TOP MANAGEMENT TEAM DECISION-MAKING:

A MULTI-LEVEL APPROACH TO UNDERSTANDING DEMOGRAPHIC AND COGNITIVE VARIATION, TEAM PROCESSES AND DECISION BELIEF

Tracey Anne Swift

Thesis submitted to the University of Sheffield in part fulfilment of the requirements of the degree of Doctor of Philosophy

Sheffield University Management School

December 2005

DEDICATION

With gratitude
I dedicate this thesis
to the three people
who have made me whole.

They are:

my mother,
Lorna June Heys,
who gave me life,
and blessed me with self belief and Faith;

my father,
Graham Martin Heys,
who taught me to read,
and thus endowed me with eternal Hope;

and

my husband,
Alleyne John Lloyd,
who championed me,
and showered me with enduring Love.

SUMMARY

Based within the 'upper echelons' tradition, the starting premise for this thesis is that demographic attributes such as age, functional background, educational attainment, gender, and tenure, influence the decisions made by top management teams (TMTs) (Pfeffer, 1983; Hambrick & Mason, 1984). Unlike most studies, which use public archival data, artificial teams, or retrospective interviews with a couple of selected senior executives, this research design (which is unprecedented in the TMT literature), investigated the decision making processes, in real time, of 23 authentic and fully functioning TMTs in the UK manufacturing sector using a state-of-the-art business simulation.

From a concentrated literature review which focused exclusively on TMTs, and disentangled the constructs of *dissimilarity* (individual level differences) and *diversity* (team level differences), a series of propositions were established. These hypothesized that demographic variation would lead to cognitive variation, that both these types of variation would influence team processes, which in turn would affect decision belief.

Despite the meticulous precision with which the constructs were measured in this research, and even with the application of sophisticated multi-level modeling techniques, only limited and sporadic support was observed for these predictions. Although there were slightly more findings than one would expect by chance alone (27 from a possible 177), these tended to be isolated and formed no clear pattern. Moreover, when one went beyond tests of simple statistical significance and reviewed effect sizes, all 27 results were tiny. The conclusion of this research is that demographic attributes are not nearly as influential in real TMTs as 'upper echelons' theory (Hambrick & Mason, 1984) supposes.

It is argued that the lack of convincing results is due to over-riding and inherent social factors in authentic TMTs, so that individual demographic differences cease to be novel or important during strategic decision-making discussions. The practical, theoretical and methodological implications of retaining the global null hypothesis are discussed in the final chapters.

PUBLICATIONS ARISING

BOOKS AND MONOGRAPHS

Swift, T. A. (1997). Top Management Teams: A Development Workshop. Trainer's Manual. Sheffield, UK: ESRC Centre for Organisation and Innovation.

Swift, T. A., & West, M .A. (1998). Reflexivity and Group Processes, Sheffield, UK: ESRC Centre for Organisation and Innovation. ISBN: 0 9533504 1X

JOURNAL ARTICLES

Swift, T. A. (1998). Creative decision-making in top management teams, Journal of High Performance Teams, **3,** (3), pp 34-39.

CONFERENCE PAPERS

Swift, T. A. (1998). Reflexivity and Naturalistic Decision-Making: Implications for Top Management Teams in Real-World Organizations Paper presented to the British Academy of Management, Nottingham, UK, September.

Swift, T. A. (1999). Team Reflexivity and Effective Decision-Making: Implications for Naturalistic Decision-Making Research, Paper presented to the 6th European Congress of Psychology, Rome, Italy, July.

Swift, T. A., & West, M. A. (1998). The Development of Top Management Teams, Video Presentation to IAAP, San Francisco, USA, August.

WORKING PAPERS

Stride, C. B., Swift, T. A., & Wall, T. D. (2000) Team Diversity in Equal Measures. Working Paper, Institute of Work Psychology, University of Sheffield, Sheffield, UK.

Swift, T. A. (1997). Coding Reflexivity in Group Discussions. Working Paper, Institute of Work Psychology, University of Sheffield, Sheffield, UK.

Swift, T. A., & Davidson, O. G. (1997). Development Workshops: Methods of Research and Data Collection. Working Paper, Institute of Work Psychology, University of Sheffield, Sheffield, UK.

REPORTS TO COMPANIES

Swift, T. A., & Davidson, O. G. (1997). Effective Decision-Making in Top Management Teams. An 8 page report for each of the 23 TMTs that participated, based on their responses to the Top Management Team Questionnaire.

ACKNOWLEDGEMENTS

"In the end, the world breaks everyone. And some get strong at the broken places".

George Bernard Shaw

My PhD journey began one glorious summer day in Queensland in 1996. Browsing in the University bookshop, I came across a monograph called Experimental and Quasi-Experimental Designs for Research by Campbell and Stanley. The fact that it had originally been published the year I was born and yet despite its vintage, was still current and on sale, piqued my interest as to its merit. Coming from a strong ethnographic research background as a management consultant, I was, at that time, quite disillusioned with the rigour and generalizability of much of the research within my profession, particularly social impact assessments at community level. I bought the monograph for the grand sum of $2.95, sat under the shade of a spectacular Poinciana tree on the banks of the Brisbane River, and read it from cover to cover. I was inspired. My focused, high achieving lifestyle became wildly peripatetic after that auspicious day nine years ago. The single constant was my hobby of doing a PhD, something I would take up every now and then, during brief periods of stillness. I have finally settled into comfortable domesticity and taken up winemaking, so the cherished, yellowing, piles of references, and tomes about multi-level modelling, must now make way for demijohns, airlocks and campden tablets.

I owe enormous thanks to a host of people without whom, this thesis would never have been written; to a succession of supervisors, but particularly Cathy Cassell and Toby Wall, both of whom stepped in some years into the study; to Cathy for offering guidance whilst trying to make sense of all that had gone before, and for so graciously accommodating necessary delays due to my career advancement, relocations and my mother's contracting terminal cancer; to Toby for his tenacity and integrity, but especially for continuously pushing me to probe deeper and teaching me the value of persistent inquiry, the document is considerably enhanced by his attention to detail; to Malcolm Patterson for taking such a huge gamble and allowing me access to the CEP group of companies; to Chris Stride for developing a bounded dispersion index so that I did not have to misuse the coefficient of variation like everyone else does; to Jon Rasbash and Andy Jones for advising me so freely on the multi-level modelling in the study and guiding me through the application of MLWin; to Valerie Sessa, at the Center for Creative Leadership, Greensboro, North Carolina for allowing me to use the PEAK Selection Simulation©; to Lee & Barbara Beach for accommodating me in their home in Tucson, Arizona and for the interesting discussions about Image Theory and reflexivity; to Bob Pritchard and Sandy Simpson for their friendship and love both here in the UK and under the Texan sun; to Dave and Frances Owen for their constancy, and particularly for never allowing me to take myself too seriously; to Oliver Davidson for sharing the burden for, and the excitement of, delivering the first 12 workshops; to Carsten De Dreu for partnering me at a workshop in Sheffield; to Patrick Flood for a memorable, army checkpoint experience somewhere deep in the border country at midnight, en route to a workshop near Belfast; to Vicki Butcher and Craig Hartley for poring for hours over the videotapes and rating the behaviour of every team; and humbly, to the 130 senior executives in 23 top management teams located all over the UK who gave so generously of their time to participate in this study.

Thanks also to two very special people: to my husband Alleyne, a most wonderful human being; for gluing me back together at a time when I was broken down; for financial support to finally finish; and for never tiring of discussing my PhD data, ideas and dilemmas; and to my Dad for proof reading the entire document and getting me to the first draft stage, a new common purpose as we emerged physically depleted and emotionally drained from a most traumatic experience.

Ultimately, I am indebted to my mother, who enthusiastically championed my desire to embark on this journey, and whose faith in my success was absolute.

Tracey A. Swift
Cambridge
December 2005

CONTENTS

CHAPTER 3:
THEORY & RESEARCH: COGNITION, TMT PROCESSES & DECISION BELIEF

CHAPTER 4:
OBJECTIVES OF THE STUDY IN DETAIL

CHAPTER 5:
RESEARCH DESIGN AND METHODOLOGY

CHAPTER 6:
MEASURES AND DESCRIPTIVE STATISTICS

CHAPTER 7: FINDINGS: HYPOTHESES 1 AND 2

CHAPTER 8:
FINDINGS: HYPOTHESIS 3

CHAPTER 9:
FINDINGS: HYPOTHESES 4 - 6

CHAPTER 10:

FINDINGS: HYPOTHESES 7 - 9

CHAPTER 11:

DISCUSSION

CHAPTER 12:

CONCLUSIONS AND FUTURE DIRECTIONS

T A B L E S

F I G U R E S

CHAPTER 1

Overview of Thesis Content and Structure

1.0 Orientation

This thesis focuses on an important conundrum in the literature on top management teams (TMTs). On the one hand, there is an almost universal belief that demographic variation in TMTs (e.g. with regard to age, functional background, education, gender, and tenure) will lead to better decision-making, and in turn, better organisational performance. The underlying premise for this is that demographic differences manifest themselves in different perspectives brought to bear by participants on strategic decision-making issues.

On the other hand, empirical evidence for the above beliefs is less than convincing, largely because of limitations in research to date. Leaving aside that findings have been inconsistent, there are methodological and conceptual weaknesses. Three problems are paramount. The first is that there has been a lack of consistency and precision around the notion of TMT demographic variation. Some investigators under the banner of '*demography*' have only investigated the similarity or central tendency of a team on a particular attribute (e.g. the mean age of a TMT) and related this to aspects of organisational performance (see for example, Hermann & Datta, 2005). Others have focused on how individuals' *dissimilarity* from their team colleagues, measured for instance by an index of Euclidean distance, relates to their behaviour such as their propensity to leave the team (see for example, Wagner, Pfeffer & O'Reilly, 1984). Still others have investigated *diversity*, that is to say, the degree of demographic variation at the team level (e.g. variance in age within teams) and related that to outcomes such as market share and profitability (see for example, Hambrick et al., 1996). As will be made explicit in Chapter 2, demography, dissimilarity and diversity are different constructs, with different meanings, that require different measures. The tendency to confound the terms, and to treat them as a single construct, has not helped a clear picture to emerge from research findings. It is not surprising therefore, that the current literature has been characterised as "uninterpretable" (Priem, Dess & Lyon, 1999).

