotheses inherent in Tajfel's social identity theory is "extended" to encompass between-individual social comparison outcomes as well, the modified theory can be used to account for all of the self-esteem results obtained in the present study. It has also been argued that the extended social identity theory can provide a complete account of all self-esteem and "dynamic" consequences of social comparison outcomes, at both the between-individual and the between-group levels of analysis, and when both levels of analysis need to be considered together. Thus, if either theory is to supersede the other, it appears that it must be social identity theory which supersedes the self-evaluation maintenance model rather than the other way around, especially when the "extended" version of social identity theory is employed.

There were undoubtedly methodological weaknesses in the present study, however, and considerably more research is needed into the relative (and absolute) strengths and weaknesses of both social identity theory and the self-evaluation maintenance model before it is sensible to call for the abandonment of either. In the next chapter, the final empirical one in this thesis, a number of aspects of social

identity theory receive further attention before evaluating the theory as a whole in the concluding chapter.

CHAPTER 7: STUDY 6 - THE "LOOSE-ENDS" QUESTIONNAIRE

CHAPTER OVERVIEW

The study reported in this chapter employed a questionnaire designed to address a number of issues within social identity theory (Tajfel & Turner, 1979).

Supporters of charitable and of political groups each engaged in mean evaluative in-group preference. This was accompanied by mean out-group derogation, however, only when political group members were asked to evaluate particular out-groups against whom their political in-groups had a tradition of conflict. This suggests that social categorization plus both identification as an in-group member and identification with an in-group (see Chapter 1, pp. 41-42, for this distinction) may be sufficient for in-group preference, but not for out-group derogation.

Positive in-group distinctiveness without a conflict of values was sufficient to provide respondents with positive self-esteem, regardless of whether or not intergroup discrimination was used. Intergroup discrimination per se was not so sufficient. This suggests that Abrams & Hogg (1988) are incorrect to postulate that levels of intergroup discrimination will be positively correlated with subsequent levels of positive self-esteem.

Social identity contingent self-esteem was shown to be affected by movement toward or away from absolute or relational in-group goals, as well as by attainment or non-attainment of positive in-group distinctiveness.

In general, regardless of respondents' prior levels of social identity contingent self-esteem, charitable group supporters were universally unwilling to employ intergroup discrimination, while political group supporters were universally unwilling not to. This suggests that "conflicts of values" play an important mediating role between low or threatened self-esteem and subsequent use of intergroup discrimination.

INTRODUCTION

The purpose of this chapter is to report the results of a questionnaire designed to address a number of issues within social identity theory (e.g. Tajfel, 1978a). These issues, several of which have already been raised earlier in the thesis, include: (i) the nature of intergroup discrimination; (ii) how in-groups are evaluated; and (iii) the two corollaries of the self-esteem hypothesis within social identity theory.

The nature of intergroup discrimination

Social identity theory suggests that group members often employ intergroup discrimination in order to bolster evaluation of the in-group and thereby enhance their social identity and self-esteem. Such in-group bias tends to take the form of in-group preference rather than out-group derogation, however, such that both in-group and out-group are dealt with in a relatively positive fashion, with the in-group merely receiving a greater degree of positive treatment than the out-group (Brewer, 1979). In Brown et al.'s (1986) study of intergroup discrimination in a paper factory, for example, sub-groups within the factory tended to evaluate all sub-groups within the factory positively, even though each tended to evaluate their own sub-group more positively than they evaluated each of the others. This relatively "benign" form of "intergroup discrimination" perhaps restricts the validity of employing social identity theory to harsher or more "vicious" forms of intergroup discrimination: i.e. those in which genuine out-group derogation occurs (as in, for example, Sherif's summer camp studies, e.g. Sherif & Sherif, 1953, 1967, or in the St. Pauls riots investigated by Reicher and his colleagues, e.g. Reicher & Potter, 1985; Reicher, 1984, 1987). The theoretical and empirical question which such considerations raise is: on what occasions will intergroup

discrimination take the form of, or be accompanied by, out-group derogation?

There are two related possible answers to this question. The first, similar to propositions within Sherif's "realistic conflict theory" (Campbell, 1965; Sherif & Sherif, 1967), is that strong intergroup discrimination will occur when one group is negatively dependent upon another for the achievements of its own goals. That is, when the progress of one group toward its goals is perceived by group members as likely to be enhanced by out-group derogation, such out-group derogation will take place. The second possibility is that over time particular groups develop and adopt a culture of strong discrimination against particular other groups. That is, for whatever reason (perhaps because of a history of perceived negative dependence), some groups develop particular out-groups against whom they more or less standardly employ out-group derogation.¹

Such reasoning, if well grounded and empirically supported, might justify modifying and extending the standard in-group/out-group distinction made within social identity theory, such that the phrase "out-group" could be restricted to those particular groups toward which particular in-groups have a "culture" or a "tradition" of hostility and/or against whom the in-group has a particular interest in competing against at a given moment. Other groups, against whom in-groups have little or no interest in competing, might then be referred to simply as non-in-groups. If such a distinction were accepted then social identity theory might be held to predict that: (i) in-group members will employ only "weak" intergroup discrimination (i.e. in-group preference without out-group derogation) against non-in-groups and their

¹ This possibility is also consistent with realistic conflict theory, which says that hostile in-group attitudes and behaviour toward particular out-groups can exist far beyond the circumstance which originally gave rise to such hostility, and also with the work of some social identity theorists, who stress the need to take account of the "ideology" of groups, particularly with respect to intergroup relations over time between in-groups and particular out-groups (e.g. Billig, 1975).

members; and, (ii) in-groups will employ both "weak" (i.e. in-group preference) and "strong" (i.e. out-group derogation) against out-groups.²

How in-groups are evaluated

Social identity theory claims that in-groups are evaluated primarily or exclusively in terms of their position relative to other groups on dimensions of value to the in-groups (e.g. Tajfel & Turner, 1979: 40). That is, in-groups are usually evaluated positively if they are perceived as superior to (i.e. positively distinct from) relevant comparison groups on such dimensions and negatively if they are perceived as inferior to (i.e. negatively distinct from) them on such dimensions.

Hinkle & Brown (1990) draw upon Abrams (1984) to suggest that in-group evaluation may be affected by "objective" as well as by between-group social comparison outcomes (see also Brown et al, 1991, 1992). That is, groups may sometimes evaluate their performances or situations (and thereby themselves) in relation to some "objective" criterion (e.g. where a political party evaluates itself in terms of how close it is to achieving full employment), instead of or as well as by comparing the in-group with relevant comparison groups (e.g. where a political party evaluates itself in terms of whether or not they are closer to achieving full employment than another political party). In other words, "objective" evaluation occurs by an in-group comparing its actual situation against a "possible" (e.g. past, future, ideal, etc.) in-group situation, rather than by an in-group comparing its situation against a particular out-group situation.

² Another way of making the same point is to employ terms used in Chapter 3 whereby in-groups show relational intergroup discrimination in favour of the in-group toward non-in-groups, but employ both relational and absolute intergroup discrimination against out-groups.

If individual-level theorizing about the effects of comparison outcomes on self-esteem can be extrapolated to the effects of group-level comparison outcomes on in-group evaluation (as Tajfel believed, e.g. Tajfel, 1981a: 337), then Tajfel himself raised this possibility, among several others. Tajfel (1978b: 9) wrote that:

a person's self-image is essentially based on certain kinds of comparisons, and it consists to a large extent of the outcomes of these comparisons. The comparisons may go in a number of directions such as: one's expectations, wishes or hopes as related to [one's] achievement, actual or subjectively assessed; a person's past as related to [their] present; one's characteristics (again, objectively ascertainable or subjectively assessed) as related to those of other people with whom meaningful comparisons can be made.

Extrapolating to the group-level, this quote suggests that in-group evaluation can be affected: by comparing an in-group's present situation to its previously expected, wished for, or hoped for future situation; by comparing an in-group's present situation with a past in-group situation; and/or, by comparing an in-group's present situation with the present situation of relevant comparison groups. Only the latter of these possibilities precludes "objective" evaluation, and this is the only option which has been taken seriously within social identity theory research.

If such reasoning is valid, rather than the simple "superiority good, inferiority bad" assertion dominant within social identity theory, it can be predicted that levels of in-group evaluation (and hence social identity and social identity contingent self-esteem) will benefit from perceived achievement of or movement toward any objective, temporal or social "in-group goal" (of which positive in-group distinctiveness may be only one), and will suffer from failure to achieve or movement away from such goals.

The self-esteem hypothesis within social identity theory

Hogg & Abrams (1990: 33) identify two corollaries of social identity theory's self-esteem hypothesis. These are: (i) that "successful intergroup discrimination...elevates self-esteem"; and, (ii) that "depressed or threatened self-esteem promotes intergroup discrimination". As has already been argued in Chapter 5, these corollaries inadequately specify social identity theory's self-esteem hypothesis.

Leaving aside issues of in-group salience, out-group relevance and comparison dimension importance, with respect to the first corollary of the self-esteem hypothesis social identity theory says that *positive in-group distinctiveness unaccompanied by a conflict of values* results in positive evaluation of the in-group and a positive contribution being made to in-group members' social identity contingent self-esteem. Thus, intergroup discrimination which successfully achieves or increases such positive in-group distinctiveness will elevate that aspect of overall self-esteem. Intergroup discrimination which maintains or protects such positive in-group distinctiveness will "only" make that aspect of self-esteem more *secure*. Intergroup discrimination will only reliably elevate self-esteem, therefore, if it: (i) *achieves or increases positive in-group distinctiveness* (as opposed to making it more secure); and, (ii) does so without a conflict of values. Even then, it is only *social identity contingent* self-esteem which is elevated, not necessarily "overall" self-esteem.

Again leaving aside issues of in-group salience, out-group relevance and comparison dimension importance, with respect to the second corollary of the self-esteem hypothesis social identity theory says that *depressed* (i.e. "negative") or *threatened* (i.e. "insecure") *social identity contingent* self-esteem will promote intergroup discrimination if and only if: (i) there are "cognitive alternatives" available (i.e. the discrimination is likely to result in new, enhanced, or more

secure positive in-group distinctiveness); (ii) the discrimination is not perceived as likely to result in a conflict of values; and, (iii) the discrimination does not preclude more attractive routes to meeting self-esteem needs (e.g. social mobility or exit). Depressed or threatened self-esteem will not reliably promote intergroup discrimination if: (i) the esteem is not social identity contingent; (ii) cognitive alternatives are not available; (iii) a conflict of values is involved; and/or, (iv) discrimination precludes more attractive routes to self-esteem improvement.

The present study

In the present study it was assumed that there may be essential differences between the consequences of belonging to and/or supporting two potentially rather different "types" of in-groups: one charitable and one political. Political group affiliation tends to occur within essentially competitive intergroup arenas. That is, pairs or sets of political parties "traditionally" have negatively interdependent goals and to a large extent (although not necessarily exclusively) are evaluated in terms of their successes or failures relative to other particular political groups. Thus, social identity theory's hypotheses are perhaps most likely to be supported with respect to such groups (e.g. Kelly, 1988: 330). Charitable group affiliation seems likely to be somewhat different in that charitable groups: (i) tend to be rather more focused toward specific "objective" goals (e.g. raising money, recruiting members, reducing poverty, etc.); and, (ii) are not "essentially" or "traditionally" in conflict with particular comparison groups. In the terminology suggested above, charitable groups, unlike political groups, do not tend to have out-groups, only non-in-groups. Further, charitable groups are perhaps more (although not necessarily exclusively) reliant on non-socially comparative evaluation than are political groups. These assumed differences allow a number of hypotheses to be formulated and tested.

Additionally, respondents' evaluation of various hypothetical situations involving between-sex differences were investigated (see below).

Hypotheses

The nature of intergroup discrimination

- H1a Supporters of charitable groups will employ only "weak" mean intergroup discrimination against comparison groups (i.e. in-group preference, with in-groups receiving the best and positive mean evaluation, but with comparison groups still receiving neutral or positive mean evaluation).
- H1b Supporters of political groups will employ both "weak" (i.e. in-group preference) and "strong" (i.e. out-group derogation) mean intergroup discrimination against out-groups against whom the in-groups "traditionally" compete, but will employ only "weak" mean intergroup discrimination against other comparison groups.

In-group evaluation

- H2a Increasing progress toward or exceeding of objective or relational in-group goals will result in increasingly positive (or less negative) mean evaluation of the in-group.
- H2b Increasing movement away from objective or relational in-group goals will result in increasingly negative (or less positive) mean evaluation of the in-group.
- H2c Achievement of objective or relative in-group goals will always result in positive mean evaluation of the in-group.
- H2d Non-achievement of objective or relative in-group goals will result in mean negative evaluation of the in-group when and only when that non-achievement is accompanied by (or indicates) movement away from those goals.

The self-esteem hypothesis: corollary 1

- H3a Respondents' mean self-esteem will be positive in situations where positive in-group distinctiveness is achieved without a conflict of values.
- H3b Respondents' mean self-esteem will be negative in situations where a conflict of values occurs and positive in-group distinctiveness is not achieved.

H3c Respondents' mean self-esteem will not be significantly different from zero in situations where (i) no conflict of values occurs but positive in-group distinctiveness is not achieved, or (ii) positive in-group distinctiveness is achieved but is accompanied by a conflict of values.

The self-esteem hypothesis: corollary 2

H4a Respondents will be willing to employ intergroup discrimination when in-group distinctiveness is insecure (i.e. when cognitive alternatives are available) and exit is unavailable.

H4b Respondents will not be willing to employ intergroup discrimination when in-group distinctiveness is secure.

H4c When in-group distinctiveness is insecure and exit is available, respondents who are members of superior groups will employ intergroup discrimination but members of inferior groups will not.

H4d Exit will be employed, when available, considerably more frequently by members of securely inferior in-groups than by members of positive or insecurely negative groups (with members of securely positive groups not employing exit at all).

METHOD

Respondents and procedure

Over two days approximately 120 first-year psychology undergraduates at Keele University were approached during tutorials and asked if they would volunteer to take away and complete questionnaires concerning "the effects of group membership", which would to be collected during the following week's tutorial. It was stressed that the questionnaire required quite a lot of thought and effort and that volunteers should complete it alone when they had at least half an hour to spare.

101 respondents returned complete or near-complete questionnaires. Of these, 23 were male, 76 were female and 2 did not indicate their sex. Respondents' ages ranged from 18-49

years old, with a mean of 22.6 (SD = 6.1, mode = 19, median = 20).

Materials

A questionnaire was designed to address each of the issues of interest and was piloted on two separate samples: one of health care professionals (primarily nurses) attending courses held at Keele University, and another of students attending an Open University Summer School, again at Keele University. Minor amendments to and clarifications of particular items on the questionnaire were made after each pilot, with the major change being to more or less abandon a third section of the pilot questionnaires which dealt with sex category membership. The structure of the final questionnaire used was as follows (see Appendix 14).

Section A dealt with charitable group affiliation; Section B dealt with political group affiliation; and Section C contained "miscellaneous" questions, including one (Question C1) dealing with sex category membership. In each of the first two main sections questions followed the same sequence.

Questions A1 and B1 asked about psychological affiliation (or non-affiliation) with various charitable and political groups, respectively. Questions A2 and B2 asked respondents to indicate "how they felt" about individual members each of those groups on 15-point scales, ranging from -7 = very negative to +7 = very positive. These questions were designed to investigate the nature of intergroup discrimination by supporters of each type of group against various comparison groups and/or their members.³

³ The response scales included negative numbers to allow the possibility of respondents derogating target groups and/or their members.

Questions A3, A4, B3 and B4 asked respondents to indicate "how they would feel" (on identical scales as before) in a variety of hypothetical situations. Question A3 informed respondents that the previous year their favourite charity had achieved only 70% of its annual fund-raising target and then asked them to evaluate situations in which this year the same charity had achieved 50%, 70%, 90%, 100% and 120% of its annual fund-raising target. Question A4 informed respondents that the previous year their favourite charity had achieved 140% of its annual fund-raising target and then asked them to evaluate situations in which this year the same charity had achieved 90%, 100%, 120%, 140% and 160% of its annual fund-raising target. These questions were designed to investigate evaluation of situations in which achievement or non-achievement of, and movement toward or movement away from, in-group goals could be "objectively" assessed.

Question B3 informed respondents that at the last election the political party they supported was runner-up and received 500 votes less than the winning party. They were then asked to evaluate situations in which at the latest election their party received 1000, 500 or 250 votes less than the winning party, the same number of votes as the (jointly) winning party, or 250 votes more than the nearest runner-up. Question B4 informed respondents that at the last election the political party they supported won, and received 500 votes more than its nearest rival. They were then asked to evaluate situations in which at the latest election their party received 250 votes less than their nearest rival (who won the election), the same number of votes as them (joint winners of the election), or 250, 500 or 1000 votes more than them (with their own party winning). These questions were designed to investigate evaluation of situations in which achievement or non-achievement of, and movement toward or movement away from, in-group goals could be "relationally" (i.e. socially comparatively) assessed.

Questions A5, B5 and C1 investigated corollary 1 of social identity theory's self-esteem hypothesis by asking respondents

how they would feel (using the same scales as before) in various situations involving either positive or negative in-group distinctiveness, use or non-use of intergroup discrimination, and the presence or absence of a conflict of values.

Questions A6 and B6 investigated corollary 2 of social identity theory's self-esteem hypothesis by asking respondents how likely they would be to employ intergroup discrimination (on 15-point scales ranging from -7 = definitely not to +7 = definitely) in various situations of either positive or negative in-group distinctiveness, and where cognitive alternatives and/or exit either were or were not available. (Where exit was available respondents could opt for it by placing an "X" instead of a number between -7 and +7 in response to that particular item.)

Questions C2 and C3 asked respondents for their sex and age, respectively. Questions C4 and C5, respectively, asked respondents to indicate their favourite charitable and political groups (open-ended).

"Question" C6 invited respondents to make any comments or criticisms they wished on any aspect of the questionnaire and/or clarifications to any of the answers they had given.