The second problem is that the mechanisms or processes through which demographic variation have their effects remain largely unexplored. Investigators have tended to *assume* that if say, functional diversity has a positive relationship with the outcome variable (e.g. innovation), that it is due to the effects of that diversity on the team's collective cognitive capabilities and team processes (see for example, Bantel & Jackson, 1989). Yet, almost invariably, they have not directly measured those mediating variables. Being able to demonstrate such a link would considerably strengthen the findings. More than that however, because such a link has not been established, the traditional input-process-output model that underpins TMT demographic research (even though it typically excludes process), does not have a solid foundation. That is to say, if the link between input and process is not established, the input-output relationship that presumes such an intervening link is present, is at best, flawed.

Furthermore, outcome measures in this literature are typically limited to aspects, often financial, of organisational performance that are published in annual reports. This practice continues despite the methodological issues concerning the influence of environmental considerations, and the inevitable lag between the TMT's devising of strategy and any observable financial performance effect (Murray, 1989). Research is needed that: (a) uses additional outcome measures to financial performance; (b) limits the lag between cause (TMT demographic variation) and effect (outcome); and (c) measures directly the link between demographic variation and process.

The third problem is that little research to date has involved the direct study of authentic top management teams in real time, and so is of unknown external validity (Shadish, Cook & Campbell, 2001). Rather, investigators have placed much reliance on the use of secondary archival data (see for example, Jackson et al., 1991) inference from other types of team (e.g. Kilduff et al., 2000), studies of individual senior managers (e.g. Kirchmeyer, 1995) or supervisor-subordinate dyads (Tsui & O'Reilly, 1989).

Thus, Pettigrew's (1992) observation over a decade ago remains relevant to today:

"The more damning indictment of the demography based top management team research is that no one has ever been anywhere near a top team in an organizational setting, either to directly observe a team in action, or to interview the members about the links between their characteristics and structure and processes of communication and decision making and their impact and performance" (pp. 175).

This thesis is responsive to all three of these limitations. The problem of conceptual precision is addressed by distinguishing clearly between the team level construct of similarity as measured by demography, and the two constructs focussing on demographic differences, dissimilarity at the individual level and diversity at the team level. Those distinctions are then used not only to interpret findings from the extant research literature, but also as the rationale for the use of multi-level modelling in the new empirical work. Examining the dissimilarity and diversity constructs together helps to make the differences between them clearer; and the use of multi-level modelling is not only particularly appropriate, but also novel in this domain (Hodgkinson, 2001). The issue of mechanism is tackled by including cognitive and team process variables in the empirical study, to determine how these are associated with demographic variation and mediate relationships with outcomes. Cognitive variation is measured at the individual (dissimilarity) and team (diversity) levels. The outcome measure is decision belief. The issue of external validity is addressed by using a sample of intact, authentic top management teams in the UK undertaking a realistic decision-making task in real time (23 TMTs and 130 executives).

1.1 Content and Structure of the Thesis

In Chapter 2, two particularly influential theories 'upper echelons' and 'organisational demography' are reviewed along with the research that has followed in each tradition. This serves to anchor the subsequent discussion which illustrates the apparent confusion in the literature over demographic variation and the way in which it has selectively ignored cognitive variation and process variables. The small number of studies that have exclusively investigated top management teams is then reviewed. The purpose here is to distil what is known specifically about such teams, but also to derive propositions concerning top management team demographic and cognitive variation.

In Chapter 3, this process is taken a stage further by means of an examination of the small group research literature that concerns managerial cognition in strategic decision making and team processes, with particular relevance to TMTs. The relationships between variation (demographic and cognitive) and four team processes, frequency of team meetings (as a proxy for communication), procedural rationality, reflexivity and psychological safety, are explored. These processes are deemed to have special relevance to complex decision-making teams, and TMTs in particular. Reasons for these suggested relationships will be investigated and explored. Ways in which the current study proposes to extend existing knowledge on these aspects of top management team functioning is outlined.

In Chapter 4, the threads of the argument developed in the previous two chapters are drawn together in the form of a guiding model, against which the nine hypotheses are mapped. The model is specifically designed to act as a framework for investigation and analysis and illustrates the proposed relationships between demographic variation, cognitive variation, team processes and decision belief.

Chapter 5 presents the research design and methodology and describes the empirical study with 23 real top management teams. It explains how limitations noted in other studies were overcome using a simulated decision-making task and bespoke coding guides. Justification for the definition and selection of the sample is presented. Description of data collection and rationale for using existing team process questionnaires offered.

Chapter 6 outlines the measures and the descriptive statistics. Chapter 7 discusses and evaluates the results in relation to the first two hypotheses, concerning the relationship between individual demographic dissimilarity and cognitive dissimilarity and between team demographic diversity and cognitive diversity. Finally, this chapter summarises the contribution which this study makes to the discipline of psychology and organizational behaviour theory.

Chapter 8 begins by outlining the procedure for multi-level modelling and then goes on to apply this to hypothesis 3, which concerns the individual decisions made by demographically dissimilar individuals in demographically diverse teams.

Chapter 9 presents the analysis and findings for hypotheses 4 – 6 which concern the relationship between demographic variation and team processes. Hypothesis 6 transcends the individual and team level, so multi-level modelling is once again employed. There is less detail concerning the model in this chapter, as the procedure and lengthy explanation of the process is presented in Chapter 8.

Chapter 10 presents the findings for hypotheses 7 – 9 which are concerned with the relationships of demographic variation, cognitive variation and team processes respectively with decision belief.

Chapter 11 observes that there are slightly more statistically significant results than one could expect to occur by chance alone and debates whether to follow precedent and massage the isolated findings from the previous four chapters into a coherent story. However, as there are over 170 relationships tested in the study, correction techniques for the purpose of establishing overall study-wise significance are reviewed. As a result of applying Cohen's Standard, the global null hypothesis is retained. That is to say, this study finds that there is no systemic effect of demographic and cognitive variation on team processes or decision belief. However, the study has considerable merit owing to the methodological advancement made in directly investigating authentic TMTs, and the precision with which the constructs were measured and analysed. Implications for 'upper echelons' theory are discussed in relation to critiques raised in earlier chapter.

Chapter 12 sets out a future agenda for research comprised of five themes which emerged in Chapter 11. The Chapter concludes by encouraging future researchers to be bold in seeking to overcome access difficulties in order to get very close to TMTs.

CHAPTER 2

Theory and Research Concerning TMT Demographic Variation

2.0 Overview

This chapter covers the underlying premise on which decades of top management team (TMT) demographic research has been based. That is, a TMT that varies on demographic attributes will have increased cognitive capabilities, which in turn will lead to enhanced performance. The two traditional perspectives for studying relationships between TMT demographic variation and outcomes are explored, 'upper echelons' and 'organisational demography'. The first, 'upper echelons' posits a direct link between demographic variation and organizational performance. The second, 'organizational demography', recognises that team processes mediate this supposed relationship.

This chapter finds that the research literature is problematic, first, because many of the studies supposedly about this topic are not actually about TMTs. Indeed, as will become clear, in nearly three decades of research, no study has come face to face with whole, intact, TMTs. The second problem with the literature is that three discrete types of demographic variation are confounded. This chapter will disentangle the differences between the distinct constructs of demography, dissimilarity and diversity previously conflated and used interchangeably in the literature. Furthermore, it will demonstrate that each construct is based on distinctly different types of measure. Methodological limitations of TMT demographic variation research are then discussed. The chapter concludes by illustrating where the current research fits in relation to existing literature.

2.1 Introduction

The dictionary definition of demography is the 'science of population statistics'. Specific applications of this science have been made to organizations (Pfeffer, 1983; Wagner, Pfeffer & O'Reilly, 1984; Zenger & Lawrence, 1989; Wiersma & Bird, 1993), in relation to superiors (Tsui & O'Reilly, 1989; Tsui, Egan & O'Reilly, 1992), and to TMTs (OReilly & Flatt, 1989; Wiersma & Bantel, 1992; O'Reilly, Snyder & Boothe, 1993; Smith et al., 1994), so that 'organizational demography', and 'TMT demography' have common parlance in TMT research literature.

Top management teams are possibly the most influential groups in organizations. They are primarily responsible for directing and guiding a business in order to maximise shareholder wealth. Hence, they have attracted a great deal of academic interest as corporate boards and other stakeholders want to know what makes them effective or ineffective.

The 1980s witnessed a burgeoning interest in the role that demographic background characteristics play in CEOs effecting strategic change. Such research was founded on an underlying premise that older CEOs made more conservative decisions, and younger managers were more adventurous (Hambrick, & Mason, 1984). The next logical line of enquiry was to investigate the background characteristics of entire top management teams. This led to the advent of two theories, 'upper echelons' and 'organisational demography'.

2.2 Upper Echelons Theory

The most famous and arguably the simplest theory concerning demographic variation in TMTs and organisational performance was put forward by Hambrick & Mason (1984). Although it post-dates the theory espoused by Pfeffer (1983), its simplicity means that it easier to address first. Moreover, this is the only theory that applies exclusively to top management teams. Hambrick & Mason (1984) took a between team perspective, positing that the *"central tendencies of the entire top management team"* (pp 196) have an effect on the strategic choices they make on behalf of their organizations. In developing 'upper echelons' theory, the authors drew on a body of existing research which had exclusively investigated the background characteristics of CEOs in relation to organisational performance. Hambrick & Mason (1984) upped the ante to address entire TMTs. They posited that *"relatively straightforward demographic* *data on managers may be potent predictors of strategies and performance levels*" (Hambrick & Mason, 1984, pp 204).