RESULTS

The nature of intergroup discrimination

Table 7.1 below shows mean evaluation of various charitable groups (from Question A2a-g), both for the sample overall and according to respondents' self-professed favourite charity (with standard deviations shown in brackets).⁴

Eval of	Favourite charity							Mean n=101
	OXFAM n=9	RSPCA n=10	G'Pce n=9	AI n=7	Shelt n=6	Other n=28	None n=9	
OXFAM	6.33 (0.7)	3.90 (2.0)	3.67 (2.3)	2.14 (4.2)	4.00 (2.4)	4.46 (3.0)	3.67 (2.2)	4.15 (2.7)
RSPCA	5.33 (2.2)	6.90 (0.3)	4.00 (2.1)	4.43 (1.9)	2.83 (1.9)	5.57 (1.8)	4.56 (2.6)	5.10 (2.3)
G'Pce	4.47 (2.3)	2.40 (3.8)	5.67 (2.3)	4.71 (2.0)	4.67 (2.3)	3.57 (2.8)	3.44 (2.7)	3.77 (3.0)
AI	4.44 (2.7)	4.80 (1.4)	4.33 (2.3)	6.43 (0.5)	3.83 (1.7)	3.96 (2.6)	3.78 (2.5)	4.24 (2.4)
Shelt	4.78 (2.7)	3.30 (4.6)	3.56 (3.0)	4.57 (1.8)	6.67 (0.5)	4.61 (2.3)	3.56 (2.5)	4.38 (2.7)
MIND	3.67 (2.9)	3.00 (4.3)	2.67 (2.6)	2.57 (2.8)	2.50 (2.1)	3.32 (3.2)	2.67 (2.9)	3.38 (3.0)
BHF	4.47 (2.2)	4.60 (2.3)	2.89 (3.1)	4.29 (2.3)	3.83 (1.2)	4.61 (2.3)	3.44 (2.7)	4.24 (2.2)

Key: G'Pce - Greenpeace; AI - Amnesty International; Shelt - Shelter; BHF - British Heart Foundation.
Evaluation on a 15-point scale from -7 = Very Negative to +7 = Very Positive

Table 7.1: Mean evaluation of various charitable groups by favoured charitable group

⁴ As indicated on question C5. The evaluations of subjects whose favourite charitable group was MIND (n=3) or the British Heart Foundation (n=2), or who indicated more than one favourite charity (n=5), or who did not respond to this item (n=6) are included in the overall means (right-hand column) but are not given separate columns. Also, it should be noted that similar analyses were calculated on the basis of subjects' responses on question A1a-g, and very similar results were obtained. These results are not reported here (i) to avoid redundancy, and (ii) because the results shown are somewhat easier to interpret.

A 7 X 7 MANOVA with favourite charity as a between-subjects factor, the charitable groups to be evaluated as a within-subjects factor, and evaluation of those charities as the dependent variable, revealed no significant between-subjects main effect ($F_{(6, 71)} = 0.56, p = 0.762$).⁵ Multivariate tests, however, revealed both a significant within-subjects main effect (Hotellings $p = 0.000$) and a significant interaction (Hotellings $p = 0.001$).

Inspection of means in the right-hand column of Table 7.1 shows that on average the RSPCA received the most favourable evaluation overall, and MIND and Shelter received the least favourable overall evaluation. The bold figures in the main body of Table 7.1 show that respondents evaluated their own favourite charity (where this was possible) positively, and tended to evaluate them more favourably than they evaluated other target charities. No target charity, however, received negative (or even neutral) mean evaluation from respondents who indicated favouring a particular charity. Thus, Hypothesis A1a was fully supported: supporters of particular charitable groups employed only "weak" mean intergroup discrimination against comparison groups (i.e. in-group preference: with in-groups receiving the best and positive mean evaluation, but with comparison groups still receiving neutral or positive mean evaluation).

It can also be noted that all target charities also received positive mean evaluation from respondents who explicitly said that they had no favourite charity.

Table 7.2 shows mean evaluation (with standard deviations shown in brackets) of various political parties (from Question B2a-g),

⁵ Within-subjects ANOVA designs can only be calculated within the MANOVA procedure of the statistic package used (SPSS).

both for the sample overall and according to respondents' self-professed favourite political party.⁶

Eval of	Favoured political party					Mean n=101
	Labour n=33	Lib-Dem n=16	Con n=8	Green n=7	None n=24	
Comm	-1.55 (3.5)	-3.38 (3.2)	-2.88 (2.8)	-2.57 (3.6)	-2.75 (3.6)	-2.66 (3.5)
Labour	4.94 (2.6)	0.69 (2.2)	-1.13 (3.6)	1.43 (2.9)	-0.13 (3.3)	1.77 (3.7)
Lib-Dem	1.73 (2.7)	4.50 (1.5)	1.25 (1.8)	3.57 (2.0)	1.25 (2.3)	2.13 (2.5)
Con	-4.94 (2.7)	-0.75 (3.3)	2.75 (1.8)	-3.43 (2.1)	-1.18 (3.4)	-1.94 (3.8)
BNP	-6.49 (1.6)	-4.31 (3.8)	-2.25 (4.2)	-5.29 (3.3)	-4.08 (3.3)	-5.12 (3.1)
Green	1.58 (3.4)	1.38 (3.4)	1.50 (3.2)	5.86 (0.9)	2.50 (2.8)	2.80 (3.4)
MRL	-2.09 (3.3)	-2.88 (3.5)	-1.00 (2.6)	-0.29 (1.0)	-1.50 (2.9)	-1.83 (3.3)

Key: Comm - British Communist Party; Labour - Labour Party; Lib-Dem - Liberal Democrat Party; Con - Conservative Party; BNP - British National Party; MRL - Monster Raving Loony Party. Evaluation on a 15-point scale from -7 = Very Negative to +7 = Very Positive

Table 7.2: Mean evaluation of various political parties by favoured political party

A 5 X 7 MANOVA with favourite political party as a between-subjects factor, the political parties to be evaluated as a within-subjects factor, and evaluation of those parties as the dependent variable, revealed no significant between-subjects main effect ($F_{(4, 83)} = 0.76, p = 0.553$). Multivariate tests, however,

⁶ As indicated on question C5. No subjects indicated that they favoured either the British Communist Party, the British National Party, or the Monster Raving Loony Party. The evaluations of subjects who favoured more than one political party (n=8), who favoured a political party not examined here (n=1), or who did not respond to item C5 (n=6) are included in the overall means (right-hand column) but are not given separate columns. As before, similar analyses calculated on responses to question bla-g gave similar results and are not reported.

revealed both a significant within-subjects main effect and a significant interaction (Hotellings $p = 0.000$ for each).

Inspection of means in the right-hand column of Table 7.2 show that on average the Labour Party and the Green Party received the most favourable evaluations overall, and the British National Party received the least favourable overall evaluation. The bold figures in the main body of Table 2 show that respondents' evaluation of their own favoured political group (where such evaluation was possible) was positive and tended to be more favourable than their evaluation of other target political groups.

Most importantly for present purposes, though, and in marked contrast to the results shown in Table 7.1, Table 7.2 shows considerable evidence of negative mean evaluation of particular target groups by supporters of particular favoured groups. Labour party supporters, for example, made clearly negative mean evaluation of the Conservative and British National parties, while Conservatives made (less pronounced) negative mean evaluation of the Labour and British National parties. Thus, strong support was found for Hypothesis 1b: supporters of political groups employed both "weak" (i.e. in-group preference) and "strong" (i.e. out-group derogation) mean intergroup discrimination against out-groups against whom the in-groups "traditionally" compete, but employed only "weak" mean intergroup discrimination against other comparison groups.

It can also be noted that respondents who explicitly indicated that they favoured no particular political party made negative mean evaluation of certain target parties (e.g. the British National Party).

How in-groups are evaluated

Question A3

Percentage of annual target reached	Evaluation
50%	-1.48 (3.95)
70%	0.90 (3.45)
90%	3.40 (2.92)
100%	5.25 (2.47)
120%	6.07 (2.56)

Evaluation on a 15-point scale from -7 = Very Negative to +7 = Very Positive.

Table 7.3: Mean evaluation according to percentage of charitable in-group's annual target reached, when 70% of the annual target was reached the previous year

Table 7.3 shows that as the in-group's success in meeting its annual fund-raising target improved, so too did respondents' mean evaluation of the situation (from Question A3). A MANOVA with the various situations to be evaluated as a within-subjects factor and evaluation of those situations as the dependent variable revealed a significant multivariate difference between the mean evaluation (Hotellings $p = 0.000$). Further, a planned polynomial contrast revealed a significant linear trend, such that as in-group success increased, so too did respondents' evaluation ($F_{(1,99)} = 233.36, p = 0.000$).⁷

In situations where the in-group both achieved its annual target and improved on its previous year's performance, mean evaluation of those situations was significantly above zero by two-tailed one-sample t-tests ($p < 0.001$ in each case).

⁷ There were also significant quadratic ($F_{(1,99)} = 30.86, p = 0.000$) and cubic trends ($F_{(1,99)} = 6.88, p = 0.010$). However, examination of the sums of squares clearly indicated that the linear trend (SS = 3896.68) explained far more of the variance than did either the quadratic (SS = 96.73) or the cubic (SS = 12.77) trends. It can be noted that this result is particularly striking as the percentage increase in the in-group's success in reaching its annual target was not an interval scale increase.

Where the in-group did not achieve its annual target but improved upon its performance the previous year (i.e. when the in-group achieved 90% of its annual target), mean evaluation of the situation was also significantly greater than zero ($t_{(100)} = 11.66$, $p < 0.001$).

If the in-group failed both to reach its annual target and to improve on the previous year's performance (i.e. when the in-group again reached only 70% of its annual target), mean evaluation of the situation was still significantly (if only slightly) positive ($t_{(100)} = 2.63$, $p = 0.01$).

When the in-group both failed to reach its annual target and its performance worsened compared to the previous year's performance (i.e. the in-group achieved only 50% of its target), mean evaluation of the situation was significantly negative ($t_{(100)} = -3.76$, $p < 0.001$).

Question A4

Percentage of annual target reached	Evaluation	
90%	0.78	(4.06)
100%	2.86	(3.31)
120%	4.07	(2.80)
140%	5.04	(2.72)
160%	5.88	(2.70)

Evaluation on a 15-point scale from -7 = Very Negative to +7 = Very Positive.

Table 7.4: Mean evaluation according to percentage of charitable in-group's annual target reached, when 140% of the annual target was reached the previous year

Table 7.4 again clearly shows that as in-group success in meeting its annual fund-raising target improved, so too do respondents'

mean evaluation of the situation (Hotellings $p = 0.000$, linear contrast $F_{(1,100)} = 131.99$, $p = 0.000$).⁸

Where the in-group both reached its annual target and exceeded its previous year's performance (i.e. the in-group achieved 160% of its annual target), mean evaluation of the situation was significantly positive by a two-tailed one-sample t-test ($t_{(100)} = 21.90$, $p < 0.001$).

If the in-group reached its annual target but its performance did not change compared to the previous year (i.e. the in-group again achieved 140% of its annual target), mean evaluation of the situation was again significantly above zero ($t_{(100)} = 18.62$, $p < 0.001$).

When the in-group achieved its annual target but had inferior performances to those obtained the previous year (i.e. where the in-group achieved 100% or 120% of its annual target), mean evaluation of the situations was still significantly positive ($p < 0.001$ in each case).

Where the in-group both failed to achieve its annual target and its performance worsened compared to the previous year (i.e. where the in-group achieved only 90% of its annual target), mean evaluation of the situation was not (quite) significantly different from zero ($t_{(100)} = 1.94$, $p = 0.055$).

Further, mean evaluation of the situation when the in-group achieved 90% of its target after it had achieved 140% of its target the previous year was significantly lower than when the in-group achieved the same percentage of its annual target after

⁸ Again there were significant quadratic ($F_{(1,100)} = 26.85$, $p = 0.000$) and cubic ($F_{(1,100)} = 7.53$, $p = 0.007$) trends, but again examination of the sums of squares makes it clear that the linear trend is the one explaining the majority of the variance ($SS_{lin} = 1547.03$, $SS_{quad} = 53.09$, $SS_{cub} = 5.57$). As above, it can be noted that this result is particularly striking as the percentage increase in the in-group's success in reaching its annual target was not an interval scale increase.

achieving only 70% of its target the previous year ($\bar{X}90\%_{(post\ 140\%)} = 0.78$, $\bar{X}90\%_{(post\ 70\%)} = 3.40$, $t_{(99)} = 6.90$, two-tailed $p = 0.000$). Similarly, mean evaluation of the situation when the in-group reached its annual target was significantly lower when the in-group had exceeded its target the previous year than when it had failed to reach the previous year's target ($\bar{X}100\%_{(post\ 140\%)} = 2.86$, $\bar{X}100\%_{(post\ 70\%)} = 5.25$, $t_{(100)} = 8.45$, two-tailed $p = 0.000$). Finally, mean evaluation of the situation was significantly lower after the in-group exceeded its annual target when the in-group had failed to equal the previous year's surplus than when it had clearly converted the previous year's deficit into an excess in the current year ($\bar{X}120\%_{(post\ 140\%)} = 4.07$, $\bar{X}120\%_{(post\ 70\%)} = 6.07$, $t_{(100)} = 7.40$, two-tailed $p = 0.000$).

Question B3

Number of votes relative to previously winning party	Evaluation
1000 less	-3.97 (2.95)
500 less	-2.06 (3.08)
250 less	-0.41 (3.67)
same	3.11 (3.19)
250 more	5.88 (1.72)

Evaluation on a 15-point scale from -7 = Very Negative to +7 = Very Positive.

Table 7.5: Mean evaluation according to political in-group's electoral performance, when it had been beaten by the winning party by 500 votes the previous year

Table 7.5 shows that respondents' mean evaluation of the situation increased with improving in-group performance relative to the previously winning party. A MANOVA with planned polynomial contrasts confirmed this linear trend as significant (Hotellings $p = 0.000$, linear $F_{(1,99)} = 609.57$, $p = 0.000$).⁹

⁹ There was also a significant quadratic (but not cubic) trend ($F_{(1,99)} = 12.23$, $p = 0.001$), but the relevant sums of squares revealed that the linear trend ($SS_{lin} = 6185.17$) explained far more of the variance than the quadratic one ($SS_{quad} = 92.06$).

Where the in-group achieved positive in-group distinctiveness and improved on the previous election's performance (i.e. where the in-group obtained 250 votes more than its rival), mean evaluation of the situation was significantly above zero by two-tailed one-sample t-test ($t_{(99)} = 36.24, p < 0.001$).

If the in-group failed to achieve positive in-group distinctiveness but improved on its previous performance, mean evaluation of the situation was also significantly positive ($t_{(99)} = 9.74, p < 0.001$), as long as the in-group was not negatively distinct from the comparison group (i.e. where the in-group obtained the same number of votes as the jointly winning party).

When the in-group improved on its previous performance but remained negatively distinct (i.e. where the in-group obtained 250 votes less than the winning party), mean evaluation of the situation was not significantly different from zero ($t_{(99)} = -1.12, p = 0.266$).

Where the in-group failed to achieve positive in-group distinctiveness and its performance did not change compared to the previous election (i.e. where the in-group again obtained 500 votes less than the winning party), mean evaluation of the situation was significantly negative ($t_{(99)} = -6.68, p < 0.001$), as was the case when the in-group both failed to achieve positive in-group distinctiveness and had a worsened performance relative to the previous election (i.e. where the in-group obtained 1000 votes less than the winning party) ($t_{(99)} = -13.45, p < 0.001$).

Again, it can be noted that this result is particularly striking as the percentage increase in the votes received by the in-group was not an interval scale increase.

Question B4

Number of votes relative to previously beaten party	Evaluation	
250 less	-3.60	(3.33)
same	0.75	(3.57)
250 more	4.55	(2.12)
500 more	5.27	(1.68)
1000 more	6.09	(1.26)

Evaluation on a 15-point scale from -7 = Very Negative to +7 = Very Positive.

Table 7.6: Mean evaluation according to political in-group's electoral performance, when it had beaten its nearest rival party by 500 votes the previous year

Table 7.6 again clearly demonstrates improving respondent mean evaluation with improving in-group performances relative to its previous nearest rival (Hotellings $p = 0.000$, linear $F_{(1,98)} = 646.38$, $p = 0.000$).¹⁰

The in-group both achieving positive in-group distinctiveness and improving its performance from the previous election (i.e. where the in-group obtained 1000 votes more than its rival) resulted in a significantly positive mean evaluation of the situation by a two-tailed one-sample t-test ($t_{(98)} = 48.01$, $p < 0.001$).

The in-group achieving positive in-group distinctiveness and equalling its previous performance (i.e. the in-group again obtaining 500 votes more than its rival) also resulted in a significantly positive mean evaluation of the situation ($t_{(98)} = 31.17$, $p < 0.001$).

¹⁰ Again there was a significant quadratic (but not cubic) trend ($F_{(1,98)} = 179.77$, $p = 0.000$), but again examination of the relevant sums of squares revealed that the linear trend ($SS_{lin} = 5654.50$) explained much more of the variance than the quadratic trend ($SS_{quad} = 724.39$).

Where the in-group's performance worsened from the previous election but positive in-group distinctiveness was still obtained (i.e. where the in-group received 250 votes more than its rival), mean evaluation of the situation was still significantly positive ($t_{(98)} = 21.38, p < 0.001$).

Where the in-group performance worsened from the previous election and positive in-group distinctiveness was not achieved the situation received a significantly positive mean evaluation as long as the in-group was not negatively distinct (i.e. where the in-group received the same amount of votes as the jointing winning party) ($t_{(98)} = 2.08, p < 0.05$). Where the in-group performance worsened from the previous election and the in-group was negatively distinct, however, (i.e. where the in-group obtained 250 votes less than the winning party), mean evaluation of the situation was significantly negative ($t_{(98)} = -10.76, p < 0.001$).

Further, obtaining 250 votes less than the nearest rival resulted in significantly lower evaluation when the in-group had previously beaten the out-group than when the in-group's negative in-group distinctiveness was nevertheless an improvement on previous between-group comparison outcomes ($\bar{X}_{(250 \text{ less: previously } 500 \text{ more})} = -3.60, \bar{X}_{(250 \text{ less: previously } 500 \text{ less})} = -0.41, t_{(98)} = 8.92, \text{ two-tailed } p = 0.000$). Similarly, obtaining the same amount of votes as the out-group resulted in significantly poorer evaluation when the in-group had previously beaten the out-group than when the out-group had previously beaten the in-group ($\bar{X}_{(\text{same: previously } 500 \text{ more})} = 0.75, \bar{X}_{(\text{same: previously } 500 \text{ less})} = 3.11, t_{(98)} = 7.29, \text{ two-tailed } p = 0.000$). Finally, the in-group obtaining 250 votes more than its nearest rival resulted in significantly lower evaluation when the in-group had previously beaten the out-group by 500 votes than when the out-group had previously beaten the in-group by the same amount ($\bar{X}_{(250 \text{ more: previously } 500 \text{ more})} = 4.55, \bar{X}_{(250 \text{ more: previously } 500 \text{ less})} = 5.88, t_{(98)} = 7.46, \text{ two-tailed } p = 0.000$).

This pattern of results provides strong support for hypotheses H2a to H2c. In line with hypothesis H2a increasing progress toward or exceeding of objective or relative in-group goals resulted in increasingly positive (or less negative) mean evaluation of the in-group; in line with hypothesis H2b increasing movement away from objective or relative in-group goals resulted in increasingly negative (or less positive) mean evaluation of the in-group; and, in line with hypothesis H2c achievement of objective or relative in-group goals always resulted in positive mean evaluation of the in-group.

It was not the case, however, that non-achievement of objective or relative in-group goals only resulted in negative mean evaluation of the in-group when that non-achievement was accompanied by movement away from those goals. In the case of the political in-groups striving to achieve the relational goal of positive in-group distinctiveness, negative in-group distinctiveness resulted in negative mean in-group evaluation when accompanied by no progress toward or movement away from that goal (i.e. when the in-group remained as negatively distinct as it had previously been). Thus, against hypothesis H2d, in-groups were not only negatively evaluated when non-achievement of relational in-group goals was accompanied by movement away from those goals.¹¹

Also, failure to achieve the relational goal of positive in-group distinctiveness did not result in negative mean evaluation of the in-group, even when this failure represented movement away from that goal (i.e. when the in-group had previously achieved positive in-group distinctiveness), as long as the in-group's performance had only deteriorated to non-distinctiveness (i.e.