Specifically, Hambrick & Mason (1984), in a purely theoretical piece, put forward 21 propositions, 18 of which concerned central tendencies on demographic factors including age, functional background, tenure, educational experience and socio-economic background that were expected to affect organisational performance. For example,

"*Firms with young managers will be more inclined to pursue risky strategies than will firms with older managers*" (pp 198);

"*The degree of output-function experience of top mangers will be positively associated with growth*" (pp 199);

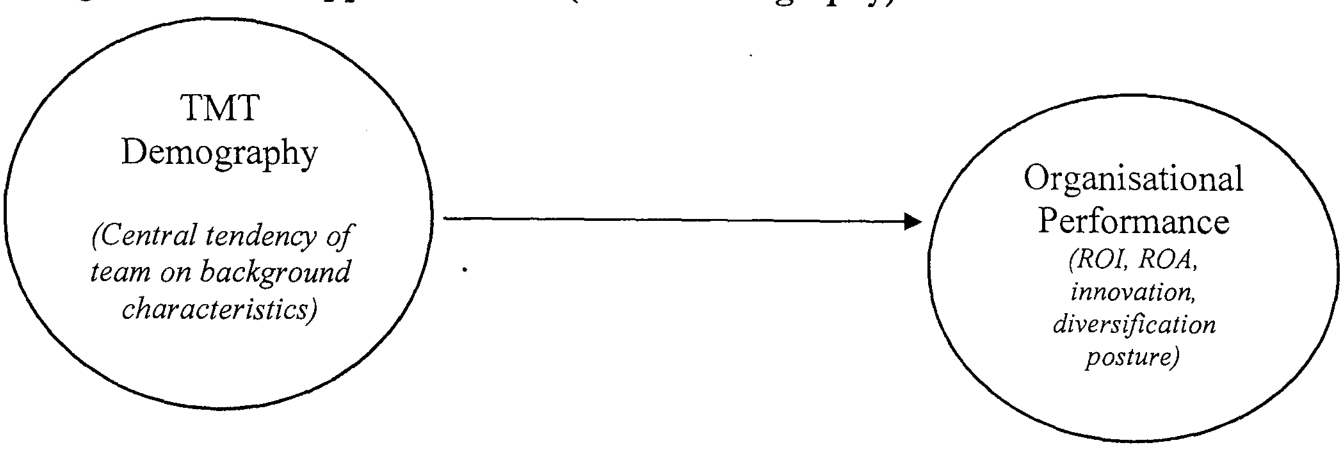
"*The amount but not the type of formal education of a management team will be positively associated with innovation*" (pp 200); and

"*Firms whose top managers come disproportionately from lower socioeconomic groups will tend to pursue strategies of acquisition and unrelated diversification*" (pp 201).

There was no mention by Hambrick & Mason (1984) of the team processes inherent in TMT decision-making, their focus was exclusively on similarity of TMT members and aspects of strategic choice or performance. Their interpretation of strategic choice was based upon earlier research concerning bounded rationality (Hodgkinson, 2003), a phrase which refers to decision-making when the consequences of a decision are unknown or uncertain. Hambrick & Mason (1984) accepted bounded rationality as a given, they were not concerned with the processual issues of decision-making. Furthermore, they were content to accept that the way in which individuals approached a decision was directly attributable to their demographic traits (Hambrick et al., 1996; Hodgkinson, 2003). Although their chief concern was about the central tendency of the TMT as a whole, they did touch on diversity within teams as evidenced from their statement that "*study of an entire team has the added advantage of allowing inquiry into dispersion characteristics such as homogeneity and balance*" (pp 197).

Research in the 'upper echelons' tradition assumes a sequential and linear relationship between TMT demography (the central tendency on background characteristics) and whatever performance outcomes are under investigation (see for example Michel & Hambrick, 1992; Finkelstein, 1992), to the exclusion of process. Figure 2.1 illustrates the direction between the two constituent parts. Various aspects of organisational performance have been studied, including propensity for high level strategic change (Wiersma & Bantel, 1992); diversification posture (Carpenter, 2000); levels of innovation (Bantel & Jackson, 1989); return on investment (ROI), sales growth and return on capital expenditure (ROCE) (Smith et al., 1994) amongst others.

Figure 2.1: The Upper Echelons (TMT Demography) Model



A second influential theory, which adds another component to the demography model will be discussed in the next section.

2.3 Organisational Demography

Organisational demography was first introduced by Pfeffer (1981; 1983), who recognised that without the collection of demographic statistics on an entire population, whether that population be a team or an organisation, there could be no investigation of similarity or differences. At first blush, 'organisational demography' and 'upper echelons' theories seem almost identical and they are quite often lumped together as a result. Like Hambrick & Mason (1984), Pfeffer (1983) asserted that organisational decision-makers have a cognitive base which is a function of their age, tenure, education and functional background. However, there are three specific differences between the two theories that are worthy of note.

The first is that Pfeffer's (1983) theory could be applied to any decision-making group, and even the organisation as whole, rather than TMTs per se. But, because CEOs and TMTs are the primary organisational decision-makers, Pfeffer's (1983) theory is particularly appropriate to, and widely held to be, pertinent to them.

The second difference between 'upper echelons' and 'organisational demography' theories is that Pfeffer (1983) urged that it was the distributional properties of team demographics, rather than the central tendency, that would be critical to an understanding of how demographic attributes affects organizational processes and outcomes. Despite his admonition to the contrary, there are still examples (e.g. Hermann & Datta, 2005) of empirical work following Hambrick & Mason (1984) which has solely investigated TMT demography using simple means to determine central tendency.

Pfeffer's (1983) concern was that by relying on the mean, research was primarily concerned with investigating the degree of overall similarity of team members and that a richer line of inquiry was about how demographic differences affected organisational performance. Pfeffer's understanding of this area was based in large part on his work concerning the impact of aging workforces on organisations (Pfeffer, 1981). Pfeffer's further work on this topic (Wagner, Pfeffer & O'Reilly, 1984), shows that it was dissimilarity, that is to say, individual differences within a group, in which he was particularly interested. He still referred to individual dissimilarity as 'organizational demography', which on the one hand is unfortunate, because it has led to confusion in the literature (a topic that will be returned to later in this chapter), but on the other, shows that demography is the starting point for understanding differences both at the individual and team levels. It is of interest to note on careful reading of the TMT literature, that those who study demography (central tendencies) tend to subscribe to 'upper echelons' theory, whereas those who study demographic differences, that is to say, individual level dissimilarity or team level diversity, follow Pfeffer. Indeed, as a result of Pfeffer's (1983) theory, a whole raft of TMT research has been about heterogeneity or diversity (see for example, Wiersma & Bird, 1993; Hambrick et al., 1996; Carpenter, 2000). These terms will be clarified in the next section, but suffice it to say at this point, Pfeffer's work was about demographic differences within and across teams, whereas Hambrick & Mason's concentrated on similarities within teams and differences across them.

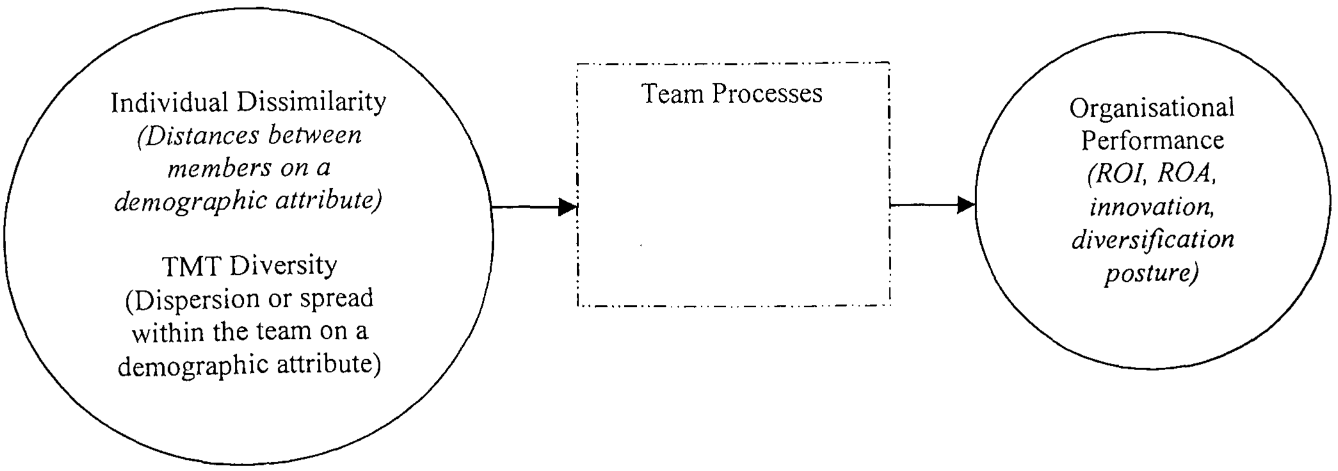
The third difference is that Pfeffer's (1983) theory suggested that demographic variation with regard to age, gender, education, tenure, race, etc would have effects on internal processes. Pfeffer based this part of his theory on small group research, which had long asserted that demographic differences result in a greater breadth of perspectives, but that the cognitive biases people held as a result demographic differences, were hard to accommodate in group decision-making.

Pfeffer (1983) argued that it is not necessary to study directly the intervening nature of team processes, because demographic variation is a good enough proxy. Hence it is extremely common in the TMT literature for performance outcomes to act as proxies for what is later inferred about team processes (Priem, Lyon & Dess, 1999; O'Reilly et al., 1993; Smith et al., 1994).

To illustrate the way in which process is inferred but not studied directly, consider the following two examples. Weirsma & Bantel (1992) argued that demographic diversity in top management teams results in more creative and innovative decision-making processes. Although links were suggested between corporate strategic change and how innovative the reported changes were (the variables studied), the actual decision-making or implementation processes used by the teams were not measured.

Similarly, Bantel & Jackson (1989) maintained that at the individual level, dissimilar persons are affected by conflict within the team, to the point where they leave *"presumably because members find the increased conflict and decreased communication to be stressful"* (Bantel & Jackson, 1989, pp 118). Propensity to leave the TMT (called turnover) is a second outcome measure typically studied in this literature. Thus, turnover acts as a proxy for understanding the degree to which demographic variation complicates social interaction processes within the top management team. As turnover is usually an individual action relative to the rest of the TMT, demographic variation in relation to turnover is also appropriately studied at the individual level, that is to say, individual demographic dissimilarity. Conflict over ideas, perceptions and judgements is argued to put pressure on the most dissimilar person who will voluntarily exit the team (Jackson et al., 1991).

Figure 2.2: The TMT Demographic Differences/Process Model



Renewed interests in 'upper echelons' and 'organisational demography' traditions has resulted in a recent rash of further prescriptions (Edmondson, Roberto & Watkins, 2003; Dewett, 2004; Carson, Mosley & Boyar, 2004). Typically, these lament the omission of intervening process models (Edmondson et al., 2003), whilst still concentrating on demographic variation and performance (Carson et al., 2004). One places emphasis on diversity rather than characteristics (Carson et al., 2004), whilst two focus on the process of decision-making (Edmondson et al., 2003; Dewett, 2004).

2.4 Key Distinction in Terms

Most traditional reviews of the demographic variation literature, usually under the slippery rubric of diversity, catalogue studies according to demographic attribute (see for example, Milliken & Martins, 1996; Williams & O'Reilly, 1998). Yet there is a profound confusion over what constitutes diversity (Guzzo, 1996). Moreover, demography and diversity (as they are the two most studied concepts) are often used synonymously and as a result, the constructs have been conflated (Tsui et al., 1992). As mentioned earlier, this is also true concerning dissimilarity which refers to the distances between team members on a particular trait (Jackson et al., 1991). In this section the differences between demography, dissimilarity and diversity are explained.

2.4.1 Demography

Demography simply refers to the central tendency of a TMT on a particular trait based on simple descriptive statistics such as mean, median or mode. Research using this construct (using the illustrative trait of age), addresses the question of whether it is predominantly older or younger teams that tend to perform better. It assumes that such a central tendency is characteristic of the team, and ignores what variation there may be.

The next two constructs in the literature are concerned with demographic differences, the first at the individual level of analysis, and the second at the team level.

2.4.2 Dissimilarity

Dissimilarity is about how different an individual is compared to his or her peers. This individual level construct is measured using Euclidean distances. Using the same illustrative trait, it takes the individual's age as the absolute starting point and computes the distances between his and every other person's age on the team. Dissimilarity research may conclude that if one person is significantly dissimilar in age to the rest of the team, then there will be an effect on the turnover of the individual in relation to the rest of the team. It is not necessarily about who is the oldest or the youngest on the team, but who is the most distant, that is to say, most dissimilar, from colleagues. This type of research (as the following review will demonstrate), is very rarely conducted in TMTs. Yet, the insights it generates are extremely precise. For example, Jackson et al. (1991), found that educationally dissimilar executives are the most likely individuals to leave a TMT.

2.4.3 Diversity

Diversity is a team level phenomenon, and is measured using indices of dispersion. Some of the confusion in the literature appears to arise from a perception that diversity is merely the aggregate of dissimilar individuals on the team. It is important to realise that this is not the case. Diversity is concerned with the amount of heterogeneity at the team level. That is to say (using the same illustrative trait), the spread of different ages represented across the team. Diversity research assumes that it is not the direction of the average age that is important. Nor is it the degree to which a single individual is different, but the degree of variation within the team (i.e. the proportion of difference across the individuals' ages) that leads to the team generating different perspectives and thus making more or less creative decisions. There are at least five dispersion indices used in the literature (Stride, Swift & Wall, 2000). Those for categorical variables (such as functional background) work on the proportion of categories represented. In most cases, they result in a score between 0 and 1, where 0 means that everyone on the team is in the same category (total homogeneity) and a high score means greater representation across the categories, or diversity. For continuous variables (such as age or tenure), a coefficient of variation (the standard deviation divided by either the mean or median) statistic is used to determine dispersion. Again, a high score means more dispersion and low score more similarity. The findings about diversity can only apply to teams (not individuals) and cannot reflect the direction of the trait in terms of older or younger (i.e. central tendency or demography). For example, the finding that educationally diverse TMTs achieve greater ROI for their companies than educationally homogenous teams (Smith et al., 1994) is not about more highly or less well educated teams, but about the variety within the team concerning time spent in education.

A lot of confusion exists concerning demographic variation in TMTs which could be avoided by researchers being clear about which of the three constructs they are studying. As has been shown in this section, each construct relies on distinct measures. The next section illustrates some of the problems of conflating the terms and trying to meld incompatible measures.

2.4.4 Problems of Conflating Demography, Dissimilarity and Diversity

Studies within the broad spectrum of TMT demographic variation research often investigate demography, diversity and/or dissimilarity constructs at the same time. The problem is that they then use the terms interchangeably, or make assertions about dissimilar individuals in teams when they have only investigated the team level phenomenon of diversity. These then get repeated in new research, only exacerbating the confusion further. Examples of conflation abound in the literature. To illustrate, using two cases in point, O'Reilly et al. (1993) measure tenure diversity (heterogeneity at the team level) using the coefficient of variation (Allison, 1978), but call their measure demography (central tendency); whilst Wagner et al. (1984) measure individuals' tenure dissimilarity using Euclidean distances, yet also refer to their measure as demography. As described above, the constructs are different and rely upon discrete types of measurement, but only a very few studies reflect this precision (see for example, Wiersma & Bird, 1993).

Conflating the constructs exacerbates confusion over the findings, leading to claims that the results of this body of knowledge are 'largely uninterpretable' (Priem et al., 1999) and 'noisy and unreliable' (West & Schwenk, 1997).

The second major problem concerns errors of measurement, which, with deeper understanding and greater precision concerning the particular construct being studied, can be resolved. Two of the worst examples are those reported by Boone et al. (2004) and West & Schwenk (1997).

Boone et al. (2004) correctly apply a Euclidean distance measure to the continuous variable age, in order to arrive at a dissimilarity score for each individual. Unusually, their variable for tenure is categorical, as is their measure for education. As there is no measure of dissimilarity (individual distances) for non-ordered categorical variables, they invert Blau's (1977) index of *team* level dispersion (diversity), which is based upon the number of categories represented within a team, and improperly claim that it is a measure of "*an individual's similarity to the rest of her or his team*" (pp 640 italics added). Subtracting the value from a constant of 1 (the highest level of diversity on Blau's index) simply changes the direction to a similarity index, but it is still at the team level (Hambrick et al., 1996). This means that every individual takes the same value on the index for their own team. In the case in point there were 5 teams, so Blau's (1977) index will only generate 5 scores for education and 5 scores for tenure. Boone et al. (2004) refer to their team level similarity index of homogeneity in their analysis as an individual distance measure (i.e. dissimilarity). To make matters worse, they then went on to compute a global measure of team level diversity in which they summed the Euclidean distances for the individuals' ages (n = 53) together with the standardized team level dispersion values (n = 5), applied another dispersion index (Herfindal-Hirschman) and called it "team heterogeneity". The problem is that the findings which were reported simply cannot be trusted because the measures they used belie the constructs they purport to have investigated.

The second example is problematic for slightly different reasons and also involved Blau's (1997) index. This time the authors, West & Schwenk (1997), computed Blau's (1977) index for 12 categorical demographic variables (which are unidentified), summed them, then subtracted from a positive constant (which is unstated). At least they realised (unlike the previously cited authors) that they were computing a team level similarity index, which was actually desirable, as they were investigating demographic homogeneity and wanted higher scores to reflect similarity across the team. The problem in this case comes from not identifying the demographic variables used. Undoubtedly West & Schwenk (1997) would have had to change some of the regular continuous variables such as age and tenure into categorical responses in order to use Blau's (1977) index. They would have had to do this because dispersion values based on interval data cannot be added to dispersion values derived from non-ordered categorical responses. Unfortunately, there is no such thing as a universal measure of diversity, although it remains the holy grail of

TMT demographic variation research. Indeed, it is well documented in this literature that different types of demographic variation (e.g. education and tenure diversity) have non-uniform and unexpected effects on outcome variables (see for example, Knight et al., 1997) with seemingly contradictory results being reported for the same dependent variables within the same study. For this reason alone, it is inadvisable to add all the types of diversity together. The study which came closest to achieving a global measure was Murray (1989), who like West & Schwenk (1997), wanted to prove that several forms of diversity constituted a single construct. Murray (1989), however, conducted a very thorough analysis using multiple diversity indices and principal components analyses. The best fit was two factors, in that age and tenure loaded together (temporal diversity) and education and functional background loaded together (occupational diversity) which he found had opposing effects on his dependent variables. Due to the inappropriate reduction of data into categories, and the lumping of all the diversity variables together, it is perhaps hardly surprising that West & Schwenk (1997) ended up with "a report of resounding non-findings".

None of the findings arising from the two problematic studies just critiqued is included in the review which follows, as the measurement is so fundamentally flawed so as to make the results and the authors' interpretations about demographic variation totally unusable.