¹¹ In line with hypothesis H2d, however, when failure to achieve the relational in-group goal was accompanied by progress toward that goal (i.e. lessened negative in-group distinctiveness or "achievement" of zero in-group distinctiveness from the out-group), this did not result in significantly negative evaluations of the in-group (and in the latter case resulted in significantly positive in-group evaluation).

equal performance) rather than to negative in-group distinctiveness. Similarly, charitable groups "just" failing to reach the objective goal of an annual fund-raising target did not result in negative mean evaluation, even when this failure was clearly accompanied by worsened in-group performance.¹² Thus, again against hypothesis H2d, failure to achieve objective or relational in-group goals, accompanied by movement away from those goals, did not reliably result in negative mean evaluation of the in-group, as long as the failure to reach the respective goals was not "too" stark.

The self-esteem hypothesis: Corollary 1

Question A5

Discrimination				No discrimination			
PGD		NGD		PGD		NGD	
No COV	COV	No COV	COV	No COV	COV	No COV	COV
6.34 (1.3)	-2.33 (4.1)	0.20 (4.0)	-4.78 (2.9)	5.22 (2.5)	1.83 (4.1)	0.19 (3.7)	-3.62 (3.2)

Key: PGD = positive in-group distinctiveness; NGD = negative in-group distinctiveness; COV = conflict of values. 15-point scale ranging from -7 = very negative to +7 = very positive. N = 100.

Table 7.7: Mean self-esteem levels for charitable group members according to use or non-use of intergroup discrimination, positive or negative in-group distinctiveness, and presence or absence of a conflict of values

A 2 X 2 X 2 MANOVA with use or non-use of intergroup discrimination, positive or negative in-group distinctiveness, and presence or absence of a conflict of values as within-subjects factors, and respondents' self-esteem as the dependent

¹² Although deteriorating charitable in-group performance plus a "clear" failure to reach an objective in-group goal did result in such negative evaluation.

variable, revealed significant main, first-order, and second-order interaction effects.

Averaged across in-group distinctiveness and conflicts of value, respondents who did not employ intergroup discrimination had significantly higher mean self-esteem (+0.91) than those who did (-0.14, $F_{(1,99)} = 23.57$, $p = 0.000$); averaged across use of intergroup discrimination and conflicts of value, respondents in positively distinct groups had significantly higher mean self-esteem (+2.8) than those in negatively distinct groups (-2.0, $F_{(1,99)} = 316.03$, $p = 0.000$); and, averaged across in-group distinctiveness and use of intergroup discrimination, respondents experiencing a conflict of values had significantly lower mean self-esteem (-2.2) than those who experienced no conflict of values (+3.0, $F_{(1,99)} = 306.26$, $p = 0.000$).

All first-order effects were also significant. Averaged across conflicts of value, positive in-group distinctiveness was accompanied by significantly higher mean self-esteem when intergroup discrimination was not employed (+3.5) than when it was (+2.0), and negative in-group distinctiveness was accompanied by significantly lower mean self-esteem when accompanied by intergroup discrimination (-2.3) than when unaccompanied by such discrimination (-1.7, $F_{(1,99)} = 9.88$, $p = 0.002$). Averaged across in-group distinctiveness, use of intergroup discrimination accompanied by a conflict of values lead to negative mean self-esteem (-3.6) but unaccompanied by a conflict of values lead to positive mean self-esteem (+3.3), and non-use of intergroup discrimination lead to relatively neutral mean self-esteem when accompanied by a conflict of values (-0.9) but raised mean self-esteem when unaccompanied by such a conflict (+2.7, $F_{(1,99)} = 59.23$, $p = 0.000$). Averaged across use of intergroup discrimination, positive in-group distinctiveness lead to relatively neutral mean self-esteem when accompanied by a conflict of values (-0.3) but resulted in positive mean self-esteem when unaccompanied by such a conflict (+5.8), while negative in-group distinctiveness lead to relatively neutral mean

self-esteem when unaccompanied by a conflict of values (+0.2) but resulted in negative mean self-esteem when there was such a conflict (-4.2, $F_{(1,99)} = 17.15$, $p = 0.000$).

These results are qualified by a significant second-order interaction ($F_{(1,99)} = 47.66$, $p = 0.000$). Positive in-group distinctiveness without a conflict of values resulted in positive mean self-esteem whether intergroup discrimination was (+6.34) or was not (+5.22) employed, while positive group distinctiveness with a conflict of values resulted in only slightly positive mean self-esteem when intergroup discrimination was not employed (+1.83) and resulted in slightly negative mean self-esteem when intergroup discrimination was employed (-2.33). Negative in-group distinctiveness without a conflict of values resulted in more or less neutral mean self-esteem whether intergroup discrimination was (+0.20) or was not (+0.19) employed, while negative in-group distinctiveness with a conflict of values resulted in slightly more negative mean self-esteem when intergroup discrimination was employed (-4.78) than when it was not (-3.62).

Question B5

PGD		NGD	
COV	No COV	COV	No COV
-3.53 (3.5)	6.30 (1.2)	-4.27 (3.4)	2.13 (3.4)

Key: PGD = positive in-group distinctiveness; NGD = negative in-group distinctiveness; COV = conflict of values. 15-point scale ranging from -7 = very negative to +7 = very positive. N = 100.

Table 7.8: Mean self-esteem levels for political party supporters according to positive or negative in-group distinctiveness and presence or absence of a conflict of values

A 2 X 2 MANOVA with in-group distinctiveness (positive or negative) and conflict of values (present or absent) as within-

subjects factors and respondents' self-esteem as the dependent variable revealed significant main and interaction effects.

Averaged across group distinctiveness respondents experiencing a conflict of values had significantly lower mean self-esteem (-3.90) than respondents without a conflict of interest (+4.22, $F_{(1,99)} = 587.11$, $p = 0.000$), and averaged across conflicts of value respondents in positively distinct groups had significantly higher mean self-esteem (+2.97) than respondents in negatively distinct groups (-1.07, $F_{(1,99)} = 65.78$, $p = 0.000$).

These results are qualified by a significant interaction ($F_{(1,99)} = 42.04$, $p = 0.000$). As can be seen from Table 7.8, there was little difference in mean negative self-esteem levels according to positive (-3.53) or negative (-4.27) in-group distinctiveness when that distinctiveness was accompanied by a conflict of values. However, when no conflict of values was involved mean self-esteem was more positive for positively distinct groups (+6.30) than for negatively distinct ones (+2.13).

Question C1

Own sex's earnings relative to other sex's earnings	Evaluation
Own sex more	-4.36 (3.93)
Same	6.15 (2.16)
Own sex less	-6.17 (2.50)

Evaluation on a 15-point scale from -7 = Very Negative to +7 = Very Positive.

Table 7.9: Mean evaluations according to own sex's earnings relative to other sex's earnings for the same work

Table 7.9 shows that respondents gave negative mean evaluation to situations in which either sex earned more than the other for the same work (which was assumed to involve a conflict of equalitarian values), and gave positive mean evaluation when each

sex earned equal money for equal work (Hotellings $p = 0.000$).¹³ Planned polynomial contrasts revealed a significant quadratic trend ($F_{(1,97)} = 666.78$, $p = 0.000$).¹⁴ It should be noted, however, that mean evaluation was significantly more negative when the other sex out-earned respondents' own sex than when the reverse was true ($t_{(99)} = 3.99$, two-tailed $p = 0.000$).

The above results provide strong support for hypotheses H3a and H3b. As predicted by hypothesis H3a positive in-group distinctiveness without a conflict of values resulted in positive mean self-esteem whether or not intergroup discrimination was employed. Similarly, as predicted by hypothesis H3b, negative in-group distinctiveness accompanied by a conflict of values resulted in negative mean self-esteem, again, whether or not intergroup discrimination was employed.

The results relating to hypothesis H3c are more complicated. Negative in-group distinctiveness accompanied by no conflict of values resulted in neutral mean self-esteem, as predicted by the hypothesis, but there is quite a lot of evidence that positive in-group distinctiveness accompanied by a conflict of values resulted in negative mean self-esteem (except when the conflict of values came about by not employing intergroup discrimination - where self-esteem was more or less neutral, as predicted).

¹³ A MANOVA with the three evaluations as a within-subject factor and sex as a between-subjects factor revealed no significant main or interaction effects for sex.

¹⁴ Running from "own sex more", through "same", to "other sex more". There was also a significant linear trend ($F_{(1,97)} = 13.82$, $p = 0.000$). However, sums of squares revealed that the quadratic trend ($SS_{\text{quad}} = 666.78$) explained considerably more of the variance than the linear one ($SS_{\text{lin}} = 134.82$).

The self-esteem hypothesis: Corollary 2**Question A6**

Negative in-group distinctiveness				Positive in-group distinctiveness			
Exit		No Exit		Exit		No Exit	
CA	No CA	CA	No CA	CA	No CA	CA	No CA
-0.61 (4.5)	-2.06 (3.4)	-0.23 (4.7)	-2.59 (3.9)	-1.14 (4.5)	-2.06 (4.4)	-0.92 (4.8)	-2.51 (4.7)

Number of respondents who chose exit:-

7	26			10	12		
---	----	--	--	----	----	--	--

Key: Exit = exit available; No Exit = exit unavailable; CA = cognitive alternatives available; No CA = Cognitive alternatives unavailable. 15-point scale ranging from -7 = strongly oppose the "poaching" suggestion and +7 = strongly support the "poaching" suggestion. N = 98.

Table 7.10: Levels of use of exit and/or intergroup discrimination by charitable group members according to positive or negative in-group distinctiveness, availability or non-availability of exit, and availability or non-availability of cognitive alternatives

A 2 X 2 X 2 MANOVA with availability or non-availability of cognitive alternatives, availability or non-availability of exit, and positive or negative in-group distinctiveness as within-subjects factors, and willingness to engage in intergroup discrimination as the dependent variable, revealed a significant main effect for group distinctiveness and a significant first-order effect for the interaction between group distinctiveness and availability of exit. No other main or interaction effects were significant.

Averaged across availability of exit and of cognitive alternatives, positive in-group distinctiveness resulted in a slightly stronger mean disinclination to employ intergroup discrimination (-1.66) than negative in-group distinctiveness (-1.4, $F_{(1,97)} = 20.02$, $p = 0.000$).

This was qualified by the significant first-order interaction between in-group distinctiveness and availability of exit ($F_{(1,97)} = 9.91, p = 0.002$). For both positively and negatively distinct groups, mean disinclination to employ intergroup discrimination was slightly stronger when exit was unavailable than when it was available, and mean disinclination to employ such discrimination was stronger in positively distinct groups than in negatively distinct ones (PGD+No Exit = -1.72, PGD+Exit = -1.60, NGD+No Exit = -1.41, NGD+Exit = -1.34).

It can also be noted that where exit was available it was employed most when in-groups were negatively distinct and cognitive alternatives were unavailable. The differences between the numbers of respondents employing exit (where it was available) according to in-group distinctiveness and availability of cognitive alternatives was not significant however ($\chi^2 = 2.587, p > 0.05$).

Question B6

Negative in-group distinctiveness				Positive in-group distinctiveness			
Exit		No Exit		Exit		No Exit	
CA	No CA	CA	No CA	CA	No CA	CA	No CA
5.36 (2.6)	0.87 (3.0)	5.99 (2.0)	3.52 (4.7)	5.56 (2.9)	5.69 (2.4)	5.83 (2.4)	6.94 (6.8)

Number of respondents who chose exit:-

10	71			7	13		
----	----	--	--	---	----	--	--

Key: Exit = exit available; No Exit = exit unavailable; CA = cognitive alternatives available; No CA = Cognitive alternatives unavailable. 15-point scale ranging from -7 = strongly oppose the "poaching" suggestion and +7 = strongly support the "poaching" suggestion. N = 93.

Table 7.11: Levels of use of exit and/or intergroup discrimination by political party supporters according to positive or negative in-group distinctiveness, availability or non-availability of exit, and availability or non-availability of cognitive alternatives

A 2 X 2 X 2 MANOVA with availability or non-availability of cognitive alternatives, availability or non-availability of exit, and positive or negative in-group distinctiveness as within-subjects factors, and willingness to engage in intergroup discrimination as the dependent variable, revealed significant main and first-order interaction effects. Averaged across group distinctiveness and availability of exit, respondents for whom cognitive alternatives were available were significantly more willing to employ intergroup discrimination (+5.69) than respondents for whom cognitive alternatives were unavailable (+4.26, $F_{(1,92)} = 78.51$, $p = 0.000$); averaged across group distinctiveness and availability of cognitive alternatives, respondents were significantly more willing to engage in intergroup discrimination when exit was unavailable (+5.57) than when it was available (+4.37, $F_{(1,92)} = 23.47$, $p = 0.000$); and, averaged across availability of cognitive alternatives and of exit, respondents were significantly more willing to engage in intergroup discrimination when their group was positively distinct (+6.01) than when it was negatively distinct (+3.94, $F_{(1,92)} = 31.58$, $p = 0.000$).

These main effects are qualified by significant first-order interactions. Averaged across group distinctiveness, when cognitive alternatives were available there was very little difference between respondents' mean willingness to employ intergroup discrimination according to whether exit was available (+5.46) or unavailable (+5.91), but when cognitive alternatives were unavailable respondents were more willing to employ intergroup discrimination when exit was unavailable (+5.23) than when it was (+0.94, $F_{(1,92)} = 4.62$, $p = 0.034$). Averaged across availability of exit, when cognitive alternatives were available there was very little difference between respondents' mean willingness to employ intergroup discrimination according to whether their group was positively (+5.70) or negatively (+5.68) distinct, but when cognitive alternatives were unavailable respondents were far more willing to employ intergroup discrimination when their group was positively distinct (+6.32)

than when it was negatively distinct (+2.20, $F_{(1,92)} = 56.54$, $p = 0.000$). Finally, averaged across availability of cognitive alternatives, for both positively and negatively distinct groups, mean inclination to employ intergroup discrimination was slightly stronger when exit was unavailable than when it was available, and mean inclination to employ such discrimination was stronger in positively distinct groups than in negatively distinct ones (PGD+No Exit = +6.40, PGD+Exit = +5.63, NGD+No Exit = +4.80, NGD+Exit = +3.12, $F_{(1,92)} = 10.74$, $p = 0.001$). The second-order interaction was not significant ($F_{(1,92)} = 1.54$, $p = 0.218$).

It can also be noted that when exit was available it was employed considerably and significantly more by respondents in negatively distinct groups for whom cognitive alternatives were not available (71) than by respondents in positively distinct groups for whom cognitive alternatives were available (7) or unavailable (13), or by respondents who were in negatively distinct groups with cognitive alternatives available (10) ($\chi^2 = 4.360$, $p = 0.05$).

In evaluating these results with respect to hypotheses H4a to H4d, it is clear that very different results were obtained for charitable groups than for political ones. No significant intergroup discrimination was indulged in by members of the former groups whereas intergroup discrimination was the norm for members of the latter ones, almost irrespective of in-group distinctiveness and availability of cognitive alternatives and of exit. For this reason the results of each type of group will be evaluated separately.

For charitable groups, the overwhelming result with respect to corollary 2 of the self-esteem hypothesis was that across conditions respondents were never willing to employ intergroup discrimination. The effects of in-group distinctiveness and availability of exit and of cognitive alternatives were on relative disinclination to employ such discrimination, and these effects, although significant, were small. Further, willingness

to employ exit (where available) was similarly slight and did not differ significantly according to in-group distinctiveness and availability of cognitive alternatives. Rather than try to evaluate hypotheses H4a to H4d in the light of these results, it seems more sensible to simply accept that members of charitable groups were generally not willing to employ either intergroup discrimination or exit, irrespective of in-group distinctiveness and availability of cognitive alternatives and exit.

The results from the political groups are much more amenable for evaluating hypotheses H4a to H4d. In support of hypothesis H4a respondents were willing to employ intergroup discrimination when cognitive alternatives were available but exit was not, irrespective of in-group distinctiveness. However, contrary to hypothesis H4b, respondents were also generally willing to employ such discrimination when cognitive alternatives were not available, regardless of in-group distinctiveness and the availability of exit.

Hypothesis H4c was not supported by the political group members. This hypothesis stated that when both cognitive alternatives and exit were available, members of positively distinct groups would discriminate but members of negatively distinct groups would not. In fact, although in such situations members of positively distinct group were slightly more willing to engage in intergroup discrimination than members of negatively distinct groups, members of both groups were willing to employ high levels of mean discrimination.

Hypothesis H4d received strong support from political group members in the sense that when exit was available it was used considerably more by members of securely negative inferior groups than by any other respondents. It should be noted, however, that there was little difference in the use of exit in the other three relevant conditions. In particular, members of securely positive groups did not employ significantly less exit than members of insecurely positive or negative groups.

It appears then that while members of charitable groups were universally unwilling to employ intergroup discrimination, for members of political groups intergroup discrimination was the norm. Thus, while it is to some extent possible to evaluate the hypotheses derived from corollary 2 of social identity theory's self-esteem hypothesis, it seems that there was an unaccounted for "extraneous" variable driving the main results, such that non-discrimination and non-exit was the norm for charitable group members and discrimination and non-exit was the norm for political group members.

DISCUSSION

A number of interesting results were obtained from the present study, many of which have considerable implications for the formulation of social identity theory.

The nature of intergroup discrimination

It was found in the present study that although supporters of charitable groups on average employed "weak" intergroup discrimination in the form of in-group preference, this in-group preference was accompanied by mean positive evaluation of both in-groups and out-groups. In the terminology introduced in Chapter 3, supporters of charitable groups engaged in relative discrimination in favour of the in-group, but not absolute discrimination against either in-groups or out-groups. Supporters of political groups, by contrast, employed "strong" absolute discrimination against certain selected out-groups. That is, against these out-groups political group members adopted both in-group preference and out-group derogation. Against other out-groups, however, political group supporters employed only "weak" discrimination. The determining factor in whether group supporters employed weak or strong intergroup discrimination

seemed to be whether or not the in-groups concerned had a "tradition" or a "culture" of competition with the relevant comparison groups. Where such a tradition existed strong intergroup discrimination took place, and where it did not, only weak discrimination was employed.

Social identity theory says that to the extent that social identities tied to particular groups are salient in multigroup contexts, there will be a need for individuals to perceive their in-groups as positively distinct from relevant out-groups. Weak intergroup discrimination (i.e. in-group preference without out-group derogation) would clearly serve such needs. The problem for social identity theory, if indeed there is one, is to explain when, if ever, social identity needs will promote strong intergroup discrimination, i.e. in-group preference with out-group derogation.