This section began by asserting that there are three discrete constructs in the TMT demographic variation literature which must be understood in order to make sense of the field. In summary, demography refers to the central tendency of a team on a demographic attribute it is measured using means, medians or modes. Dissimilarity refers to the extent to which each individual is different to the rest of the individuals on the team and is measured using Euclidean distances. Diversity refers to the degree to which there is variation within the team on demographic attributes. It is measured using team level dispersion indices such as the coefficient of variation for continuous variables (Allison, 1978) or Blau's (1977) proportional representation index for categorical variables. The next section gets into the specifics of what is known about TMT demographic variation whilst paying heed to the different constructs.

2.5 Research into demographic variation in top management teams

A broad church of material exists in the 'upper echelons' and 'organisational demography' traditions. In order to sift this vast literature, and to try to make sense of the uneven findings in relation to demographic variation in TMTs, a two step classification method is adopted.

First, only sources that use real TMTs as the unit of analysis are reviewed. This criterion serves to exclude research that deals with superior-subordinate relationships, chief executive officers to the exclusion of other members of a top management team, and with synthetic teams made up of senior executives for the purpose of role play (see for example, Kilduff et al., 1997). By way of contrast, it is interesting to note that of the 38 studies that could be broadly classed as 'upper echelons' research, reviewed by Milliken & Martin (1996), only eight actually were concerned specifically with TMTs.

The second classification step used here is to catalogue the findings as to whether the independent variables are demography, dissimilarity or diversity. It is of interest to note that not all studies include a range of demographic attributes (e.g. Finkelstein, 1992 examines only functional background), and some measure attributes such as age in one or more ways (see for example, Jackson et al., 1991, who measured dissimilarity and diversity).

Only 24 studies met the criteria of dealing with demographic variation in real TMTs, and a summary of all demographic attributes studied, at which level, the outcome variables, the control variables and a summary of results is presented in Table 2.1. Four studies were dropped from the review which follows, two owing to fundamental flaws in measurement as described above. Another used aspects of demographic diversity as control variables in studying TMT size and CEO dominance (Haleblian & Finkelstein, 1993), hence relationships concerning diversity were not reported. A further study (Clark et al., 1997) that developed fascinating descriptions of TMTs such as "shaky alliance", "headless group" and "autocracy", is also included in Table 2.1 for completeness as the component variables included demographic diversity. However, as there were no results in terms of dependent and independent variables, it is not included in the review. This same study has been reported elsewhere (Smith et al., 1994; and Knight et al., 1997) and is included appropriately in both the tables and the review. This means that there are 20 studies which form the core literature on TMT demographic variation.

Table 2.1 Variables Studied in TMT Demographic Variation Research

AUTHOR/DATE	DEMOGRAPHY					DISSIMILARITY					DIVERSITY					OUTCOME		CONTROLS	RESULTS
	AGE	PROF	EDU	SEX	TEN	AGE	PROF	EDU	SEX	TEN	AGE	PROF	EDU	SEX	TEN	T/O	PERF		
Bantel & Jackson, 1989	-	-	X	-	-	-	-	-	-	-	X	X	X	-	X	-	X	Size (Firm & TMT), Location	Education Demog +ve relationship with tot and technical innovation, Education Diversity no relationship. Function Diversity +ve with total and admi innovation.
Boone et al., 2004	-	-	-	-	-	X[a]	-	-	-	-	-	-	X[a]	-	X[a]	-	X	Major consolidation events, Firm Size, Team Diversity	Used an inverted version of Blau's index, which they claim (incorrectly) is a dissimilarity measure! Totally inappropriate and unnecessary when range of distance measures exist, i.e. Binary Euclidean Distance.
Carpenter, 2000	-	-	-	-	X	-	-	-	-	-	-	X	X	-	X	-	X	Firm Size, Industry, Average Tenure, Team Size, International Work Experience, Nationality	Functional and Tenure Diversity positive predictors of low level internationalization, negative predictors of high level internationalization. Educational Diversity positive predictor of low and high level internationalization. Short tenure TMT (demography) predictor of internationalization.
Clark et al.,1997	-	-	-	-	-	-	-	-	-	-	-	X	X	-	X	-	X	Not stated	8 Clusters or configurations of TMT devised. High, low and average levels of diversity important in defining clusters. Results are unclear.
Finkelstein & Hambrick, 1990	-	-	-	-	X	-	-	-	-	-	-	-	-	-	-	-	X	TMT Size	Short tenure teams pursue novel strategies that deviate widely from industry norms.
Finkelstein, 1992	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-	X	Power of individuals in TMT	Proportion of TMT members with finance background is marginally +ve predictor of diversification posture. (Used bespoke measure of proportion of dissimilarity for individuals).
Glick et al., 1993	-	-	-	-	-	-	-	-	-	-	X	X	-	-	X	-	X	Not stated	Functional Background Diversity +ve predictor of cognitive diversity about efficacy of spending money on advertising and rich communication
Haleblian & Finkelstein, 1993	-	-	-	-	-	-	-	-	-	-	-	X[b]	-	-	X[b]	-	X	Diversity vars, borrowing capacity, efficiency, environment, firm size, strategic unrelatedness.	Large teams with lessdominant CEOs are profitable in turbulent environments (computer industry) than stable environments (natural gas distribution)

Table 2.1 Continued…/

AUTHOR/DATE	DEMOGRAPHY					DISSIMILARITY					DIVERSITY					OUTCOME		CONTROLS	RESULTS
	AGE	PROF	EDU	SEX	TEN	AGE	PROF	EDU	SEX	TEN	AGE	PROF	EDU	SEX	TEN	T/O	PERF		
Hermann & Datta, 2005	X	X	X	-	X	-	-	-	-	-	-	-	-	-	-	-	X	Firm Size, ROA, R&D intensity	Higher levels of education positively associated with international diversification. Longer tenure and older age negatively associated with international diversification.
Hambrick et al. 1996	-	-	X	-	X	-	-	-	-	-	-	X	X	-	X		X	Firm Size, TMT Size, Average Educational Level.	All Diversity +ve predictors market share and profitability. Demography not associated.
Jackson et al. 1991		-	-	-	X	X	-	X	-	X	X	-	X	-	X	X	-	Age (proxy for retirement), Firm Size TMT Size	All Diversity +ve predictor turnover. Education Dissimilarity +ve predictor of turnover.
Knight et al. 1997	-	-	-	-	-	-	-	-	-	-	X	X	X	-	X	-	X	Not Stated	Functional, educational diversity negative predictors of consensus, tenure diversity positive predictor of consensus. Functional diversity +ve predictor of interpersonal conflict, age diversity –ve predictor of agreement seeking.
Krishnan et al. 1997	-	-	-	-	-	-	-	-	-	-	-	X	-	-	-	X	X	Firm Size, Prior Firm Performance, Industry Profitability	Functional homogeneity +ve with post-acquisition performance. Functional homogeneity is a –ve predictor of turnover.
Michel & Hambrick, 1992	-	X	-	-	X		-	-	-		-	X[c]	-	-	X[c]	-	X	Firm Size, Firm Age.	Functional homogeneity (legal and finance) +ve predictor of high interdependence. Low tenure homogeneity +ve predictor of high interdependence. Functional homogeneity strong –ve predictor of profitability & diversification strategy in high interdependent firms. Average tenure +ve predictor of strategic change.
Murray, 1989	-	-	-	-	-	-	-	-	-	-	X	X	X	-	X	-	X	Not stated.	Factor Analysis on diversity – age & tenure together (temporal diversity), function and education (occupational diversity). Temporal diversity –ve predictor of change in market share. Occupational diversity negative predictor of long term performance.
Norburn & Birley, 1988	X	-	X	X	X	-	-	-	-	-	-	-	-	-	-	-	X	Not stated.	Mixed results, but study across 5 industries with firms in various stages of growth. Demography strongest predictor of growth when measured intra-industry.
O'Reilly et al. 1993	-	-	-	-	-	-	-	-	-	-	-	-	-	-	X	X	X	Firm Age, Firm Size, TMT Size.	Diversity strong –ve predictor of team dynamics. Diversity also +ve predictor of turnover.
Simons, 1995	-	-	-	-	-	-	-	-	-	-	X	X	X	-	X	-	X	TMT Size, Firm Age, Firm Size, TMT Tenure	Functional Background Diversity interaction with debate predicted increased profitability. Interaction of debate and educational diversity positively associated with decision comprehensiveness

Table 2.1 Continued.../

AUTHOR/DATE	DEMOGRAPHY					DISSIMILARITY					DIVERSITY					OUTCOME		CONTROLS	RESULTS
	AGE	PROF	EDU	SEX	TEN	AGE	PROF	EDU	SEX	TEN	AGE	PROF	EDU	SEX	TEN	T/O	PERF		
Smith et al., 1994	-	-	-	-	X	-	-	-	-	-	-	X	X	-	X	-	X	ROI, Firm Size, Industry Growth Rate,	Educational Diversity +ve predictor of performance. Tenure diversity (experience) – ve predictor of social integration and informal communication. Team size –ve predictor of social integration. Tenure demography
Wagner et al., 1984	-	-	-	-	-	X[d]	-	-	-	X[d]	X	-	-	-	X	X	-	Firm Age	Tenure Diversity +ve predictor turnover Age dissimilarity predicts turnover.
West & Anderson, (1996)	-	-	-	-	X	-	-	-	-	-	-	-	-	-	-	-	X	None	Longer Tenure positively associated with effect of innovation on staff well-being.
West & Schwenk, 1996	-	-	-	-	-	-	-	-	-	-	-[e]	-	-	-	-	-	X	Not stated	'Resoundingly non-significant results'. Problem with this study is that it inappropriately summed Blau's index for 12 unidentified different demographic measures then subtracted from a +ve constant and called it 'demographic homogeneity'.
Wiersma & Bantel., 1992	X	-	X	-	X	-	-	-	-	-	X	-	X	-	X	-	X	Firm Size, TMT Size, Prior Firm Perf.	Low age, long team tenure, high educational level, high educational specialization, +ve predictor of corporate strategic change.
Wiersma & Bird, 1993	X	-	-	-	X	X	-	X	-	X	X		X	-	X	X	-	TMT Mean Age, Org. Perf., Industry	Educational Diversity +ve predictor of turnover. After controlling for TMT mean age, age diversity and educational diversity predict turnover.

[a] Boone et al. took an inverted team level diversity index and then claimed it was an individual dissimilarity index. The study is included here for completeness, but spurious results arising are not presented.

[b] Haleblian and Finkelstein included diversity variables as control variables. The study is tabled here for completeness

[c] Michel & Hambrick were interested in homogeneity, so inverted the coefficient of variation to measure homogeneity.

[d] Wagner et al. used a version of the Gini index to devise a measure of individual similarity to which they then applied the coefficient of variation to give a diversity measure.

[e] West & Schwenk used a global measure of inverted diversity to measure homogeneity, based on 12 unidentified demographic categories. There were no statistically significant effects.

The most frequently used concept is that of TMT demography, the central tendency of the TMT as regards tenure (10 times), education (5 times), functional background and age (4 times each) and gender (once). The second most studied concept is TMT diversity, with particular reference to tenure diversity (11 times), functional diversity (9 times), educational diversity (6 times) and age diversity (6 times). Gender diversity has not been studied at all. Dissimilarity is the least studied construct in TMTs, with tenure dissimilarity and age dissimilarity investigated only three times, and educational dissimilarity studied twice. Functional dissimilarity and gender dissimilarity have not been studied at all in TMTs.

Some of the names and the types of demographic attribute studied are slightly different, so where one researcher might use industry experience as the term for number of years in the industry, another might use company tenure or TMT tenure. For the purposes of this review, these have all been grouped under tenure. There are also subtleties in the way various demographic attributes have been measured. With regard to education, for example, Smith et al. (1994) measure number of years spent being educated, whereas Jackson et al. (1991) measure highest attainment gained, curriculum and prestige of university. The point here is not how each measure was derived, but which construct and at which level it was studied. Furthermore, the inverted measures of diversity used by Michel & Hambrick (1992) to study homogeneity have also been included under diversity. In table 2.1 and in the text that follows, clarification is made as appropriate so as to preserve the integrity concerning the direction of the results.

A very few studies (e.g. Jackson et al., 1991) included unusual measures of dissimilarity and diversity, such as military experience which have not been included as they are not representative of the field as a whole.

The two dependent variables typically studied in TMT demographic variation research are turnover and organisational performance, but again, not all studies include both with most only investigating one or the other.

Each of the 20 studies has been positioned according to which demographic variable and corresponding outcome variable was investigated in Table 2.2.

2.5.1 Findings Concerning Top Management Team Demography

Twelve of the 20 studies under review considered one or more aspects of TMT demography and organizational performance, and one considered TMT demography and turnover. Each aspect of demography will now be considered in turn starting with those using organizational performance as the dependent variable.

2.5.1.1 Age Demography and Organizational Performance

Age is deemed to be important in 'upper echelons' research generally because age is deemed to be a proxy for a person's life stage, experience, values and outlook on life. It is widely held that the central tendency on age of a TMT will have an impact on creativity and team decision-making (Hambrick & Mason, 1984; Williams & O'Reilly, 1998). Youthful managers are thought to be more innovative, creative and more prepared to take risks than older managers who may be more conservative in decision-making (Hambrick & Mason, 1984). Thus, it has been suggested that age similarity will result in team members getting along better, but less creative decisions (Williams & O'Reilly, 1998).

Four studies investigated age demography and organizational performance. Three suggested some effects with regard to the lower average age of the TMT. Hermann & Datta (2005) found that younger TMTs were more likely to actively seek to diversify; Wiersma & Bantel (1992) found that younger TMTs were more likely to be involved in strategic change; and Norburn & Birley (1988) found that younger TMTs outstrip inter-industry means for sales and employee growth. A fourth study (Bantel & Jackson, 1989) found no effects.

Table 2.2 TMT Demographic Variation Studies by Dependent Variables

Team Turnover			Organisational Performance		
Age Demography Jackson et al. (1991) Wiersma & Bird (1993)	Age Dissimilarity Jackson et al. (1991) Wagner et al. (1984) Wiersma & Bird (1993)	Age Diversity Jackson et al. (1991) Wagner et al.(1984) Wiersma & Bantel (1992) Wiersma & Bird (1993)	Age Demography Bantel & Jackson, (1989) Norburn & Birley (1988) Wiersma & Bantel (1992) Hermann & Datta (2005)	Age Dissimilarity Nil	Age Diversity Bantel & Jackson (1989) Glick et al.(1993) Knight et al.(1997) Murray, (1989) Wiersma & Bantel (1992)
Functional Demography Nil	Functional Dissimilarity Nil	Functional Diversity Krishnan et al. (1997)	Functional Demography Finkelstein (1992) Michel & Hambrick (1992) Norburn & Birley (1988) Hermann & Datta (2005)	Functional Dissimilarity Nil	Functional Diversity Bantel & Jackson (1989) Carpenter, (2000) Glick et al. (1993) Knight et al.(1997) Hambrick et al.(1996) Krishnan et al.(1997) Michel & Hambrick, (1992) Murray (1989) Smith et al.(1994) Wiersma & Bantel (1992)
Educational Demography Nil	Educational Dissimilarity Jackson et al. (1991) Wiersma & Bird (1993)	Educational Diversity Jackson et al. (1991) Wiersma & Bantel (1992) Wiersma & Bird (1993)	Educational Demography Bantel & Jackson (1989) Norburn & Birley (1988) Hambrick et al. (1996) Wiersma & Bantel (1992) Hermann & Datta (2005)	Educational Dissimilarity Nil	Educational Diversity Bantel & Jackson, (1989) Carpenter, (2000) Hambrick et al.(1996) Knight et al.(1997) Murray, (1989) Smith et al. (1994) Wiersma & Bantel (1992)
Gender Demography Nil	Gender Dissimilarity Nil	Gender Diversity Nil	Gender Demography Norburn & Birley (1988)	Gender Dissimilarity Nil	Gender Diversity Nil
Tenure Demography Jackson et al. (1991) Wiersma & Bantel (1992)	Tenure Dissimilarity Jackson et al. (1991) Wagner et al. (1984) Wiersma & Bird (1993)	Tenure Diversity O'Reilly et al.(1993) Jackson et al. (1991) Wagner et al. 1984 Wiersma & Bird (1993)	Tenure Demography Carpenter, (2000) Finkelstein & Hambrick (1990) Hambrick et al.(1996) Jackson et al. (1991) Michel & Hambrick (1992) Norburn & Birley (1988) Smith et al.(1994) West & Anderson, (1996) Wiersma & Bantel (1992) Hermann & Datta (2005)	Tenure Dissimilarity Nil	Tenure Diversity Bantel & Jackson, (1989) Carpenter, (2000) Glick et al. (1993) Hambrick et al.(1996) Knight et al.(1997) Michel & Hambrick, Murray, (1989) O'Reilly et al.(1993) Smith et al.(1994) Wiersma & Bantel (1992)

2.5.1.2 Functional Demography and Organizational Performance

Hambrick & Mason (1984) argued that functional background shapes behaviour and the way in which people identify and formulate problems. They based this proposition on prior research into individual managers which suggested that single managers are rarely able to take a conceptual view of organizational problems, but instead, define organizational problems only according to the functional area in which they operate. To illustrate, they argued that managers from backgrounds in output functions (i.e. marketing and sales) would view growth as important and promote that outcome accordingly. By way of contrast, operations managers from input functions (ie. production) would likely emphasise efficiency.

Four studies investigated functional demography and organisational performance. Generally this has taken the form of testing the number of executives in throughput functions such as production, and output functions such as finance and sales and marketing. Demography in this context is concerned with the dominant background of the executives in the team. Norburn & Birley (1988) looked at the relationship between which types of functional background accounted for improved financial performance. They found that TMTs with a predominantly finance or marketing background were more likely to head up companies which had higher levels of productivity (sales per employee). Similarly, Finkelstein (1992) noted a positive relationship (albeit marginally statistically significant) between the number of top management team members having a finance background and diversification posture. This latter finding is strengthened by the substantial and statistically significant finding of Michel & Hambrick (1992) that TMTs populated by more finance and legal executives were more likely to implement strategic change in highly interdependent firms. The fourth study investigating the relationship between functional background and international diversification found no relationship (Hermann & Datta, 2005).

2.5.1.3 Educational Demography and Organizational Performance

Formal education, particularly tertiary education, aims to provide a person with systematic processes for assessing problems and analysing information. Within the different disciplines, knowledge and problem solving tools are often valued differently. For example, the hard sciences conventionally value immutable laws and objective, analytical approaches to knowledge ordering and framing problems. By contrast, arts and humanities traditionally seek to overturn previous knowledge, placing greater emphasis on innovation and creativity. It has been suggested that educational similarity in teams will result in information being assimilated much more quickly, meaning that a decision can be made speedily (Hambrick & Mason, 1984). Five studies investigated educational demography and organisational performance. Findings reported include: greater education amongst top management team members is associated with: (a) enhanced firm performance and productivity (Norburn & Birley, 1988); (b) greater propensity to change corporate strategy (Wiersma & Bantel, 1992); (c) a propensity to engage in technical and total innovation (Bantel & Jackson, 1989); (d) a propensity towards international diversification (Hermann & Datta, 2005); and (e) enhanced scoping of strategy and better execution of strategic plans (Hambrick, Cho & Chen, 1996).