There are several possible responses to this problem which might be given on social identity theory's behalf. The first is simply to argue that the theory does not attempt to explain strong intergroup discrimination. Rather, its role is to supplement Sherif's realistic conflict theory by pointing out that whenever social identities are salient in-group members will need their in-groups to be positively distinct from all relevant comparison groups. Thus, social identity theory supplements realistic conflict theory by predicting that in-group members will usually display or employ in-group preference when comparing salient in-groups with relevant out-groups. Further, this can be done without out-group derogation, even when between-group social comparison outcomes reveal that the in-group is negatively distinct from relevant out-groups (remembering that social identity theory conceives of in-group status as an outcome variable, Tajfel & Turner, 1979: 43): i.e. by simply giving more "goods" to the in-group than to the out-groups, thus improving the between-group situation in favour of the in-group. Such an approach would then maintain that explanations of strong

intergroup discrimination fall within the province of Sherif's theory.¹⁵

While such a response is possible (and indeed seems close to the position eventually taken by Tajfel - see Tajfel & Turner, 1979: 46-47), it does seem to rather reduce the impact of social identity theory as an explanatory theory of intergroup conflict:¹⁶ the theory is relegated to explaining situations in which in-group members are simply "nicer" to in-groups and their members than they are to out-groups and theirs, which applies only to very tame instances of intergroup "discrimination" and intergroup "conflict".

A second response might be to argue that supporting charities is very different to supporting political parties, in that each is associated with a different aspect of an individual's self-concept. Such an argument might claim that supporting a political party involves cognitions to the effect that "I am a member of this particular political group", whereas supporting a charity involves rather different cognitions to the effect that "I approve of and support this particular charity". The difference, of course, is that the former is claimed to involve a self-definition as a group member, thus invoking social identity, whereas the latter involves a self-definition in terms of a particular individual characteristic, thus invoking personal

¹⁵ An exception occurs "when a group's action for positive distinctiveness is frustrated, impeded, or in any way actively prevented by an out-group [which] will promote overt conflict and hostility between the groups" (Tajfel & Turner, 1979: 46). Such an exception points to the fact that social identity theory does not, in the final analysis, conceive of in-group status solely as an outcome variable, but also conceives of it as a group goal. The processes resulting from conceiving positive in-group distinctiveness as a group goal, however, are already incorporated within Sherif's theory, and so examination of social identity theory's distinctive contribution must be restricted to situations in which in-group status is (at least in the first instance) an outcome variable. These points will be returned to in the final chapter of this thesis.

¹⁶ The title of Tajfel & Turner (1979) is 'An integrative theory of intergroup conflict'.

identity. Thus, such an account might continue, supporting a political party leads to processes tied up with social identity needs (i.e. social identity, social comparison, and intergroup discrimination), whereas supporting a charity does not. Therefore, such an account would conclude, social identity theory has nothing to say about the behaviour of people who support charities, but does apply to supporters of political parties by predicting that they will engage (in appropriate circumstances) in social competition.

There are a number of things to say in response to such a possibility. First, it may rely on a false distinction between supporters of each of the groups. To the extent that a particular charity can be identified and numbers of people are known to support it (i.e. identify with it), it seems that social identity theory would in fact predict that supporters of such a charity would identify themselves as group members, and act accordingly (see Tajfel & Turner, 1979: 40). At the very least, it is a matter of empirical investigation to determine whether or not there is a legitimate distinction to be made between supporters of political parties and supporters of charities in terms of their self-conceptions as psychological group members and/or their social identities. More importantly, even if for the sake of argument it is conceded that social identity theory only applies in the present study to supporters of political groups, it is still the case that such supporters employed both weak and strong average intergroup discrimination relative to particular out-groups. Therefore it is also still the case either that social identity theory must be restricted to explaining in-group preference rather than stronger forms of intergroup discrimination, or that something within social identity theory must be identified to enable prediction and explanation of each form of discrimination.

A third possibility is to attempt to explain the differences in strength of discrimination in terms of differences in the relevance of the comparison groups. Such an account would argue

that other charities are not relevant comparison groups for any given charity and so strong intergroup discrimination against such groups would not be anticipated. Further, not all political parties are relevant comparison groups for any given political party, and strong intergroup discrimination would be anticipated only against those which are. According to such an interpretation, the results of the present study wholly support social identity theory, as (strong) intergroup discrimination occurs only when respondents identify with in-groups and compare them with relevant out-groups.

Apart from perhaps more or less removing social identity theory as an explanation of weak intergroup discrimination (i.e. in-group preference without out-group derogation), the main problem with such a response is that it is far from clear how to identify whether or not any given comparison group is a relevant comparison group. Social identity theory identifies "similarity, proximity, and situational salience [as] among the variables that determine out-group comparability" (Tajfel & Turner, 1979: 41), but it is difficult to see how such variables would apply differentially across the out-groups of interest here. All of the comparison groups in the present study, for both political parties and charitable groups, seem to be roughly equivalent to their respective in-groups in terms of their similarity, proximity, and situational salience. Perhaps the only way to defend the notion of out-groups being relevant or irrelevant comparison groups is therefore to reintroduce the idea that in-groups can have "traditional" out-groups, against whom there is a "culture" of strong intergroup discrimination. This, of course, returns us to the position we started out with.

It seems, then, that social identity theory is supported in so far as multigroup social categorization plus identification by in-group members (both as members of one in-group but not of others, and with the in-group) seems sufficient to promote those members to engage in weak intergroup discrimination, i.e. in-group preference without out-group derogation. Explanations of

strong intergroup discrimination or intergroup conflict, however, involving out-group derogation as well as in-group preference, seem to require that in-group "customs", "traditions", or "norms" of behaviour toward particular comparison groups be considered.¹⁷

Such a conclusion still leaves at least two important questions unanswered. First, where do such in-group norms come from? Second, are such norms, in advocating out-group derogation rather than simple in-group preference, concerned only (or even necessarily) with the achievement and maintenance of positive in-group distinctiveness from particular "relevant" out-groups: that is, do they easily fall within the province of Tajfel's social identity theory? These questions will be addressed below.

The level at which social identity theory operates

It is worth reiterating that derogation of particular political groups was not restricted to members of other groups. People who explicitly disclaimed political group membership still engaged in derogation of specific target groups. The importance of this for social identity theory is that it highlights the possibility of "cross-level" behaviour (broadly conceived). That is, it is possible for *individuals* to take particular stances (e.g. attitudes, behaviours) toward or against groups. Individuals may hate the BNP without needing to affiliate themselves to any distinct political group. By the same token, it seems possible for particular groups to adopt particular stances toward or against specific individuals (e.g. "fan clubs").

¹⁷ Here *established* group "norms" are being investigated and discussed. Although, as mentioned in the Introduction to this chapter, it is also likely that in-group norms of strong intergroup discrimination will result when *situational* factors suggest that such behaviour is necessary and/or appropriate. (See also footnote 15, p. 352.)

Such reasoning suggests that in attempting to explain and predict widespread discrimination against particular groups and group members, social identity theory could potentially benefit from abandoning its exclusive focus on between-group phenomena to allow for the possibility of "across-level" phenomena (as well as between-individual phenomena, as suggested in the previous chapter. See also Horton, 1993).

In-group evaluation

In the present study it was found that evaluation of in-groups involved much more than simple attainment or non-attainment of positive in-group distinctiveness from relevant comparison groups on dimensions positively valued by in-groups. Such between-group social comparison outcomes did affect in-group evaluations, but so did: (i) outcomes of comparing in-groups against temporal or objective standards (e.g. previous in-group situations or desired criteria); and, (ii) progress toward or away from (temporal, objective, or between-group relational) in-group goals.

One might want to argue that the temporal and objective comparison outcomes had the effects they did on in-group evaluations only because they were indicative to respondents of how in-group positions were changing relative to specific relevant out-groups. Thus, it is possible that discovering an in-group has achieved certain temporal or absolute goals may not be directly beneficial for positive in-group evaluations, but rather may be indirectly beneficial because such achievement suggests that the in-group is also achieving or enhancing positive in-group distinctiveness (or reducing negative in-group distinctiveness) with respect to some relevant out-group. Similarly, perceiving that an in-group has failed to achieve certain temporal or absolute goals may only be indirectly detrimental to in-group evaluation because such failure suggests decreased positive in-group distinctiveness or increased negative in-group distinctiveness with respect to some relevant out-group.

That is, achievement or non-achievement of absolute and/or temporal group goals may only be relevant to in-group evaluations to the extent that such achievement or non-achievement reveals or suggests changed *between-group* social comparison outcomes. (A similar argument can be made with respect to movement toward or away from absolute and/or temporal goals having only indirect effects on in-group evaluations to the extent that such movement is relevant in revealing or suggesting altered *between-group* social comparison outcomes relative to some relevant out-group.)

The first thing to note about such a possibility is that it concedes that in-group evaluations can be affected by temporal *as well as* straightforward relational *between-group* social comparison outcomes. That is, *change* in valued in-group distinctiveness is conceded to affect in-group evaluations in *addition* to valued in-group distinctiveness *per se*. Evidence of exactly this point was obtained in the present study where respondents evaluated their political in-groups in situations where those groups won an election by obtaining 250 votes more than their nearest rivals. There were two such situations: one where the in-group had previously been beaten by its rival by 500 votes, and one where the in-group had previously beaten its rival by a similar margin. If it were *only* positive in-group distinctiveness which affected in-group evaluation the in-group evaluations in each of these situations would have been identical. However, while the in-group received positive evaluations in both situations, where the in-group both beat its rival and reversed previous negative in-group distinctiveness the in-group was evaluated significantly more positively than where the in-group beat its rival by a narrower margin than it had on a previous occasion. Thus, the *changes* in the respective intergroup situations (i.e. the temporal comparison outcomes) made contributions to in-group evaluation other than those

provided by the in-group distinctiveness per se (i.e. the between-groups social comparison outcome).¹⁸

The second thing to note in response to the claim that temporal and absolute comparison outcomes may affect in-group evaluation only to the extent that they reveal or suggest particular between-group social comparison outcomes (and/or changes in such outcomes) is that such a claim is potentially unfalsifiable. For example, in the present study it was found that supporters of charitable groups evaluated those groups positively if the groups were deemed to have achieved and/or to have made progress toward absolute goals. In this chapter such a finding was accepted at face value, but it could nevertheless have been argued that such achievements only affected in-group evaluations to the extent that they were relevant in revealing or suggesting changed intergroup relationships between the charities and some relevant (but not easily identified) out-group(s). But why might one want to make such a claim? That between-group social comparison outcomes undoubtedly affect in-group evaluations in some circumstances does not entail that in-group evaluations can only be affected by such between-group social comparison outcomes. Neither is it enough that one could always claim that absolute or temporal comparison outcomes only affect in-group evaluations to the extent that they reveal or suggest changed between-group social comparison outcomes. Unless it is a dogmatic assertion, such a claim is an assumption: an assumption which the present study was set up in part to test. And this test revealed exactly the same sort of evidence that absolute and temporal comparison outcomes affect in-group evaluations as was found to support the claim that between-group social comparison outcomes do. To then insist that the temporal and absolute comparison outcomes had the effects they did only in so far as they revealed or suggested to

¹⁸ It cannot be determined from the present study whether the reversed negative in-group distinctiveness made an additional positive contribution and/or whether the attenuated positive in-group distinctiveness made a detrimental contribution to the positive in-group evaluations resulting from the positive in-group distinctiveness per se.

respondents changed between-group relationships is no more legitimate than to insist that between-group social comparison outcomes have the effects they do only to the extent that they reveal or suggest changed absolute or temporal comparison outcomes.

Therefore, unless and until evidence can be obtained which indicates that effects on in-group evaluation from *all* comparison outcomes are (perhaps "at root") actually effects of actual or inferred *between-group* social comparison outcomes, the present study provides evidence that in-group evaluation can be affected by attainment or non-attainment, or progress toward or away from, between-group relational *and/or* temporal or absolute in-group goals.¹⁹ Broadly speaking, it is argued here that the evidence from this study supports the claim that achievement of *and/or* movement toward in-group goals tends to provide positive contributions to in-group evaluation, and non-achievement of *and/or* movement away from such goals tends to make negative contributions to in-group evaluation - be those goals temporal, objective, or between-group relational (i.e. socially comparative).

The common thread to each of these evaluative phenomena is the presence and importance of in-group goals, of which positive in-group distinctiveness seems to be only one. This raises the possibility that in some situations in-groups might be very positively evaluated even in the absence of social comparisons, and perhaps even when the in-group is negatively distinct to specific relevant out-groups. So, for example, it is possible

¹⁹ Positive in-group distinctiveness *and/or* the absence of negative in-group distinctiveness of salient in-groups from relevant out-groups is conceived here as an in-group goal. Although not usually conceived in such terms within social identity theory, such a formulation is at least consistent with the theory. See footnote 15, p. 352. There is a difference between what is claimed here and what is claimed in social identity theory, however. In social identity theory positive in-group distinctiveness is conceived as a *universal* goal for groups *and/or* their members, whereas it is considered here as a *possible* (albeit common) one.

that an in-group might be positively evaluated by its members if it both achieved (or exceeded) what it had set out to do and showed signs of constant improvement (perhaps closing the gap between itself and one or more rival out-groups), even if at a particular moment it happened to be negatively distinct on in-group valued comparison dimensions from one or more relevant out-groups. In such an instance, it is also possible that in-group members would derive positive social identity contingent self-esteem from membership of the group, again, regardless of the current negative in-group distinctiveness.

One is reminded here of Tajfel's notion of "satisfactory self-realization" (Tajfel, 1978b: 9). This is a phrase Tajfel considered using to refer to the motivation behind using comparison outcomes for self-evaluation purposes. That is, Tajfel thought that people desired "satisfactory self-realization" and used comparisons to determine the extent to which they had achieved it. Tajfel rejected the term as "hopelessly vague" and "synthetic", because it "can mean so much" as to be "in danger of meaning very little at all". Tajfel adopted instead the notion of satisfactory (or positive) self-esteem (or self-regard, self-image, self-concept, etc.) (e.g. see Tajfel 1978a, 61; 1978b, 9; Tajfel & Turner, 1979: 40). It is a moot point whether such a notion is any less vague, synthetic or meaningless than satisfactory self-realization (see previous chapters), but, more importantly, such a preference appears misplaced in as much as: (i) it seems likely that it is the knowledge or belief that one has not achieved satisfactory self-realization which results in unsatisfactory self-esteem (and the motivation to do something about it) in the first place (that is, self-esteem is "merely" a mediating variable between assessment of satisfactory self-realization and behaviour intended to bring such satisfactory self-realization about); and, (ii) unsatisfactory self-realization can be conceived as a much more general term than self-esteem. With respect to the latter point, the knowledge that one has yet to achieve one's goals may not affect one's self-esteem (if, for example, one can make an

external attribution for one's failure so far), but may well provide a motivation to continue trying to bring such "satisfactory self-realization" about.

In attempting to develop this line of thought, it is possible to coin the phrase "satisfactory in-group realization" to refer to the state where an in-group has achieved one or more of its goals. Achievement of (or acceptable progress toward) in-group goals (be they terminal or instrumental) can be considered to contribute to a sense of satisfactory in-group realization and non-achievement (or unacceptable progress toward/movement away from) terminal or instrumental in-group goals can be considered to contribute negatively to a sense of satisfactory in-group realization. In turn, the extent of satisfactory in-group realization can be considered to determine evaluation of the in-group (although not necessarily "in-group esteem"). Finally, to the extent that in-group members adopt in-group goals as their own, extent of satisfactory in-group realization will (partially) determine in-group members' sense of satisfactory self-realization (which may or may not affect their self-esteem).

The relevance of all this to social identity theory is manifold. Three main considerations are as follows. First, it shifts the focus of attention away from between-group social comparison outcomes per se to seeing such outcomes as one of several possible means of evaluating progress toward or away from satisfactory self-realization. Second, it allows for the possibility of non-social comparison outcomes (e.g. temporal, objective, actual-ideal) affecting evaluation of in-groups and in-group members, and in turn affecting in-group behaviour aimed at bringing about or maintaining such realization. Third, it allows for the possibility of between-group social comparison outcomes not affecting the self-esteem of in-group members, for two reasons: the comparison outcome may simply not be relevant to evaluating satisfactory in-group (and self-) realization (e.g. if the in-group goal is best evaluated via objective comparisons), and, even if such evaluation does occur, this will

not affect in-group members' self-esteem unless the self and/or the in-group are held responsible for its current position and/or the perceived nature of the in-group (and/or its members) is challenged in some way.

Put simply, the present study calls into question the traditional (or at least widely promulgated) tenet of social identity theory that positive in-group distinctiveness is necessarily good for in-group evaluation and in-group members' self-esteem while negative in-group distinctiveness is necessarily bad for in-group evaluation and in-group members' self-esteem.

All that said, however, one finding from the present study concerning in-group evaluation merits further attention. This is that when political groups failed to achieve positive in-group distinctiveness which they had previously enjoyed, this did not result in negative evaluation of the in-group as long as the in-group was still not negatively distinct. That is, even though the in-group had both failed to achieve its relational in-group goal (of positive in-group distinctiveness) and had moved away from that goal, the in-group still received (admittedly very small) positive evaluation when it had least not fallen to a position of inferiority. This perhaps attests both to the bias toward positive evaluation of the self (and by extension, identified-with in-groups) and to the possibility that it is negative in-group distinctiveness which is most damaging to in-group (and self-) evaluation, rather than a simply failure to attain positive in-group distinctiveness. That is, the possibility still seemed open to these respondents to say "at least we didn't come out worse than we did": an actual-possible in-group (downward) comparison from which respondents were able to claw a vestment of positive evaluation.

Corollary 1 of social identity theory's self-esteem hypothesis

The results from the present study bearing on corollary 1 of the self-esteem hypothesis within social identity theory strongly support the arguments that: (i) Hogg & Abrams do not give an adequate formulation of that hypothesis (with respect to corollary 1); and, (ii) a more adequate formulation of corollary 1 can be empirically supported. Thus, it is clear from the present study that it is not use of intergroup discrimination per se (or even "successful" intergroup discrimination) which leads to positive evaluation of the in-group (and of in-group membership): it is positive in-group distinctiveness without a conflict of values (regardless of whether or not intergroup discrimination is employed). Thus, successful intergroup discrimination is a means to an end (i.e. positive in-group distinctiveness without a conflict of values), not an end in itself.²⁰

Further, it seems that both negative in-group distinctiveness and/or a conflict of values contribute negatively to in-group evaluation, presumably because such occurrences indicate unsatisfactory in-group realization (i.e. both because of in-group inferiority and because of violation of cherished values).

Indeed, of the two factors (i.e. in-group distinctiveness and value conflict), it appears that it is a conflict of values which has the stronger effect, as positive in-group distinctiveness accompanied by a conflict of values so often led to negative evaluation. It can be hypothesized that this is because a conflict of values entails unsatisfactory in-group realization, in that it violates members' "in-group-concept" ("this is not the sort of group we are/should be"), whereas in-group distinctiveness only has the potential (see above) to affect

²⁰ Note: In the light of the previous discussion it should also be clear that positive in-group distinctiveness without a conflict of values is similarly not an end in itself. It is one possible indication of satisfactory in-group realization.

evaluation of in-group progress toward or movement away from a particular goal (as opposed to threatening the "very essence of *its existence*").