2.5.1.4 Gender and Ethnicity Demography and Organizational Performance

Although ethnicity and gender are obvious surface attributes of demographic difference, they are missing in the TMT literature. Moreover, neither has received anywhere near the attention of other attributes in the wider literature concerning demographic differences (Williams & O'Reilly, 1998). With regard to ethnicity, this omission is due to the fact that racial diversity is virtually non-existent in top management teams studied in the U.S.[1] (Williams & O'Reilly, 1998). As with race, the effects of gender differences on performance have not been a feature of TMT studies. The fact that homogeneous groups of white Western males tend to dominate the upper echelons of US firms and their subsidiaries in Western nations has stymied this line of inquiry (Williams & O'Reilly, 1998). Those who have tried to include gender as a measurement variable have either not reported it (see for example, Norburn & Birley, 1988), or later dropped it from analysis because of the consistently low representation of women on top management teams (see for example, Jackson et al., 1991). Gender in the wider managerial literature has typically concentrated on performance ratings of women in groups by their male counterparts

---

[1] The few exceptions are: Wiersma & Bird (1993), who wholly studied Japanese firms; the study of Irish and American firms variously reported by Smith et al. (1994), Clark et al. (1997) and Knight et al. (1999); and Norburn & Birley (1988) who conducted an Anglo-American study.

(see for example, Nieva & Gutek, 1980; Ruble, Cohen & Ruble, 1984; Swim et al., 1989), effects of being in a gender minority (Tsui et al., 1992) and the likelihood of promotion in organizations where more women are already at the top (Cohen & Bailey, 1997).

Only one study investigated gender demography and organisational performance, and found that TMTs that included women were in organisations with greater organisational growth (p <.10), particularly with regard to sales per employee (Norburn & Birley, 1988). Unfortunately, neither the details of the correlation nor the interpretation of the finding are reported, so it is unclear as to what it means. As will be explained later, Norburn & Birley (1988), unlike many other researchers in this field, did not confine themselves to the top two levels in the organisation, and they also included single managers in their analysis of TMTs.

2.5.1.5 Tenure Demography and Organizational Performance

Tenure is generally measured in terms of time spent in the top management team, time in the organisation and time in the industry. Pfeffer (1983) argued that tenure was an important determinant of process and performance due to familiarity with the organisation and with other members on the team. Specifically, he reasoned that similarity in organisational tenure would lead to increased communication and shared interpretation of information which would have positive effects on cohesive group processes. Research neither confirms nor denies Pfeffer's (1983) arguments with positive and negative findings being reported as the following illustrates.

Industry tenure, as a proxy for experience, is reasoned to colour the way in which executives view current strategic opportunities and threats (Hambrick & Mason, 1984). Organizational tenure, or the amount of time a person has spent in an organization, is related to the degree to which individuals become exposed the norms, values and practices, or organizational culture. In turn, organizational tenure is believed to affect the way in which top management members interact with one another, develop a common language and approach to problem solving (Clark et al., 1997). Team tenure measures the degree to which a team is socialised into itself, and is important to understanding top management team functioning regarding power hierarchies and communication patterns between members (Clark et al., 1997). As this study is primarily concerned with teams, it is TMT tenure that is of particular interest and will be reported on in the following paragraphs.

Ten studies investigated the effects of TMT tenure demography and organisational performance. Results suggest that TMTs with short tenures tend to: have better organisational performance[2] (Carpenter, 2000); pursue more novel strategies that deviate from industry norms (Finkelstein & Hambrick, 1990; Hambrick et al., 1996) and more enthusiastically seek international diversification (Hermann & Datta, 2005).

At the same time, longer tenured TMTs are found by several studies to be: more socially integrated, which in turn has a positive effect on longer term return on investment and sales growth (Smith et al., 1994); better at seeing through the implementation of corporate change strategies (Wiersma & Bantel, 1992; Michel & Hambrick, 1992); and to have more innovations impacting favourably on staff-well-being (West & Anderson, 1996).

Only one further study did not find any statistically significant relationships between TMT tenure and organisational performance (Norburn & Birley, 1988).

2.5.1.6 Age and Tenure Demography and Team Turnover

Three studies have investigated aspects of team demography and team turnover. Jackson et al. (1991) found that older teams (i.e. teams in which the average mean age is high) tend to have higher rates of turnover, due to older executives being closer to retirement. Wiersma & Bantel (1992) similarly found that TMTs with longer tenures were also more likely to experience more turnover, again, due to executives with the longest tenures tending to be older and closer to retirement. On the other hand, Wiersma & Bird (1993) found that younger members (i.e. those between 51 and 55 years of age) with shorter tenures were more likely to leave the team. This may be attributable to cultural differences. In the U.S. and U.K. at 55 one is considered to be in the twilight of one's career and probably settling down rather than still climbing the corporate ladder. By way of contrast, in Japan (the site of Wiersma & Bird's 1993 study), there is no forced retirement age, and therefore no natural attrition due to retirement at a particular age. As executives can go on working until they choose to stop the world is still their oyster at 55, and they may be leaving due to career advancement. As alluded to previously, most TMT studies are conducted in the U.S.,

[2] as measured by tendency towards internationalization

so the cross cultural effects of demographic variation are not particularly defined or explored.

2.5.1.7 Summary of Findings Concerning Demography and Organizational Performance

The foregoing review of TMT demography suggests that, despite decades of research into 'upper echelons' theory, what is actually known about the central tendencies of demographic attributes in top management teams is scant indeed. Encouragingly however, the literature is not quite as "uninterpretable" (Preim et al., 1999) as is made out, and the overall direction of the results for organisational demography across the studies is discernible.

On balance, the research literature suggests that TMTs made up of younger, highly educated executives, from predominantly finance backgrounds, will have a positive effect on organizational performance. TMTs of above average age will tend to have more turnover due to executives approaching retirement. TMTs which have been together longer are more socially integrated and will achieve more in terms of innovation, strategic change and profitability over the longer term. The next section reviews what is known about dissimilarity in TMTs.

2.5.2 Findings Concerning Dissimilar Individuals in Top Management Teams

Only three of the 20 studies under review considered demographic variation in terms of dissimilarity in TMTs, and the dependent variable in each case was turnover. This is appropriate because organisational performance results from the collective effort of the TMT, whereas individuals' leaving the team is ultimately an individual decision. These works conclude that the most dissimilar individual in terms of age (Wagner et al., 1984) and educational attainment (Jackson et al., 1991; Wiersma & Bird, 1993) is most likely to leave before his or her fellow TMT members. Only Wiersma & Bird (1993) failed to find that tenure dissimilarity was a statistically significant predictor of turnover, and they attributed this to cultural differences (there is no forced retirement age in Japan (which affects tenure distances amongst individuals) as there is in the U.S. the location for the other two studies).

It is quite amazing that such a tiny number of TMT studies have investigated the dissimilarity in demographic attributes. It is clear, from referring to Table 2.2, that this is an unexploited avenue for research.

The pithy nature of the findings presented in this section merely serves to endorse the construct approach used here to sorting the literature. The next section will address the studies concerning TMT demographic diversity.

2.5.3 Findings Concerning Top Management Team Diversity

Eleven of the 20 studies under review have considered aspects of demographic diversity within TMTs. The findings are more uneven than those concerning demography and dissimilarity, with several contrasting and seemingly contradictory effects being observed. Specifically, the results for organisational performance and turnover appear to be different, and it is perhaps the melding of these two together in previous studies and in reviews of the field that leads to the critics' frustration concerning uneven results.

The rest of this section will address these two outcome variables separately, taking organisational performance first. What is striking to note in the review which follows is that the findings for demographic diversity and performance are both substantial and rich, and they tell a reasonably coherent story. In contrast, the findings for demographic diversity and turnover across fewer studies are more uneven. However, it is clear that temporal diversity, that is age and tenure diversity, are the strongest predictors of team turnover.

As this is supposedly a very 'lumpy' literature which has attracted much criticism, all of the studies which have addressed each variable are included below. In the interests of transparency and thoroughness, all the results, even studies that report no associations are noted.

2.5.3.1 Age Diversity and Organizational Performance

Six studies investigated the relationship between age diversity and organisational performance. Of all of the diversity variables studied, age diversity attracts the fewest statistically significant results.

Knight et al. (1999) found that age diversity was associated with disagreement in TMTs, which corresponds with common sense and conforms to theoretical predictions that people of different ages have different viewpoints. The more age groups are represented on a team, the less likely the team is to agree on issues. Unlike most studies in this area, Murray (1989) did not assume that the effects of

TMT diversity are stable over time and are the same across industries. His study compared short-term performance (efficiency) and long term performance (adaptability to changing environment) across two industries (oil and food production). He found that temporal diversity (which he measured as age and tenure diversity together) was positively associated with long-term performance in stable but not less stable environments. Of course, because the two were measured together, the unique effect of age diversity is unclear.

Four out of the six studies, however, found no effects (Bantel & Jackson, 1989; Glick et al., 1993; Wagner et al., 1984; and Wiersma & Bantel, 1992).

2.5.3.2 Functional Diversity and Organizational Performance

Functional diversity is supposed to be important for TMT decision-making and organisational performance because different departments or functions in organisations often operate on different reward structures or operating procedures, especially when they are also geographically dispersed from the central head office (Glick et al., 1993). Thus it follows that executives heading up these departments who are members of the TMT will view organisational problems and solutions differently to their peers from other departments, thereby offering a broader perspective.

Nine of the 20 studies under review investigated the relationship between functional background diversity and organisational performance. The results have been fairly consistent, in that functional diversity is good in some situations and functional homogeneity is better in others. Only two studies failed to find a statistically significant result with this variable (Smith et al., 1994 and Wiersma & Bantel, 1992).

Functional background diversity has been found to relate to organisational performance in a variety of ways. For example, Hambrick et al. (1996) found that functional diversity was positively associated with increased market share, increased profits, and bold strategic change to gain competitive advantage (Hambrick et al., 1996). In a similar vein, Bantel & Jackson (1989) found a relationship with functional diversity and the number of administrative innovations introduced by TMTs.