As mentioned in Chapter 5, the large body of literature which reports "tests" of corollary 1 cannot be considered as evidence for or against it, as most if not all of the relevant studies investigate (at best) the link between use of intergroup discrimination and subsequent self-esteem rather than whether in-group members' (social identity contingent) self-esteem is affected by positive or negative in-group distinctiveness with or without a conflict of values.

Corollary 2 of social identity theory's self-esteem hypothesis

The results of the present study are not nearly as "clean" with respect to the second (as compared to the first) corollary of the self-esteem hypothesis within social identity theory.

The general thrust of the "reformulated" corollary 2 argued in this thesis to adequately reflect social identity theory's self-esteem hypothesis is that group members will employ intergroup discrimination when: (i) they need to (i.e. when in-groups are inferior or insecurely superior); (ii) such discrimination seems likely to achieve their desired end (i.e. secure superiority without a conflict of values); and, (iii) preferential alternatives are not available (e.g. social mobility to a more securely superior group). When any of these conditions are not in place intergroup discrimination is not predicted (at least not nearly as confidently as when they are in place).

The results from the present study indicated that members of charitable groups were more or less consensually unwilling to employ intergroup discrimination, even when all of the above conditions were in place, and they were also generally unwilling to employ alternative methods of meeting positive social identity

needs (e.g. social mobility). Members of political groups, on the other hand, were more or less universally willing to employ intergroup discrimination, although they too were generally (although not consensually) reluctant to avail themselves of available social mobility strategies.

The key to interpreting these results is perhaps the now familiar notion of "conflicts of values". The discussion in the previous sub-section makes it clear (as does a careful reading of Tajfel) that achieving positive in-group distinctiveness is unlikely to result in positive evaluation of the in-group and of in-group membership if it is achieved only at the cost of a conflict of values. By extension, group members are unlikely to employ intergroup discrimination to achieve positive in-group distinctiveness if by doing so they give rise to a conflict of values, as such an outcome is likely to be detrimental to evaluation of the in-group and of in-group membership. Unfortunately, the variable of possible value conflict was not incorporated into the present study's examination of corollary 2 of social identity theory's self-esteem hypothesis.

If one belongs to a charitable group it is probable that objective goals are considerably more important than relational (i.e. between-group) ones, especially if the other groups have objectives similar to the in-group's own. It also probable, bearing in mind the benevolent nature of charitable work, that use of intergroup discrimination against other charities striving for similar goals as the in-group would entail a conflict of values. Thus, the antecedents identified by social identity theory to lead to intergroup discrimination were not in place in the present study's attempted test of corollary 2 of the self-esteem hypothesis, at least for charitable group membership, and the results obtained are silent both on the adequacy of the argued-for formulation of that corollary and on the empirical validity of it. Put simply, the corollary and its formulation were just not tested by the items concerned with charitable group membership.

Political group membership, by contrast, has intergroup discrimination (at least in the sense of voting for one's party) as an intrinsically valuable thing (above and beyond its instrumental value in striving for one's party to be elected). Here, not employing intergroup discrimination would actually constitute a conflict of values other than in exceptional circumstances. One such exceptional circumstance is where one's own party cannot possibly win but another (similar in important respects) party can, especially when defeating a particular out-group is perceived as the over-reaching in-group goal. In the case of political group membership, therefore, we should expect that group members will standardly employ intergroup discrimination, and will only fail to do so in circumstances in which getting rid of a hated out-group overcomes the lesser goal of expressing support for one's favourite party. The results obtained in the present study are consistent with such a scenario. That is, political group members typically employed intergroup discrimination (sometimes to contribute to in-group victory, sometimes to express in-group solidarity), and only failed to do so when a more important in-group goal could be achieved by alternative action (which perhaps should be considered as indirect intergroup discrimination, rather than an absence of it).

In other words, the present study failed to take account of in-group values (and potential conflicts thereof), and thus failed to adequately investigate corollary 2 of the self-esteem hypothesis within social identity theory. It can be noted, however, that the discussion of the results considered in this sub-section, although post-hoc, are perfectly consistent with the second corollary of social identity theory's self-esteem hypothesis. The problem was with an inadequate method to test that formulation, rather than with the corollary itself.

A short note about methodology

Although it produced some interesting and potentially important findings, the results from the present study must be viewed with some caution. Because of the variety of rather disparate issues the study tried to address at once, the "manipulations" and measurements used are potentially somewhat unreliable and invalid. Two particular difficulties or limitations should be especially noted. First, many of the "manipulations" used concerned respondents imagining how they believed they *would* feel in a variety of hypothetical situations. Obviously, future research would improve considerably on the present study if real situations were manipulated and/or examined. Second, rather "lax" dependent measures were employed in the present study, several of which involved making questionable assumptions. To draw attention to just one example, in question 3 of each of the two main sections of the present questionnaire, respondents were asked to indicate "how they would feel" in a variety of situations, which was assumed to be a measure of in-group evaluation/social identity contingent self-esteem. Bearing in mind several of the arguments offered earlier in this thesis, such assumptions are questionable. Such assumptions are given some credence, however, by the fact that the pattern of results obtained supported many or most of the a priori predictions made. Nevertheless, future research would clearly benefit from using more obviously reliable and valid indicators of the various dependent variables examined in the present study.

CONCLUSIONS

Notwithstanding the methodological limitations of the present study just noted, the main conclusion to be made is clearly that social identity theory (or at least common understandings, applications and tests of it) is considerably too simplistic with regard to several of its key notions and propositions.

Specifically, it pays scant attention to: (i) potential differences between particular "sorts" of "out-groups" (and in-groups); (ii) potential variations in types of intergroup discrimination toward these different "out-groups"; (iii) phenomena (including group derogation) which may not occur at a strictly between-group level; (iv) possible alternative sources of in-group (and in-group membership) evaluation other than in-group distinctiveness; (v) the potential complexity of the relationship between in-group distinctiveness and in-group (and in-group membership) evaluation; (vi) alternatives to "self-esteem" as determinants and consequences of intergroup discrimination (and other phenomena); and (vii) the complexity of the motivational account commonly known as the "self-esteem" hypothesis, particularly with respect to the potentially crucial variable of "conflicts of value".

Of these issues, it appears that the notion of in-group goals, (broadly conceived to include in-group values), is perhaps the most important. It is arguable that social identity theory cannot adequately explain or predict group (or group members') behaviour without taking such goals (and values) into consideration. This argument will be returned to in the concluding chapter of this thesis.

CHAPTER 8: CONCLUSION - DEVELOPING SOCIAL IDENTITY THEORY

INTRODUCTION

This chapter attempts to provide the link between the material presented in this thesis on the one hand, and future research in the area of social identity theory on the other. The chapter begins with a brief review of the main results obtained (and conclusions reached) during the empirical studies reported in earlier chapters. This is followed by a section considering the possible implications of those main findings, to the extent that they are valid, for various aspects of social identity theory. The next section examines the possible methodological and theoretical limitations which potentially threaten the validity of those main findings.

This is followed by a section in which an attempt is made to very tentatively sketch a bare-bones account of what social identity theory might look like if modified to be consistent with most relevant contemporary research (including the research presented in this thesis). This modified account is then compared and contrasted with Tajfel's original version of the theory.

The final section suggests potentially fruitful areas for future research and development of social identity theory.

REVIEW OF KEY FINDINGS

In Study 1 (reported in Chapter 2), social categorization accompanied by a "thinkaloud" procedure did not result in significant mean sample-level in-group favouritism. Further, at an individual level, subjects seemed to employ a variety of strategies, and they gave a variety of rationales for those strategies.

Neither was it the case that social categorization accompanied by thinkaloud resulted in most subjects identifying with their minimal groups and/or membership of them. Only a minority of subjects explicitly mentioned identifying with their groups, and fewer referred to such identification to explain their allocation behaviours. It is noteworthy that those subjects who *did* employ in-group favouritism all explicitly mentioned that they identified with their minimal groups (although others who identified in this way did not employ such discrimination). No subjects, however, explained their in-group preference (or any other strategy) in terms of striving to achieve positive in-group distinctiveness, positive social identity, or satisfactory self-esteem.

The main conclusions drawn from Study 1 were as follows. First, (accepted, multigroup) social categorization per se does not inevitably promote intergroup discrimination in favour of the in-group, at either individual or sample level. Second, (accepted, multigroup) social categorization per se is insufficient as an explanation of any uniformities of social behaviour, not just of intergroup discrimination. Third, (accepted, multigroup) social categorization per se is insufficient to result in people identifying with their imposed social categories and/or their membership of them (as opposed to identifying themselves as in-group members). Fourth, social identification with in-groups may be a necessary, but it is certainly not a sufficient, precondition of in-group favouritism. Fifth, behaviour in the minimal group paradigm does not seem to be motivated by social identity needs, at least not in the sense of using intergroup discrimination as a means to try and achieve positive in-group distinctiveness, and thereby satisfactory social identity contingent self-esteem.

In Study 2 (reported in Chapter 3), significant sample-level in-group favouritism did follow social categorization in the minimal group paradigm. As in Study 1, however, only a subset

of individual subjects employed such in-group preference. Other subjects pursued a variety of other, non-discriminatory, strategies.

Where subjects did engage in in-group favouritism, two distinct forms could be identified. One form accompanied in-group preference with out-group derogation. The other form strove for maximum in-group profit consistent with in-group superiority.

The main conclusions drawn from Study 2 were as follows. First, (accepted, multigroup) social categorization is not a sufficient cause of in-group favouritism at an individual level. Second, where individuals do employ in-group favouritism following social categorization, that social categorization is not sufficient to determine which form of discrimination favouring the in-group individuals will employ.

In Study 3 (reported in Chapter 4), both mean sample-level in-group favouritism and mean sample-level fairness across groups followed social categorization in the minimal group paradigm (i.e., in condition 1). Following intragroup communication mean in-group favouritism polarized, and mean fairness across groups depolarized to zero.

In all three conditions, individual subjects and particular sub-groups pursued a variety of strategies. When transcripts of sub-groups' decision-making processes in condition 2 were examined, it was possible to identify for each sub-group: (i) whether sub-group norms were established; (ii) what the contents of any established norms were, and in which contexts they applied; and, (iii) individual sub-group members' levels of commitment to any sub-group norms established. Using such a normative analysis, all individual, sub-group, and sample-level condition 2 and condition 3 behaviour could be fully explained.

The main conclusions drawn from Study 3 were as follows. First, social categorization per se is insufficient as an explanation of the complexity of behaviour which follows such social categorization. Second, if group norms can be identified, as well as group members' commitment to them, then all individual and group behaviour can be satisfactorily explained in normative terms.

In Study 4 (reported in Chapter 5), a manipulation was employed to make subjects feel either better or worse about their national group and/or their membership of it than they had moments earlier. Four different self-esteem measures were administered across conditions, with subjects receiving one measure both before and after one or the other of the experimental manipulations.

A "general trait" self-esteem measure did not detect either a rise in self-esteem as a result of the positive manipulation, or a fall in self-esteem as a result of the negative manipulation. A "trait" self-esteem measure specific to national group membership similarly did not detect changes in self-esteem from either manipulation. A "general state" self-esteem measure detected self-esteem improvement as a result of the positive manipulation, but detected no significant change in self-esteem as a result of the negative manipulation. Only a measure of "state" self-esteem specific to national group membership detected both self-esteem improvement as a result of the positive manipulation and self-esteem deterioration as a result of the negative manipulation.

Strength of social identity with one's national group did not correlate with self-esteem change on any of the measures as a result of either manipulation.

A "general state" in-group evaluation measure detected improvements in in-group evaluation as a result of the

positive manipulation, but detected no significant change in in-group evaluation as a result of the negative manipulation.

The main conclusions drawn from Study 4 were as follows. First, the self-esteem hypothesis within social identity theory is best tested using a state measure of self-esteem which is contingent on (i.e. is specific to) a particular group membership. No published tests of the self-esteem hypothesis have to date employed such a measure. Second, there is a possibility that there may not be a one-to-one correspondence between evaluations of particular groups, evaluations of members of those groups, and evaluations of oneself as a member of such groups. It is also possible that threats to evaluations of particular groups may be more easily countered than threats to one's self-esteem as a member of such groups.

In Study 5 (reported in Chapter 6), self-esteem was found to be determined by an interaction of between-group and between-individual social comparison outcomes on relevant dimensions. It seemed to be the case that, when subjects' social identities were salient, both positive in-group distinctiveness and positive individual distinctiveness from in-group others on relevant dimensions made positive contributions to subjects' self-esteem. Similarly, when subjects' social identities were salient, it seemed that both negative in-group distinctiveness and negative individual distinctiveness from in-group others on relevant dimensions made negative contributions to subjects' self-esteem.

Subjects in this study reported a willingness to aid the performance of in-group others on relevant dimensions, particularly when the in-group had previously been beaten by the out-group on such dimensions. This was so regardless of whether in-group others had previously beaten or been beaten by subjects on these dimensions.

Following intergroup competition, subjects in this study reported feeling closer to in-group others than they had before the intergroup competition, particularly when the in-group had beaten the out-group during that competition. This was so regardless of whether the in-group others' performances during the competition had been inferior or superior to the subjects' own.

When the in-group had been beaten by the out-group during the intergroup competition, subjects reduced the subjective relevance of comparison dimensions relative to pre-competition levels. When the in-group had beaten the out-group during the competition, subjects neither reduced nor raised the subjective relevance of the comparison dimensions. There was no significant effect on subjective relevance change according to whether the performance of in-group others during the intergroup competition exceeded or was exceeded by the subjects' own performances.

The main conclusions drawn from Study 5 were as follows. First, contrary to Tesser's self-evaluation maintenance model, downward individual-level social comparison outcomes on relevant dimensions seem to have a beneficial effect on self-esteem. Second, again contrary to Tesser's self-evaluation maintenance model, when social identity is salient, between-group social comparison outcomes on relevant dimensions have an effect on in-group members' self-esteem levels; indeed, a greater effect than between-individual social comparison outcomes on such dimensions. Third, even when social identity is salient, the outcomes of between-individual social comparisons with in-group others do have a significant effect on self-esteem. This effect is positive when the comparisons reveal individual superiority and negative when they reveal individual inferiority. Fourth, when social identity is salient, between-individual and between-group social comparison outcomes on relevant dimensions have an interactive effect on self-esteem. Superiority on either level makes a

positive contribution to the self-esteem of in-group members, and inferiority on either level makes a negative contribution to the self-esteem of in-group members. These effects seem to be additive. Fifth, contrary to Brewer's optimal distinctiveness model, the previous conclusion suggests that it is not necessarily the case that when social identity is salient people will derive positive contributions to their self-esteem from superior performance of in-group others on relevant dimensions. Such superiority will raise self-esteem only to the extent that it contributes to positive in-group distinctiveness (or reduced negative in-group distinctiveness). However, this positive contribution to social identity contingent self-esteem may be more than cancelled out by negative contributions to self-esteem because of the comparing individual's inferiority to the superior in-group other (both in terms of personal performance, and in terms of relative contributions to the between-group outcome). Sixth, similar thoughts suggest that, contrary to Turner's (1987) self-categorization theory, personal and social identities are not mutually antagonistic: both can be salient (and psychologically operative) at the same time.

Study 6 (reported in Chapter 7) yielded several results. First, although respondents showed evaluative in-group preference in both charitable and political contexts, such in-group preference was only accompanied by out-group derogation when political in-group members evaluated out-groups which their in-groups were "traditionally" antagonistic toward. Additionally, individuals who did not identify with any particular charitable or political in-group also made favourable or unfavourable evaluations of particular groups, including derogation of certain political groups.

Second, evaluations of in-groups were shown to be affected by much more than simple attainment or non-attainment of positive in-group distinctiveness from particular out-groups. In-group evaluations were positively affected by movement toward, as

well as by attainment or enhancement of, positive in-group distinctiveness. In-group evaluations were also positively affected by movement toward, as well as by attaining or exceeding, "absolute" or temporal in-group goals. Similarly, in-group evaluations were negatively affected by movement toward, as well as "attainment" or enhancement of, negative in-group distinctiveness. And, in-group evaluations were also negatively affected by movement away from, as well as by non-attainment of, "absolute" or temporal in-group goals.

Third, use of in-group favouritism was found not to correspond with subsequent levels of social identity contingent self-esteem. Instead, attainment of positive in-group distinctiveness without a conflict of values was shown to result in positive social identity contingent self-esteem, regardless of whether or not such positive in-group distinctiveness was achieved via discrimination in favour of in-groups. Similarly, where negative in-group distinctiveness was accompanied by a conflict of values, respondents' social identity contingent self-esteem was negative, again regardless of whether or not respondents had employed in-group favouritism. Where positive in-group distinctiveness was accompanied by a conflict of values, and where negative in-group distinctiveness did not involve a conflict of values, respondents' social identity contingent self-esteem was neutral, once again irrespective of use or non-use of in-group favouritism.

Fourth, levels of social identity contingent self-esteem were found not to correspond with subsequent use of in-group favouritism. Instead, charitable group members were found to be generally unwilling to employ such discrimination, regardless of their prior levels of social identity contingent self-esteem. Political group members, on the other hand, were found to be almost universally willing to employ in-group favouritism, again irrespective of their prior levels of social identity contingent self-esteem. The main exception to

this latter finding was that members of securely negatively distinct political groups were unwilling to employ in-group favouritism when exit from those groups to more attractive ones was available. Such respondents chose exit significantly more than when such an option was available for members of securely or insecurely superior political groups, or when members' political group inferiority was insecure.

The main conclusions drawn from Study 6 were as follows. First, social categorization per se may be sufficient for evaluative in-group preference, (or at least for evaluating in-groups as positively as, if not always more positively than, other groups), but it is not sufficient for out-group derogation. Such out-group derogation may be explained, at least in some circumstances, according to in-group norms. Second, achievement of positive in-group distinctiveness, and perhaps more importantly, avoidance of negative in-group distinctiveness, may be best thought of in terms of being in-group goals. Movement toward and exceeding such goals, as well as attaining them, seems to contribute positively to in-group evaluations and to social identity contingent self-esteem. Similarly, movement away from such goals, as well as non-attainment of them, seems to contribute negatively to in-group evaluations and to social identity contingent self-esteem. Importantly, between-group relational (i.e. in-group distinctiveness) goals are not the only goals in-groups can have. They can also have "absolute" or temporal ones. Thus, it seems to be the case that attainment of, movement toward, and/or exceeding any in-group goals makes positive contributions to in-group evaluations and social identity contingent self-esteem. Similarly, it appears that non-attainment of and/or movement away from any in-group goals may make negative contributions to in-group evaluations and social identity contingent self-esteem.

The third main conclusion drawn from Study 6 is that corollary 1 of the self-esteem hypothesis within social identity theory

(Abrams & Hogg, 1988) should be framed in terms of the relationship between securely legitimate in-group distinctiveness and subsequent social identity contingent self-esteem, rather than in terms of use of intergroup discrimination and subsequent self-esteem. Similarly, corollary 2 of the self-esteem hypothesis within social identity theory (Abrams & Hogg, 1988) should be framed in terms of social identity contingent self-esteem and subsequent efforts to achieve or retain legitimate and secure contributions to self-esteem, rather than in terms of self-esteem and subsequent use of intergroup discrimination. With respect to both of these corollaries, the possibility of "conflicts of values" accompanying actual or potential in-group distinctiveness requires particular attention. In turn, this seems once again to point toward the need for social identity theory to take in-group norms into account when explaining and predicting the consequences of group membership.

IMPLICATIONS OF KEY FINDINGS FOR TAJFEL'S THEORY

To the extent that the key findings reported in this thesis are reliable and valid, they have quite dramatic implications for Tajfel's social identity theory.

Evidence supporting social identity theory

The least dramatic findings (in the sense that they suggest that no theoretical change is required), are those which support various aspects of social identity theory. Perhaps the most important of these are the findings which support social identity theory's notion that uniformities of social behaviour can stem from in-group members acting in terms of particular common social identities (rather than in terms of

disparate personal identities, or in terms of social identities other than those associated with a common group membership). This was most clearly demonstrated in Study 3 (reported in Chapter 4). In this study it was relatively common for members of particular sub-groups to adopt unique strategies in the first individual allocation situation (i.e. Condition 1), but to act in ways more or less identical to other in-group members once intragroup communication (in Condition 2) allowed them to reach consensus concerning what constituted in-group appropriate behaviour. The notion that in-group members identifying with their social groups can lead to uniformities of social behaviour also received some support from Study 1 (reported in Chapter 2). In that study, all subjects who engaged in in-group favouritism also explicitly indicated that they identified with their imposed minimal group membership.

Another important result which supports social identity theory is the finding in Study 5 (reported in Chapter 6) that when particular social identities are salient, the self-esteem of in-group members can be at least partially determined by between-group social comparison outcomes on dimensions valued by in-group members. Positive in-group distinctiveness on such dimensions made a positive contribution to in-group members self-esteem, and negative in-group distinctiveness on such dimensions made a negative contribution to in-group members self-esteem, exactly as predicted by social identity theory. The general notion within social identity theory that individuals' self-esteem can be affected by evaluations of their in-groups also received support from Studies 4 and 6 (reported in Chapters 5 and 7, respectively).

A third area of support for social identity theory came from the finding in Study 6 (reported in Chapter 7) that members of charitable and political in-groups always evaluated their own groups at least as positively as they evaluated other groups, and indeed often showed considerable preference for their own

groups over others. This supports social identity theory's postulate that people generally derive satisfaction from their in-groups and membership of them, because when they do not they either leave those groups or try to improve the situation of them (e.g. via strategies of in-group favouritism) so that satisfactory self-esteem can be restored.

This does not entail that people always evaluate every aspect of all their group memberships positively. It was found in Study 6 (reported in Chapter 7) that subjects experienced low or negative social identity contingent self-esteem when they were members of situationally inferior groups and/or were members of groups whose situation was associated with a conflict of values. This is exactly what social identity theory would predict. Further, there was also evidence from Study 6, again consistent with social identity theory, that when social identity was securely unsatisfactory for subjects, those subjects would tend to wish to leave their groups to join more favourable ones where such a strategy was possible.

Findings such as these suggest that Tajfel's social identity theory is to a considerable extent both a valid and a valuable theory. Perhaps its primary theoretical contribution is to stress the importance of social identity for: (i) bringing about and directing certain uniformities of group and intergroup behaviour (such as evaluative in-group preference except when conflicts of values or secure and legitimate negative in-group distinctiveness - or perhaps other forms of perceived in-group "failure" - make such in-group preference impossible); and, relatedly, (ii) making group members' psychological well-being at least partially dependent upon the fate of their in-groups. In addition to the results above supporting various aspects of social identity theory, however, many more findings suggest that other aspects of that theory require and/or could benefit from revision or development.

Evidence suggesting necessary and/or desirable revisions to social identity theory

Social categorization and subsequent in-group favouritism

At repeated points in this thesis social categorization per se (plus identification as group members by those categorized) has been shown to be insufficient to promote (and therefore explain) intergroup discrimination (see Studies 1, 2, and 3). This is consistent with previous research where: (i) sample-level in-group favouritism is not inevitably found following mere social categorization (see Mullen et al., 1992); and (ii) substantial proportions of minimal group subjects do not individually engage in the in-group favouritism which occurs at a mean sample-level (e.g. Tajfel et al., 1971).

Such findings present certain difficulties for social identity theory. First, the theory must "explain away" all non-discriminatory behaviour following mere social categorization, for example in terms of individual differences in self-esteem. This is unfortunate considering the wide variety of individual behaviour which seems to regularly follow such mere social categorization. Additionally, where in-group favouritism does follow social categorization per se, it appears that at least two distinct forms of intergroup discrimination can be distinguished (see below). Thus, even where social categorization does result in discrimination in favour of in-groups, it appears that something other than social categorization per se is needed to explain which sort of intergroup discrimination occurs.

Second, the claim that mere social categorization is a sufficient cause of intergroup discrimination was one of the main justifications for needing social identity theory as a supplement to Sherif's realistic conflict theory. If, as seems to be the case, social categorization per se is not a

sufficient cause of intergroup discrimination, then the claim that negative interdependence of group goals is not a necessary cause of intergroup conflict remains unjustified.

Third, if individuals in the minimal group paradigm do not reliably engage in in-group favouritism following mere social categorization, this casts doubt on hypothesized sequence of social identity - intergroup comparison - intergroup discrimination: a sequence which forms the backbone of social identity theory. In the minimal group paradigm, it will be remembered, individuals who recognize that they are members of minimal groups (i.e. individuals for whom social identity as minimal in-group members is salient) are held to have little choice but to engage in in-group favouritism, as this is the only way they can try and ensure positive in-group distinctiveness, positive evaluation of their minimal groups and membership of them, and positive social identity contingent self-esteem. In-group favouritism in the minimal group paradigm is taken to be evidence that all of these processes are happening. But, if significant numbers of individuals do not engage in in-group favouritism following mere (accepted, multigroup) social categorization, then the empirical evidence for such a sequence is seriously undermined. Further, the fact that group members often show in-group preference outside of the minimal group paradigm cannot provide evidence of the kind required, as: (i) too much else other than mere social categorization is involved outside of that paradigm (and so explanations in terms of perceived interdependence, established norms, etc., cannot be ruled out); and, (ii) it cannot be ruled out that such in-group preference outside the minimal group paradigm represents a *reflection* of perceived existing between-group differences rather than an *action* to bring certain desired between-group differences (i.e. positive in-group distinctiveness) about.

Overall, then, if social categorization per se is not a sufficient cause of intergroup discrimination in favour of the

in-group, social identity theory needs to be modified to specify more clearly: (i) what the sufficient causes of (perhaps various types of) intergroup discrimination are; and, (ii) what the exact effects of (accepted, multigroup) social categorization are, both alone and in combination with other factors. In light of the importance of sub-group norms in Study 3 (reported in Chapter 4), and of conflicts of values in Study 6 (reported in Chapter 7), in affecting use or non-use of intergroup discrimination, it appears that in order to do these things social identity theory needs to incorporate a normative element (or to more adequately specify the normative elements already at least implicit in the theory).

Social categorization and subsequent social identity

Study 1 suggests that not only is social categorization insufficient to promote intergroup discrimination in favour of the in-group, it is also insufficient to promote social identification with the in-group. If so, this suggests that social identity theory needs to: (i) specify the conditions under which social categorization will and will not result in social identification with in-groups; and, (ii) more adequately conceptualize the distinction between social categories (which people identify themselves as members of but do not necessarily identify with - see Tajfel & Turner, 1979: 41) and social groups (which people both identify themselves as members of and necessarily identify with - see Tajfel, 1982: 485-506).

In-group distinctiveness and subsequent in-group evaluation

Various results from Study 6 strongly suggest that social identity theory needs to re-conceptualize its proposed link between in-group distinctiveness and evaluations of the in-group. It seems simply not to be the case that positive in-

group distinctiveness entails positive in-group evaluations and negative in-group distinctiveness entails negative in-group evaluations. First, consistent with the exposition of social identity theory given in Chapter 1, only in certain situations do positive and negative in-group distinctiveness reliably make positive or negative contributions to in-group evaluations, (e.g. where there are no conflicts of values - see Tajfel & Turner, 1979: 45-46), even when such in-group distinctiveness obtains relative to relevant out-groups on in-group valued dimensions. Second, in-group evaluations clearly seem to be affected by processes other than between-group social comparison outcomes (cf. Brown et al., 1991).

In-group evaluation and subsequent social identity contingent self-esteem

Studies 4 and 5 each seem to suggest that there may not be a one-to-one correspondence between evaluations of particular in-groups and the self-esteem of in-group members which is contingent upon those particular in-group memberships. This may be so because self-esteem derived from membership of a particular in-group is determined *both* by evaluation of the in-group "as a whole" (e.g. "this is a good in-group") and by one's unique position within that group relative to other in-group members (e.g. "this is not such a good in-group, but at least I have high status *within* the group"). Similarly, evaluation of an in-group "as a whole" may not affect evaluation of oneself as a member of that group if one feels that one is not personally responsible for the in-group's situation (e.g. "my group lost the competition, but not because of me"). Conversely, it is possible that an in-group member's social identity contingent self-esteem may suffer from a particular situation when in-group evaluation does not (e.g. "my group won, no thanks to me").

Such possibilities also suggest that social identity theory needs to clarify the personal identity/social identity distinction. Having a certain prestige *within* a particular group is an individual or "personal" situation which one cannot share with all other in-group members, but it still contributes to one's evaluation of one's own membership of that group (i.e. one's "social" identity).

Self-esteem and subsequent intergroup discrimination

The arguments above, and the results of Studies 4 and 6, strongly suggest that it is not the case that low or threatened self-esteem inevitably promotes intergroup discrimination, even when the self-esteem in question is contingent upon a particular in-group membership. Rather, low or threatened social identity contingent self-esteem will promote in-group favouritism when, and only when, such discrimination is perceived as part of a viable strategy for obtaining secure and *legitimate* positive contributions to in-group members' self-esteem. Discrimination will not always provide such an option (e.g. when no cognitive alternatives to the current situation exist), and alternative strategies for obtaining secure and legitimate positive contributions to in-group members' self-esteem will often be preferred to in-group favouritism (e.g. exit, or renewed efforts toward non-intergroup relational in-group goals). Social identity theory therefore needs to more clearly specify the antecedents of (perhaps various types of) in-group favouritism motivated by in-group members' self-esteem needs (i.e. social competition).

Further, Studies 1, 3 and 6 suggest that in-group favouritism may sometimes be motivated by factors associated with in-group membership/social identity other than attempts to achieve secure and legitimate social identity contingent self-esteem. In-group members may sometimes discriminate simply because they perceive it to be expected of them as group members,

perhaps even when their self-esteem suffers as a result (e.g. the subject in Study 1 who felt "pretty bad" about discriminating - see also Tajfel, 1982: 503).

Intergroup discrimination and subsequent self-esteem

The same reasoning and results strongly suggest that it is not the case that intergroup discrimination (however conceptualized) inevitably raises in-group members' self-esteem, even that part of their self-esteem which is derived from membership of discriminating in-groups. If discrimination fails to achieve secure and legitimate positive in-group distinctiveness, for example, or if intergroup discrimination is motivated by factors other than social identity contingent self-esteem needs, enhanced self-esteem for in-group members should not be expected. At the very least, it seems to be the case that it is secure and legitimate positive in-group distinctiveness, rather than intergroup discrimination per se, which contributes positively to social identity contingent self-esteem. But even this may be too bold a claim, as it has been argued above that: (i) positive in-group distinctiveness (even when legitimate and secure) is not the only route to positive in-group evaluation; and, (ii) there may not be a direct correspondence between in-group evaluation and social identity contingent self-esteem.

The implication of this and of the previous sub-section for social identity theory is clearly that the theory needs to do much more than postulate that unsatisfactory self-esteem leads to intergroup discrimination, and that intergroup discrimination leads to satisfactory self-esteem.

Conceptualizing social identity and self-esteem

Various results from the present thesis, but particularly those from Studies 4 and 5, suggest that social identity theory needs to be much more rigorous in its conceptualization of social identity, and also in its conceptualization of the link between social identity and self-esteem. In particular, care needs to be taken to delineate the specificity and temporality of social identity and self-esteem referred to in particular contexts. To give just one example, a member of a particular low-status group may have corresponding negative contributions to their social identity and self-esteem, but their overall social identity and self-esteem may be extremely satisfactory. In its present form, social identity theory can be used to predict both that the individual will discriminate in favour of that particular group (i.e. because of low group-membership contingent self-esteem), and that the individual will not discriminate in favour of that group (i.e. because their overall social identity and self-esteem is satisfactory).

Conceptualizing intergroup discrimination

Studies 2 and 6 suggest that social identity theory also needs to be more rigorous in its conceptualization of intergroup discrimination, intergroup competition, and intergroup conflict. In particular, differences between in-group preference which is or is not accompanied by out-group derogation need to be considered, as do the antecedents and consequences of each. Social identity theory also needs to distinguish more fully between "intergroup discrimination" which is a subjective reflection of existing states of affairs (e.g., in-group preference stemming from the fact that in-group members already evaluate the in-group at least as positively as they evaluate other groups), from intergroup discrimination which is an action to bring about desired

states of affairs (e.g. in-group favouritism engaged in to bring about new or enhanced secure and legitimate in-group superiority). This latter consideration also suggests that care should be taken to differentiate both the dimensions on which intergroup discrimination occurs (i.e. evaluative or behavioural), and the different possible beneficiaries of intergroup discrimination (i.e. the in-group as a whole, in-group members, oneself as an in-group member, an in-group product, etc.).

Levels of behaviour

Finally, it will be remembered from Study 6 that subjects who did not claim to be members of any particular charitable or political groups still managed to adopt certain attitudes towards such groups. That is, *individuals* made either positive or negative evaluations of certain groups they were not members of, even though they were not obviously thinking of themselves in terms of their membership of other groups. Thus, it seems that even when people are operating in terms of their personal identities, it is still possible for them to think of others in terms of their social identities.

This challenges Tajfel's notion of an interpersonal-intergroup continuum of social behaviour, in as much as these subjects seemed to be behaving across levels (i.e. person-to-group, as opposed to person-to-person, or group-to-group).¹ Further, the results of this study also suggest that there can be uniformities of person-to-group behaviour (in as much as certain groups were evaluated positively or negatively on average by individuals who did not claim to belong to any particular group). In turn, this suggests that some instances of seemingly intergroup behaviour may actually be instances of

¹ Group-to-person behaviour also seems to be possible, as, for example, when a religious group adopts a particular attitude towards a particular individual.

uniform person-to-group behaviour, for example when numbers of people individually derogate particular groups and/or group members. Such instances clearly involve uniformities of social behaviour, but such uniformities do not obviously stem from particular group memberships or social identities.

Perhaps the main conclusion to be drawn from the results considered in this sub-section is that social identity theory's main weakness is in failing to make explicit the importance of in-group norms for initiating and directing certain forms of genuinely collective (i.e. group and intergroup) behaviour. Such norms are absent, and cannot be generated, in the situation of social categorization per se, and so mere social categorization cannot (and does not) result in genuine collective behaviour. This both seriously undermines the claim that social identity theory is needed as a supplement to realistic conflict theory, and also seriously undermines much of social identity theory's empirical support for its notion of social competition (i.e. intergroup discrimination in the absence of perceived negative interdependence of group goals).

That in-group norms are a necessary precondition of genuinely collective behaviour is, in fact, implicit in social identity theory.² If one re-examines Diagrams 1.1 and 1.2 in Chapter 1, it is clear that for relative in-group status to promote genuinely collective behaviour, in-group norms are necessary to ensure that numbers of individuals (i.e. group members) *commonly*: value the comparison dimension; consider the out-group relevant; consider the between-group status differential illegitimate and/or unstable; consider exit inappropriate; and believe certain forms of group oriented behaviour (i.e. social competition and/or social creativity) necessary and/or desirable. This normative aspect of social identity theory has been underplayed, however, probably because of the

² Further, it is argued below that in-group norms entail in-group goals, e.g. conformity to those norms.

theory's account of the *motivation* for collective behaviour and of its corollary account of the *consequences* of such behaviour. In each case the theory relies on the notion of self-esteem, but the results considered here (consistent with many contemporary studies) suggest that the theoretical marriage between self-esteem and group behaviour is not as convenient as Tajfel originally proposed.

POTENTIAL LIMITATIONS INHERENT IN THIS THESIS

All of the considerations in the previous section relate to the implications of findings presented in this thesis for social identity theory, either in terms of supporting aspects of that theory, or in terms of suggesting necessary or desirable modifications to that theory. However, such considerations are only pertinent to the extent that the conceptual points made and the empirical results presented are valid. This section examines some of the more important potential limitations to such validity.

The main conceptual issue concerns whether or not the exposition of Tajfel's social identity theory presented in this thesis is in fact true to Tajfel's position. The only thing that can be said to address this point is that the exposition presented relies wholly on Tajfel's own writings. All the components of the presented exposition have cited particular texts by Tajfel, and indeed most if not all of the presented exposition can be verified by reference to only two papers: Tajfel & Turner (1979) (which essentially is a summary of Tajfel, 1978a), and Tajfel (1978c).

A second conceptual issue is that the present thesis has paid scant regard to the various proposed modifications of and developments to social identity theory since Tajfel (especially perhaps Turner's self-categorization theory). One

response to this is simply to say that the thesis is explicit from the title forward that it concerns only Tajfel's version of his own theory. As mentioned in Chapter 1, many of the proposed modifications of and extensions to social identity theory have failed to be adequately distinguished from or integrated with that theory. As a consequence, unless one focuses on Tajfel's version of the theory, or upon relatively clearly distinct descendants of it (such as self-categorization theory), it is very difficult to ascertain what "the" current content of social identity theory is. And, as contemporary developments of Tajfel's theory tend to contain large chunks of the original theory as component parts, and therefore rely in part for their own validity upon the validity of the original theory, the original theory was the focus of this thesis.

There is one clear conceptual limitation to the present thesis. This is that certain aspects of that theory have received either scant or no attention, e.g., Tajfel's notion of accentuation of between-group differences and of between-group similarities, and the subsequent notions of perceived in-group and/or out-group homogeneity, social stereotyping, etc.. The only possible response to this limitation is to acknowledge that it exists.

Principal methodological limitations of the present thesis concern sampling and measurement. In attempting to evaluate social identity theory (more or less) "as a whole", empirical results relating to any specific part of that thesis are clearly insufficient to do more than suggest areas in which that theory may receive support or may need modification. Sample sizes for almost all of the studies are relatively small, and the vast majority of the subjects involved were white, British university undergraduates, studying psychology. Clearly, none of the studies standing alone would do much to convince anyone of the validity of the conclusions drawn from those studies.

Additionally, many of the measures used in the studies presented in this thesis are of uncertain (or, in the case of the Tajfel matrices, contested) reliability and validity. These problems were dealt in some depth in each of the relevant chapters, so perhaps all that is needed here is an illustration. In Chapter 5 an attempt was made to develop four measures of different "types" of self-esteem from a single measure of general trait self-esteem (i.e. Rosenberg, 1965). This procedure was chosen so that the four measures would be as equivalent as possible, and because of the popularity and known psychometric qualities of that "template" measure (Wylie, 1989). Apart from calculating internal reliability, no attempt was made to investigate the reliability and validity of any of these measures. Bearing in mind the complexity of the wording which was required to try and measure such concepts as state self-esteem specific to a particular group membership, it seems unwise to place too much faith in the psychometric qualities of such measures. (Similar, or even more damning, comments could be made about the probable reliability and validity of the coding frame used in Study 1; the allocation grids employed in Study 2; the self-esteem, closeness, relevance, and performance measures used in Study 5; and the whole host of measures adopted during Study 6.)

Apart from referring to the large number of a priori hypotheses which received support during the studies reported in this thesis, all that can be done in relation to the methodological limitations is to accept that they exist, accept that any conclusions reached on the basis of single studies must therefore be tentative, and to suggest that future research employs more sophisticated measures.

TOWARD A MODIFIED SOCIAL IDENTITY THEORY

Accepting for the sake of argument that the results presented in this thesis have some validity, some speculative comments can be made about how social identity theory might look if it were modified in the light of these results (and of other research). In this section some attempt will be made to provide a "bare bones" account of a modified version of social identity theory which would be consistent with much if not most prior research in this area. Then, in the next section, this modified account will be critically compared and contrasted with Tajfel's original theory.

A modified account of social categories

At a subordinate level, social categories are (actual or potential) collections of individuals who are perceived by one or more people to have at least one characteristic in common, i.e. the mere fact of category membership. Social category members need not be perceived to have any other characteristic in common, even the extent to which they are perceived as category members (Lakoff, 1987; Rosch, 1978). Social category members (actual or potential) will each have many individual characteristics other than those determined or entailed by social category membership, and the nature of social categories will in part be determined by all of the characteristics individual social category members have.

Individuals who are categorized by others as members of a particular social category do not themselves necessarily need either to know this or, if they do know it, to accept their category membership.

When individuals accept that they are members of particular social categories, and/or when individuals believe that others perceive them to be members of particular social categories

(even if they do not themselves accept this), these cognitions can affect those individuals' subsequent cognitions, emotions, and behaviour (including social behaviour). For example, individuals may infer self-attributes from attributes perceived to be associated with social categories which they accept themselves as members of, and these self-ascriptions may entail particular cognitive, emotional, evaluative, and/or behavioural consequences. Or, individuals may believe that others who perceive them as members of particular social categories will consequently evaluate and act toward them in particular ways, and these expectations may themselves entail particular cognitive, emotional, evaluative, and/or behavioural consequences.

If numbers of individuals each accept their own membership of particular social categories, and/or believe that others perceive them to be members of particular social categories, and if each of these individuals responds to such situations in similar ways (cognitively, emotionally, evaluatively, and/or behaviourally), then we have instances of uniformities of social behaviour (broadly conceived) resulting from social categorization (broadly conceived).

At a superordinate level, social categories are cognitive entities "in their own right", with their own superordinate qualities. Social categories are not sentient, they have neither goals nor intentions, and they do not have the means to fulfil such goals and intentions were they to exist. Nevertheless, social categories do "exist in their own right". They can be evaluated and behaved toward in ways unconnected with the evaluations of and behaviours toward the members of those social categories, and they have qualities that their members do not.

A modified account of social groups

Social groups are particular types of social categories and, as such, have all of the qualities and limitations of social categories expounded in the previous sub-section, with one critical exception.³ The essential difference between social categories which are groups and those which are not is that the former are perceived to have the additional characteristic of having at least one superordinate-level goal. Social groups are no more sentient than are any other social categories, and, at a superordinate level, they neither have goals, know that they are perceived to have goals, nor have the means to strive toward such goals.

However, when one or more *individuals* perceive a social category to have one or more goals, that social category is a social group for those individuals. As with other social categories, the individuals who perceive social groups may also perceive themselves either as members or as non-members of those social groups. Again, as with other social categories, people may believe that others perceive them as members of particular social groups, but may themselves deny that they are. Similarly, again as with other social categories, when individuals accept that they are members of particular social groups, and/or when individuals believe that others perceive them to be members of particular social groups (even if they do not themselves accept this), these cognitions can affect those individuals' subsequent cognition, emotion, evaluation, and behaviour (including social behaviour). Finally, and again in common with other social categories, if numbers of individuals each accept their own membership of particular social groups, and/or believe that others perceive them to be members of particular social groups, and if each of these individuals responds to such situations in similar ways (cognitively, emotionally, evaluatively, and/or

³ But see also the sub-section below entitled "A modified account of group behaviour".

behaviourally), then we have instances of uniformities of social behaviour (broadly conceived) resulting from social group membership (broadly conceived). However, because social groups necessarily have superordinate goals associated with them, whereas other social categories necessarily do not, uniformities of social behaviour are considerably more likely as a result of social group membership than as a result of social category membership.

A modified account of social identity

Overall (i.e. "general") social identity comprises of that part of an individual's self-concept which derives from: (i) subjective membership of social categories and groups; and, (ii) beliefs that others perceive one as a member of certain social categories and groups. Overall social identity is also the totality of specific social identities an individual has. A specific social identity is that part of an individual's self-concept which derives from: (i) subjective membership of a particular social category or group; and/or, (ii) a belief that others perceive one to be a member of a particular social category or group.

Specific social identities have cognitive and evaluative components. For any specific social identity, the cognitive component of that identity consists of a belief that one is a member of a particular social group or category; and/or the belief that one is perceived by others to be a member of a particular social group or category; and also, beliefs concerning the consequences of such membership (subjective and/or other-perceived) of the social group or category of concern. For any specific social identity, the evaluative component of that identity consists of all the positive and negative consequences of membership (subjective and/or other-perceived) of the social group or category which provides the identity. (The evaluative component of specific social

identities will be determined, at least in part, by evaluating actual subjective and/or other-perceived category or group memberships relative to other possible subjective and/or other-perceived conceptions of the self.)

The cognitive and evaluative components of overall social identity are determined by the totality of cognitive and evaluative components of all the specific social identities an individual has.

A modified account of the relationship between social and personal identity

When referring to an individual's particular specific social identity, the term "personal identity" can be used to specify all the characteristics that the individual has which are not determined or entailed by membership of the particular social category or group which the specific social identity is derived from. This is potentially misleading, however, as many of these characteristics will be determined or entailed by the individual's membership of other social categories and groups.

It makes even less sense to try to distinguish between an individual's overall social and personal identities. If overall social identity is that part of an individual's self-concept which is derived from membership of all the social groups and categories they are a member of (and/or believe others perceive them to be a member of), then there is precious little left over, if anything, to call overall personal identity. What characteristics are there which cannot be conceived as associated with membership of social groups or categories (e.g. extroverts, tall people, people with swimming certificates, people called Tom, etc.)?

As conceptualized here, overall personal identity refers to the unique totality of an individual's self-concept. Overall social identity refers to that unique part of an individual's overall self-concept (i.e. overall personal identity) which is derived from membership of social categories and groups and/or from beliefs that others perceive one as a member of certain social categories and groups. Specific personal identities refer to those parts of an individual's self-concept which derive from particular characteristics that individual has. Finally, specific social identities are those parts of an individual's self-concept which derive from membership of particular social categories and groups, and/or from beliefs that others perceive one as a member of particular social categories or groups. Thus, all other identities are nested within overall personal identity, and specific social identities are simply collections of specific personal identities which are associated with membership of particular social groups or categories, and/or with beliefs that others perceive one as a member of such social categories or groups.

A modified account of group behaviour

Social groups and other social categories can have descriptive norms associated with membership of those social groups and other social categories. That is, proportions of members of social groups and other social categories can have particular characteristics in common to such an extent that those characteristics become associated with membership of those social groups and other social categories.

Social groups, but not other social categories, can also have *prescriptive* social norms. Social categories with at least one superordinate goal become social groups. Prescriptive social norms associated with superordinate social units entail superordinate goals (i.e. conformity to the prescribed norm), and so superordinate units with prescriptive superordinate

norms must therefore be social groups. Conversely, superordinate goals entail superordinate norms (i.e. pursuit of the goals). Thus, we can further distinguish social groups from social categories by saying that social groups are social categories with superordinate-level goals and prescriptive norms (i.e. group goals and prescriptive norms).

Group behaviour can now be defined as the totality of individuals' behaviour which is determined by attempts by those individuals to conform to the prescriptive norms associated with a particular group they are members of.

By this account, group behaviour is, at a subordinate level, simply a particular type of individual behaviour (or, rather, the totality of certain aspects of individuals' behaviour). Group behaviour occurs when *individuals* strive to make progress toward goals associated with particular groups those individuals belong to. Nevertheless, when engaging in group behaviour, individuals are in some sense "serving the needs" of the groups they belong to. Therefore, although it is individual group members who are providing the means for attempted satisfaction of group goals, it still makes sense to talk (at a superordinate level) about the group "itself" behaving. For example, when England declared war against Germany in 1939, it was an individual who did the declaring. Chamberlain did not personally declare war against Germany (or even against Hitler). Rather, Chamberlain declared war against Germany *on behalf of* a particular group he was a member of (i.e. England). At a superordinate level, therefore, it makes sense to say that *England* (i.e. a group) declared war on Germany.⁴

⁴ This account presupposes National-level goals which Chamberlain was pursuing (possibly including goals specifying conformity to National-level prescriptive norms). Discourse analytically inclined social psychologists might profitably study group representatives' accounts of, or justifications for, decisions they have taken on behalf of their groups, and see how often such accounts take the form of pursuing group goals and/or conforming to prescriptive group norms.

A modified account of uniformities of social behaviour

The phrase "uniformities of social behaviour" refers to numbers of individuals behaving in similar ways in social situations. Such uniformities can come about via all sorts of processes, including habit, conformity, innate or commonly learned responses to particular social stimuli, innate or commonly learned needs, etc.. However, the uniformities of social behaviour of interest here are those which stem from group membership.

When numbers of individuals strive toward superordinate goals of groups they belong to, uniformities of social behaviour are likely to result. The more salient, unambiguous, and important a group's goals, the greater the consensus among group members concerning how to pursue them, and the more group members wish to promote progress toward those goals, the more uniformities of social behaviour can be expected.

The existence of group goals promotes uniformities of social behaviour when numbers of group members try to satisfy those goals. When numbers of group members do this, group behaviour results. Only uniformities of social behaviour which result from group members striving to satisfy group goals should be considered "genuine" group behaviour. Thus, numbers of group members behaving in similar ways because of common individual responses to certain social stimuli (e.g. group membership per se), although interesting and important, should not be considered instances of group behaviour. Because such individuals are "serving their own needs", rather than "serving the needs" of the group, their behaviour is not group behaviour. (Although, of course, it may look like group behaviour, both from within and without the group, and indeed may become group behaviour, if individuals interpret the uniformities of social behaviour by other members of their

group as providing a prescription for how group members should act.)⁵

A modified account of group evaluation

Because the characteristic which distinguishes social groups from other social categories is the existence of superordinate-level (i.e. group) goals, it is possible to conceive of the notion of "satisfactory group realization". Satisfactory group realization is the superordinate-level equivalent of Tajfel's (1978: 9) notion of satisfactory self-realization (see next sub-section). Satisfactory achievement of, maintenance of, or progress toward group goals promotes satisfactory group realization. Conversely, unsatisfactory maintenance of, or unsatisfactory progress toward, group goals promotes unsatisfactory group realization.

At a superordinate level, groups will be evaluated positively to the extent that they achieve satisfactory group realization, and groups will be evaluated negatively to the extent that they do not.

Group goals can specify the sorts of situations groups would like to be in (which might be termed "terminal" group goals); the means by which groups would like to be in such situations (which might be termed "procedural" group goals); and/or, the sorts of groups groups would like themselves to be (which might be termed "existential" group goals). Satisfactory group realization, and therefore superordinate-level positive evaluation of groups, will only reliably be achieved to the extent that each of these sorts of goals (where each exists) are adequately satisfied.

⁵ As descriptive norms can become prescriptive ones.

For example, if a group suddenly and novelly finds itself in a desired situation (i.e. satisfaction of a terminal group goal), this will not reliably promote satisfactory group realization (or, therefore, positive evaluation of the group at a superordinate level) if the group wanted to be, but was not, responsible for getting itself into this desired situation (i.e. non-satisfaction of a procedural group goal). In this example, although the in-group situation (i.e. "where the in-group is") may be positively evaluated, the in-group itself is not. Similarly, if through its own efforts a group achieves a desired situation (i.e. satisfaction of terminal and procedural group goals), but this achievement is accompanied by an unacceptable change in the nature of the group (i.e. non-satisfaction of an existential group goal), this will also not reliably promote satisfactory group realization (or, therefore, positive evaluation of the group at a superordinate level). In this example it is again the case that the in-group situation is positively evaluated (both in terms of "where the in-group is" and "how it got there"), but the in-group itself is not.

At a superordinate level, groups are evaluated positively "overall" to the extent that satisfactory progress toward certain group goals exceeds unsatisfactory progress toward others. Similarly, at a superordinate level groups are evaluated negatively "overall" to the extent that unsatisfactory progress toward certain group goals exceeds satisfactory progress toward others. However, it may sometimes be misleading to talk of "overall" evaluations of groups, as certain aspects of groups may be evaluated positively while other aspects of groups are simultaneously evaluated negatively: it is possible to have ambivalent evaluations of groups. It is therefore often preferable to talk of the "totality of", rather than "overall", group evaluations.

A modified account of the consequences of social category membership

Individuals strive for "satisfactory self-realization" (Tajfel, 1978c: 9). Satisfactory achievement of, maintenance of, or progress toward an individual's goals promotes satisfactory self-realization for that individual. Unsatisfactory maintenance of an individual's goals promotes unsatisfactory self-realization, as do unsatisfactory progress toward or away from such goals.

Individuals evaluate themselves positively to the extent that they achieve satisfactory self-realization, and evaluate themselves negatively to the extent that they do not.

As with group goals, individual goals can specify the sorts of situations individuals would like to be in (i.e. "terminal" individual goals); the means by which individuals would like to be in such situations (i.e. "procedural" individual goals); and/or, the sorts of people individuals would like themselves to be (i.e. "existential" individual goals). Satisfactory self-realization, and therefore positive self-evaluation, will only reliably be achieved to the extent that each of these sorts of goals (where each exists) are adequately satisfied.

A large part of an individual's self-concept comprises of their social identity. Thus, an individual may desire membership of certain social categories (including social groups), either as an end in itself, or as a means to satisfaction of other self-goals the individual has. Similarly, but distinctly, an individual may desire to avoid membership of certain social categories (including social groups), either as an end in itself, or as a means to satisfaction of other self-goals the individual has. Individuals may also desire that others perceive them to be members of certain social categories (including social groups), again either as an end in itself, or as a means to

satisfaction of other self-goals the individual has. Finally, individuals may desire to avoid others perceiving them as members of certain social categories (including social groups), either as an end in itself, or as a means to satisfaction of other self-goals the individual has.

Thus, an individual may have membership and/or non-membership (actual and/or perceived by others) of certain social groups and of other social categories as self-goals. Therefore, membership and/or non-membership (actual and/or perceived by others) of certain social groups and of other social categories can have a strong bearing on individuals' satisfactory self-realization, and on their subsequent self-evaluation. In turn, where people have such social category membership and/or non-membership goals, those people will be motivated to make satisfactory progress toward such goals, will obtain positive contributions to their self-evaluations to the extent that they achieve satisfactory progress toward such goals, and will obtain negative contributions to their self-evaluations to the extent that they make unsatisfactory progress toward such goals.

It is because membership (and/or non-membership, actual and/or other-perceived) of certain social categories (including social groups) has a bearing on an individual's self-concept, and consequently on their satisfactory self-realization, that such memberships have the cognitive, emotional, evaluative, and/or behavioural consequences they do (see sub-section above entitled "A modified account of social categories"). But, because individuals can have wildly different personal goals, individuals can also differ greatly in their cognitive, emotional, evaluative, and/or behavioural responses to membership (and/or non-membership, actual and/or other-perceived) of any particular social group or other social category. Thus, predicting uniformities of social behaviour (broadly conceived) as a result of particular social category memberships (broadly conceived) requires either that

individuals have similar personal goals and respond to certain social category memberships in similar ways to pursue those goals,⁶ and/or that individuals have different personal goals but nevertheless respond to certain social category memberships in similar ways to pursue those goals. In each of these instances, however, we are talking about *individual* behaviour: i.e., behaviour aimed at satisfying exclusively *individual* goals. Any uniformities of social behaviour which result from social category memberships in such cases are "simply" the consequence of common individual goals and/or common action aimed at satisfying those individual goals.

Fortunately, there are circumstances in which we can predict with a fairly high degree of confidence that members of certain social categories will have common individual goals, and will also share some degree of consensus concerning how to go about satisfying such goals. These are situations in which members of social groups strive to achieve common superordinate-level (i.e. group) goals.

A modified account of the consequences of social group membership

Besides having "purely" personal goals, individuals can also identify with the goals of others. When individuals identify with the goals of other individuals, the identifying individuals' satisfactory self-realization will depend, in part, upon the perceived satisfactory self-realization of the identified-with individuals. Similarly, when individuals identify with the goals of groups, those individuals'

⁶ For example, the "popular" understanding of Tajfel's social identity theory says that individuals respond to any and all social category memberships in similar ways (i.e. intergroup discrimination) in order to satisfy common personal goals they have (i.e. obtaining positive in-group distinctiveness in order to obtain satisfactory social identity and self-esteem).

satisfactory self-realization will depend, in part, upon the satisfactory group realization of those groups. When individuals identify with the goals of others, be those others individuals or groups, the identifying individuals take on the others' goals as goals for themselves (i.e. "personal" goals). This does not necessarily entail that the identifying individuals feel any responsibility to bring about satisfactory self- or group realization of the identified-with others. It only means that the identifying individuals' satisfactory self-realization is in some sense "tied-up" with the satisfactory self- or group realization of the identified-with other. That said, it will usually be the case that when one identifies with the goals of another (individual or group), one also takes on some personal responsibility to bring those goals about, if only because one's satisfactory self-realization is partially dependent upon the satisfactory progress by the other toward their goals. When one does take on such responsibility, then one's satisfactory self-realization will be affected both by the others' progress toward or away from their own goals, and by one's own contributions to that progress.

What this means is that when individuals identify with the superordinate (i.e. group) goals of groups they perceive themselves as belonging to (i.e. in-groups), those individuals' satisfactory self-realization will be determined, in part, by the satisfactory group realization of those in-groups (i.e. by satisfactory in-group realization). To the extent in-group members take on responsibility for the attainment of satisfactory in-group realization, those group members' satisfactory self-realization will be partially determined both by satisfactory in-group realization, and by their own perceived contribution to that satisfactory in-group realization.

To the extent, therefore, that individuals identify with in-group goals, individuals will be motivated to strive for

satisfactory in-group realization, in order to achieve contributions to satisfactory self-realization. This is the mechanism by which group behaviour comes about. Numbers of individual members of a particular group accept their group's goals as "personal" goals, and strive to make satisfactory progress toward them.

It was noted above that, at the superordinate level, groups are evaluated positively to the extent that they achieve satisfactory group realization, and are evaluated negatively to the extent that they do not. It has also been noted that when individuals identify with in-group goals, those individuals' satisfactory self-realization is partially determined by satisfactory in-group realization. However, it is not the case that superordinate evaluations of in-groups fully determine either: (i) individual in-group members' subordinate-level evaluations of their in-groups; or, (ii) in-group members' evaluations of themselves as particular in-group members. In each case, this is because individual in-group members' goals are usually not fully exhausted by the superordinate-level goals of the in-groups they identify with.

Group members, as individuals, have all manner of personal goals which are associated with in-group membership, only a fraction of which are those deriving from identification with in-group superordinate goals. To take a simple example, imagine that an individual has the following personal goals associated with a particular in-group membership: (i) the achievement of satisfactory in-group realization; (ii) to be personally responsible for bringing such satisfactory in-group realization about; (iii) to remain as leader of the in-group; and, (iv) to enjoy the respect of other in-group members. Let us say that this individual takes action to ensure that the in-group achieves its goals, but in the process of doing so loses their leadership of the group and the respect of other in-group members. Because it achieves its goals, the in-group achieves satisfactory group realization and will be positively

evaluated at a superordinate level. Because the individual identifies with the group's goals, and also satisfies a personal goal of bringing those group goals about, the individual will obtain positive contributions to their satisfactory self-realization. However, in losing both the leadership of the group and the respect of other in-group members, the individual fails to satisfy two important personal goals (which are related to in-group membership), and therefore also receives negative contributions to their sense of satisfactory self-realization. Thus, it is quite possible for the individual to evaluate the in-group negatively at the subordinate level, even though the same individual accepts that the in-group achieved all of its goals (and is therefore evaluated positively at the superordinate level).

Now consider the case of an individual whose in-group achieves satisfactory group realization, but whose individual contribution to that group achievement was negligible. If that individual had a personal goal of contributing significantly to the in-group achievement, that individual may evaluate the in-group positively at both the superordinate and subordinate levels, but still evaluate themselves (as a particular in-group member) negatively.

The usual mismatch between group-level goals and the goals of group members means that in-group evaluations at superordinate and subordinate levels are rarely identical. It also means that evaluations of particular in-group members (including the self) can be distinct from either or both of these evaluations of the in-group. At a superordinate level, in-groups are evaluated positively or negatively according to whether or not they have achieved satisfactory group realization. At a subordinate level, in-groups are evaluated positively or negatively according to whether membership of them contributes positively or negatively overall to individual group members' satisfactory self-realization. Finally, in-group members evaluate themselves as particular individual group members

(i.e. "within the group") positively or negatively according to whether they contribute positively or negatively to the satisfactory group realization of their in-groups.⁷

The mis-match between group goals and the goals of individual group members has one more important consequence. This is that, even if they identify with them, group members will not strive toward perceived in-group goals which are incompatible with other goals which those individuals hold to be personally more important. Group behaviour will therefore only come about when conformity to perceived group goals by in-group members does not entail violation of other goals which those in-group members are more committed to (which may or may not be goals associated with membership of the group in question).

A modified account of the effects of in-group distinctiveness

Groups can have superordinate-level goals which involve striving to achieve, maintain, or enhance positive in-group distinctiveness from particular comparison groups on particular dimensions, and/or striving to avoid, reduce or reverse negative in-group distinctiveness from particular comparison groups on particular dimensions. Groups can also have goals which are not explicitly concerned with in-group distinctiveness, but which nevertheless consider such in-group distinctiveness as indicative of satisfactory or unsatisfactory progress towards the goals they do have. To the extent that perceptions of in-group distinctiveness entail or suggest satisfactory progress toward in-group goals, positive contributions to superordinate evaluations of in-groups will result from such perceptions. Conversely, to the extent that perceptions of in-group distinctiveness entail or

⁷ Although, of course, in-group members' evaluations of themselves as individuals (i.e. "in total") will depend on much more than this, i.e. on their "overall" satisfactory or unsatisfactory self-realization.

suggest unsatisfactory progress toward in-group goals, negative contributions to superordinate evaluations of in-groups will result from such perceptions. However, to the extent that perceptions of in-group distinctiveness neither entail nor suggest either satisfactory or unsatisfactory progress towards in-group goals, such perceptions will not influence superordinate evaluations of in-groups.

Similar considerations apply with respect to the effects of perceptions of in-group distinctiveness on subordinate evaluations of in-groups. To the extent that perceptions of in-group distinctiveness entail or suggest satisfactory progress toward group members' goals, positive contributions to subordinate evaluations of in-groups will result from such perceptions. Conversely, to the extent that perceptions of in-group distinctiveness entail or suggest unsatisfactory progress toward in-group members' goals, negative contributions to subordinate evaluations of in-groups will result from such perceptions. And, to the extent that perceptions of in-group distinctiveness neither entail nor suggest either satisfactory or unsatisfactory progress towards in-group members' goals, such perceptions will not influence subordinate evaluations of in-groups.

Finally, similar considerations apply with respect to the effects of perceptions of in-group distinctiveness on individual group members' evaluations of themselves in terms of their particular in-group memberships. To the extent that in-group members perceive that they have personally contributed toward in-group distinctiveness which entails or suggests satisfactory progress toward in-group goals, positive contributions to their self-evaluations as in-group members will result from such perceptions. To the extent that in-group members perceive that they have personally contributed toward in-group distinctiveness which entails or suggests unsatisfactory progress toward in-group goals, negative contributions to their self-evaluations as in-group members

will result from such perceptions. Finally, to the extent that in-group members believe that perceived in-group distinctiveness neither entails nor suggests either satisfactory or unsatisfactory progress towards in-group goals, such perceptions will not influence their self-evaluations as in-group members, regardless of their perceived personal contributions to that in-group distinctiveness.

It follows from the above, but should be made explicit, that evaluations of in-groups and in-group members can be affected by situations other than ones of perceived in-group distinctiveness. Any situations which entail or suggest satisfactory or unsatisfactory progress toward in-group goals will affect superordinate evaluations of in-groups; any in-group situations which entail or suggest satisfactory or unsatisfactory progress toward in-group members' goals will affect subordinate evaluations of in-groups; and, any situations which entail or suggest that individual in-group members have contributed positively or negatively to satisfactory or unsatisfactory in-group realization will affect those individual in-group members' evaluations of themselves as particular in-group members.

A modified account of the causes of intergroup discrimination

Groups will prescribe that their members should engage in intergroup discrimination (or any other form of behaviour) when such discrimination appears likely to promote satisfactory in-group realization. Groups will prescribe that their members should not engage in intergroup discrimination (or any other form of behaviour) when such discrimination appears unlikely to promote satisfactory in-group realization, especially when such discrimination seems likely to promote unsatisfactory in-group realization.

Group members will engage in intergroup discrimination (or any other form of group behaviour) when engaging in such intergroup discrimination appears to likely to promote both satisfactory in-group realization and satisfactory self-realization for those group members. (When individuals engage in behaviour intended to promote satisfactory self-realization regardless of the likely effects on in-group realization, such behaviour should not be thought of as group behaviour.) Group members will not engage in intergroup discrimination (or any other form of group behaviour) when engaging in such intergroup discrimination does not appear likely to promote both satisfactory in-group realization and satisfactory self-realization for those group members, especially when engaging in such discrimination seems likely to promote unsatisfactory self-realization for those in-group members.

As superordinate entities, groups do not have the power to promote their own goals, except by "encouraging" in-group members to strive for those goals on the groups' behalf. They achieve this to the extent that in-group members' satisfactory self-realization is or becomes dependent upon satisfactory in-group realization. However, because group goals and the goals of in-group members are rarely identical, group members' will not slavishly conform to in-groups' "wishes". Rather, group members act on behalf of their groups to the extent that doing so contributes to their satisfactory self-realization (i.e. to the extent that the consequences of acting on behalf of the group seem personally more attractive than the consequences of failing to do so). Clearly, the more that groups can improve in their favour the cost/benefit ratio for in-group members who do or do not conform to in-group wishes (such as intergroup discrimination), the more likely behaviour prescribed by the group will be.

**A COMPARISON OF THE ORIGINAL AND MODIFIED
VERSIONS OF SOCIAL IDENTITY THEORY**

Although the considerations in the above section are conceived individually as possible amendments to Tajfel's social identity theory, rather than conceived jointly as a replacement for it, for aid of exposition the account above will be referred to as "goal identification theory".

The first thing to note is that the two theories share considerable amounts of common ground. Among other similarities, both accept or claim that individuals often conceive of themselves and others in terms of their membership of social groups and categories; that membership of social groups and categories has cognitive, emotional, evaluative and behavioural consequences; that uniformities of social behaviour, including intergroup discrimination, can stem from membership of certain social categories in particular situations; that both individual and superordinate-level factors need to be used to explain and predict uniformities of social behaviour stemming from group membership (i.e. self-esteem and in-group distinctiveness in the case of social identity theory, and satisfactory self-realization and satisfactory group realization in the case of goal identification theory); that both identity and "realistic" concerns can motivate group behaviour; that between-group social comparison outcomes (e.g. perceptions of in-group distinctiveness) can affect group members' cognitions, evaluations, and behaviours; that people in unfavourable social group or category membership situations can either try to leave those groups or categories, or to take action to improve membership of them; and that group norms are important in specifying the limits of group membership motivated behaviour.

However, there are clearly several important differences between the two accounts. The first of these differences

concerns the postulated individual motivation for group behaviour: i.e. self-esteem in the case of social identity theory, and satisfactory self-realization in the case of goal identification theory. Despite the problems associated with the formulation and previous testing of the self-esteem hypothesis within social identity theory (see Chapters 5 and 7), there can be no doubt that evidence for that hypothesis is mixed at best, and that several adherents of social identity theory are now searching for alternative accounts for the motivation behind group behaviour (e.g. Abrams, 1992). Goal identification theory provides such an alternative, and does so in a way that accepts that group members' self-esteem needs can motivate and be affected by group action, but nevertheless argues that it need not necessarily be. It does this by postulating that self-esteem is one, non-necessary, component of satisfactory self-realization.

This is consistent with Tajfel's own original thoughts (Tajfel, 1978c). Tajfel, however, rejected the notion of satisfactory self-realization as "a hopelessly vague, synthetic term which can mean so much that it is in danger of meaning very little at all" (p. 9). In response to this it can be claimed that the notion of self-esteem, as employed within social identity theory, is not obviously any less vague or synthetic than the notion of satisfactory self-realization. Further, it has been argued above that satisfactory self-realization is preferable to self-esteem when discussing the antecedents and consequences of group action, as individuals can engage in group action regardless of both their prior levels of self-esteem and their anticipations of the effects that such group action will have upon their self-esteem. Thirdly, it does not seem that the notion of satisfactory self-realization is necessarily hopelessly vague and synthetic. The term simply attempts to capture the intuitive truth that people are motivated by many concerns which can differ across both individuals and situations. It therefore allows for (and indeed "insists on") the possibility for

empirical research to *identify* what motivates certain people in certain situations. This seems preferable to a blanket assumption that people are always (and only?) motivated by self-esteem needs.

A second difference between the two accounts concerns the group-level factor which is postulated to affect and motivate individuals to group action. This is the "drive" to positive in-group distinctiveness (and thus satisfactory social identity and self-esteem) in the case of social identity theory, and the "drive" to satisfactory in-group realization (and thus satisfactory social identity and self-realization) in the case of goal identification theory. Clearly satisfactory in-group realization is the more inclusive of the two concepts, just as satisfactory self-realization is more inclusive than self-esteem. And, as with self-esteem, adherents of social identity theory are beginning to question whether in-group distinctiveness is sufficient for the task social identity theory requires of it (e.g. Brown et al., 1992). Goal identification theory again offers an alternative, by postulating that positive in-group distinctiveness can be a component of satisfactory in-group realization, but it need not necessarily be.

A powerful objection to this line of argument might be to cite the vast amount of research which supports social identity theory's notion that people do tend to display in-group preference, and that people do seem to "prefer" and to be motivated to achieve positive in-group distinctiveness and to avoid negative in-group distinctiveness. Such evidence might seem to suggest that social identity theory is correct to stress the importance of in-group distinctiveness, both as a motivation for group behaviour and as a determinant of group members' psychological well-being. The response to such an objection is to agree that the notion of in-group distinctiveness is indeed important in these ways, but to deny that it is necessarily involved in motivating group action or

in affecting in-group members' psychological well-being. Positive in-group distinctiveness and, perhaps particularly, avoidance of negative in-group distinctiveness are important because such states are often either group goals in and of themselves, or taken by group members to be indicative of progress toward or away from other goals in-groups have. However, as with the debate above concerning the relative merits of self-esteem and satisfactory self-realization, goal identification theory does not assume that group members are always (and only?) motivated and psychologically affected by in-group distinctiveness. Rather, it suggests that it is a matter of empirical enquiry to determine the circumstances in which such concerns will (and will not) motivate group members and affect their psychological well-being.

A third difference between the two accounts concerns the relative centrality of group goals/norms within each theory. Goal identification theory might be held to fall foul of Occam's razor in its proliferation of (unnecessary) theoretical concepts. At first glance, social identity theory appears to rely on very few constructs to explain and predict the antecedents and consequences of (at least) all intergroup situations; namely self-esteem, social categorization, social identity, intergroup social comparison, and the unique group goal of positive in-group distinctiveness. Goal identification theory, on the other hand, as we have seen, relies on a considerably more complex set of constructs, the most prominent member of which is group goals (and norms). Clearly, if we can account for all instances of intergroup behaviour without such additional concepts, we should.

However, there are several compelling reasons to argue that group goals are a necessary, or at least desirable, component of any theoretical account of group behaviour. These have been relatively well rehearsed in this chapter and there is little to be gained by excessive repetition. However, certain points may benefit from reiterated. First, social identity

theory does, in fact, rely on many more concepts than the list given just above. In particular, its notions of "valued comparison dimensions" and of conflicts of values at least implicitly accept that group goals do need to be considered when attempting to predict and explain antecedents and consequences of group behaviour.⁸ Second, unless superordinate-level (i.e. group) phenomena are explicitly considered, there is no way to distinguish between uniformities of social behaviour which are "genuinely" the result of group behaviour and those which are "simply" the result of individuals responding in similar ways to similarly perceived stimuli. Third, without considering group-level goals and norms, social identity theory offers a rather unsatisfactory account of all instances of non-discrimination in multi-group situations (both by groups and by individual group members). Fourth, without acknowledging that groups can have various superordinate-level goals and norms, and can assign different priorities in different situations to such goals and norms, social identity theory is unable to specify when social competition (as opposed to realistic competition) can be expected (let alone when non-discriminatory group behaviour may occur). Finally, a neglect of group goals and norms deprives social identity theory (and still more self-categorization theory) of a genuinely temporal dimension. That is, social identity theory seems committed to saying that in each and every situation all that will motivate group members is the in-group's position relative to one or more particular comparison groups. This is intuitively (at least) unappealing, as it seems a relatively common phenomenon for groups to develop norms of group and intergroup behaviour which apply more or less regardless of the particular situation groups or group members happen to be in at any given time. It seems doubtful, for example, that committed members

⁸ See also the importance given to group goals and norms within referent informational influence and self-categorization theory (e.g. Turner et al., 1987, esp. Reicher's chapter).

of the British National Party would reflect on the relative position of blacks and whites in society when deciding how to behave toward or evaluate a black person who had just bumped into them...

Goal identification theory, in contrast to social identity theory, explicitly states that across groups and across situations there will be different group-level goals and norms, and different priorities among such group goals and norms, and that it is a matter of empirical investigation to determine what these might be in any given situation.

A fourth major difference between the two theories on offer, and the last one to be considered here, concerns their respective ranges of application. Social identity theory, at least in the form Tajfel presented, is an explicitly intergroup theory. Thus, it has little to say about individual, interindividual, within-group, group (as opposed to intergroup),⁹ or "across-level" phenomena (cf. Horton, 1993). Goal identification theory potentially has a much more ambitious application. Its notions of satisfactory self-realization and identification with the goals of others means that it is possible to apply the theory to all areas of human behaviour. In particular, as well as concerning itself with individuals identifying with groups, and therefore being guided in their group and intergroup behaviour, the theory allows for considerations of situations in which individuals identify with and are guided by the goals of other individuals and, possibly most interestingly, of "ideologies" and "causes" (e.g. democracy, environmentalism, etc.).

⁹ Actually, for Tajfel there appears to be no possibility of group behaviour which is not at the same time intergroup behaviour (see Tajfel, 1981: 40, 256).

POTENTIAL DIRECTIONS FOR FUTURE RESEARCH

As noted above, in attempting to investigate Tajfel's social identity theory more or less "as a whole", it has not been possible to investigate particular aspects of that theory in any great depth, theoretically or empirically. Clearly, then, the first suggestion for future research has to be that the particular concerns addressed in this thesis are returned to with greater attention to theoretical complexities and with considerably improved methodologies.

The least interesting (but still worthwhile) thing to suggest is that each of the studies presented in this thesis are re-run with improved samples, manipulations, and measures, and perhaps also with different aspects of those studies addressed separately (e.g., Study 6, reported in Chapter 7, could be broken down into at least 3 separate investigations).

Another potentially fruitful direction would be to take some of the issues raised in this thesis and see how they fare in comparison with alternative derivations from or developments of Tajfel's theory (e.g. Turner et al.'s 1987 self-categorization theory), or indeed with "competing" theories (e.g. Rabbie et al.'s 1989 behavioural interaction model).

A third possible direction for future research is the most radical. This would be to consider "goal identification theory" not as a derivative of or a modification to social identity theory as such, but rather as a separate theory in its own right. Clearly, bearing in mind the paucity of, and the limitations associated with, the empirical research "supporting" that theory at the present time, such a programme of research would require a certain leap of faith. However, the growing research suggesting that social identity theory is too simplistic in certain key areas, coupled with the promise of a genuinely social psychological theory applicable to a much wider range of phenomena than simply one form of

intergroup discrimination, suggests that such an endeavour would not be entirely without merit. If pursued, it is envisaged that goal identification theory would best be investigated in combination with related theories. Potentially fruitful links between goal identification theory and Tajfel's social identity theory have been relatively extensively explored. Other possibilities include Turner's account of the development and adoption of in-group norms in his referent informational influence model (Turner, 1981) and self-categorization theory (Turner, et al., 1987) (see Chapter 4), and Moscovici's account of the same thing in his social representations theory (Doise et al., 1993; Moscovici, 1984, 1988). Also potentially useful are theories which postulate, and offer methodologies for measuring, actual self/possible self (and, by extension actual in-group/possible in-group) discrepancies (e.g. Higgins, 1987, 1989; Markus & Nurius, 1986; Wicklund & Gollwitzer, 1982).