Whilst the above findings tend to suggest that functional diversity is beneficial for performance, research also shows that there are situations in which functional homogeneity is appropriate. For example, Murray (1989) found that functional homogeneity was positively related to short term company performance in stable environments. He argued that this result was due to homogenous TMTs of engineers (in the oil industry) being more efficient, that is to say, communicating more effectively and coordinating their actions better. This finding is further supported by Krishnan et al. (1997), who found that the functional homogeneity of TMTs is positively related to performance in post-acquisition companies. In a similar vein, Hambrick et al. (1996) found that functional homogeneity makes it easier to mobilise a company around a particular innovation.

Carpenter (2000), in a bold contribution to the field, addressed one of the conundrums which has plagued this literature, that is, is there a point at which diversity ceases to be good for performance? This goes right to the heart of the value-in-diversity (Cox et al., 1991) premise that diversity is good and more diversity is better. Consistent with expectation, Carpenter (2000) found that functional diversity predicted organisational performance (return on assets) in firms with low levels of internationalization but was a negative predictor in firms with high levels of internationalization. Almost identical findings were observed by Michel & Hambrick (1992), who found that companies with different diversification strategies benefited from different levels of functional diversity. For example, functional homogeneity (that is TMTs populated mainly by finance and legal executives) significantly increased profitability (as measured by return on assets) and resulted in strategic change in firms that were less interdependent because the TMT is steeped in knowledge about core functions. On the other hand, in highly interdependent firms, where formal controls from head office are less in evidence, the knowledge base of the TMT needed to be more diverse, and increased profitability in such firms is achieved by functionally diverse teams (Michel & Hambrick, 1992).

Glick et al. (1993) found that functional diversity meant more diverse beliefs about the efficacy of advertising and more rich communication. Both these findings are borne out by Knight et al. (1999) who also found that functional diversity led to more diverse beliefs instead of strategic consensus, and more interpersonal conflict. The high levels of disagreement found by these studies amongst functionally diverse teams provides a possible explanation for Hambrick et al.'s (1996) findings, that such TMTs are slower and less likely to respond to competitors' actions, probably because they disagree on the way forward. The increased cognitive diversity observed by

Glick et al. (1993) and Knight et al. (1997) may also account for positive relationship between functional diversity and magnitude of response to competitors actions noted by Hambrick, et al. (1996). That is to say, the various functions represented on the TMT give a solid and diverse cognitive base from which to respond strategically.

2.5.3.3 Educational Diversity and Organizational Performance

Educational diversity within TMTs is generally held to have a positive effect on the information processing capacity of teams faced with solving complex, non-routine problems (Bantel & Jackson, 1989). Information in educationally diverse teams is argued to be subjected to more careful analysis and better use of information than in homogeneous groups (Williams & O'Reilly, 1998). Seven studies investigated the relationship between educational diversity and organisational performance. The results show that educational diversity is statistically significantly associated with return on investment (Smith et al., 1994); bold strategic change (Wiersma & Bantel, 1992; Hambrick et al., 1996); and market share and profitability (Hambrick et al., 1996); and is marginally significantly associated with return on assets, especially in firms with high levels of internationalization (Carpenter, 2000).

On the other hand, the evidence suggests that educationally diverse TMTs: find strategic consensus difficult to achieve (Knight et al., 1997); are slower to act as a first movers to gain competitive advantage, are slower to respond to competitors' innovations (Hambrick et al., 1996); and are less efficient in terms of short term company performance (Murray, 1989). Only one study found no effects (Bantel, & Jackson, 1989).

2.5.3.4 Gender and Ethnic Diversity and Organizational Performance

None of the studies under review included gender diversity. Only one could be found that attempted to investigate ethnicity at this level, that by Carpenter (2000), who used TMT nationality diversity as a control variable. It was measured as the representation of non-U.S. born executives on the TMT computed using Blau's (1977) index of heterogeneity. In all 247 TMTs in Carpenter's (2000) sample, none had more than one non-U.S. born member.

2.5.3.5 Tenure Diversity and Organizational Performance

Eleven of the TMT demographic variation studies under review investigated the relationship between tenure diversity and performance. It seems that, on balance, tenure diversity is detrimental for short term performance, but beneficial in the long term.

Tenure diversity is negatively associated with: short term performance (Murray, 1989; Smith et al., 1994); strategic change in stable, competitive environments (Murray, 1989; Carpenter, 2000); informal communication (Smith et al., 1994); and propensity to respond and speed of response to competitors (Hambrick et al., 1996). All of this means that TMTs which are more similar with respect to length of service are more likely to engage in adaptive change (O'Reilly et al., 1993).

However, tenure diversity is also found to be positively associated with: strategic consensus (Knight et al., 1997); long term organisational performance (Carpenter, 2000); long term strategic change (Murray, 1989; Michel & Hambrick, 1992); and profitability and change in market share (Hambrick et al., 1996). Two further studies found no effects (Bantel & Jackson, 1989; Wiersma & Bantel, 1992).

2.5.3.6 Summary of Findings Concerning Diversity and Organizational Performance

Far from being "uninterpretable" (Priem et al., 1999), the results for the demographic diversity literature tell a remarkably coherent story. All aspects of TMT diversity have been found to have a generally unfavourable effect on short term efficiency gains, but are consistently associated with long term performance, strategic change and profitability. Very few studies have investigated the link to process, and none have investigated immediate outcomes such as decision choices.

Having considered the TMT diversity studies with regard to organisational performance, the next section of this chapter will deal with the relationships with turnover. The results across the six studies that have investigated diversity and turnover are somewhat mixed, but the temporal aspects of diversity appear to be the most reliable predictors.

2.5.3.7 Age Diversity and TMT Turnover

Three studies have investigated the relationship between age diversity and team turnover, with age diversity found to be associated with team turnover in all three (Jackson et al., 1991; Wagner et al., 1984; Wiersma & Bird, 1993).

2.5.3.8 Functional Diversity and TMT Turnover

One of the 20 TMT studies under review investigated the relationship between functional diversity and team turnover. Krishnan et al. (1997) found that functional background diversity was positively associated with team turnover. In other words, TMTs which are made up of people from similar backgrounds will tend to stick together and not have people leaving as often.

2.5.3.9 Educational Diversity and TMT Turnover

Three studies investigated the relationship between educational diversity and team turnover with mixed results. Wiersma & Bantel (1992) found no results, Jackson et al. (1991) found that diversity in curriculum had a marginally statistically significant association with individuals leaving a TMT, whilst Wiersma & Bird (1993) found that educational diversity was a statistically significant predictor of team turnover.

2.5.3.10 Gender Diversity and TMT Turnover

No studies investigated the relationship between gender diversity and team turnover.

2.5.3.1.1 Tenure Diversity and TMT Turnover

Four studies have investigated the relationship between tenure diversity and team turnover, with tenure diversity found to be associated with team turnover in three out of four (O'Reilly et al., 1993; Wiersma & Bird, 1993; Wagner et al., 1984), whilst one study found no association between team tenure diversity and team turnover (Jackson et al., 1991). The three positive findings mean that the more tenure diverse the team, the greater the proportion of team members who will leave. As to the types of persons that will leave, it is likely that those individuals that share particular socializing events in the lifetime of the firm will leave around the same time (Wagner et al., 1984).

2.5.3.12 Summary of Findings Concerning Diversity and TMT Turnover

Age and tenure diversity consistently predicts the proportion of the TMT that will leave, whilst functionally homogeneous teams stick together.

The next section will address the features and limitations in methodology in the literature.

2.6 Features of TMT Demographic Variation Research

This chapter has discussed the findings of demographic variation research with particular reference to TMTs. This section examines four important features of the existing research literature, namely:

(1) inconsistent specification of the term top management team;

(2) use of indirect research methods;

(3) relatively small sample sizes; and

(4) causal assumptions.

Each of these will now be explored in turn.

2.6.1 Feature 1: Inconsistent Specification of Top Management Team

The term top management team in common parlance, applies to a small group of influential senior executives at the strategic apex of an organization (Hambrick et al., 1996). There are two interrelated parts to this issue explored in this section. The first is which level in the organisational hierarchy is addressed. This subject was one of the major criticisms about TMT research made by Pettigrew (1992). The second is size of the TMT, and is still an issue for TMT researchers (Carpenter et al., 2004). Obviously, the higher one climbs into the strategic apex, the fewer people there are.

In empirical studies this definition has been applied so that studies of TMTs have been conducted with: senior managers from the two highest executive levels in an organisation (e.g. Wiersma & Bantel, 1992; Murray, 1989; Carpenter, 2000); selected members from amongst those identified by the chief executive officer

(CEO) as members of the TMT (e.g. Bantel & Jackson, 1989; Glick et al., 1993); or selected members from all those involved in the event or decision under investigation (e.g. Smith et al., 1994); or all managers at the vice-president level[3] or higher (e.g. Michel & Hambrick, 1992; Carpenter, 2000). By way of contrast, Norburn & Birley (1988) used TMTs that ranged in size from 1-28 persons, and clearly used several management levels lower down the hierarchy than most other researchers in this field. Also there is clearly an issue in this last cited work regarding the fact that one person cannot constitute a team, which will not be critiqued here. Suffice it say, the cited works in this paragraph support the contention that there is inconsistent specification of TMT across the literature. A further issue concerning TMT size is that researchers are not consistent about controlling for team size, with many studies failing to do this, making comparisons of results across studies difficult (Carpenter et al., 2004).

Studies which rely upon secondary archival data for the supply of statistical demographic data tend to use whole TMTs or subsets thereof. Murray (1989) and Jackson et al. (1991) both used two classifications of TMT, an exclusive or elite group which included selected titles (Chairman of the Board; Vice Chairman; Chief Executive Officer; Chief Operating Officer; President; Senior Vice President and Executive Vice President) and an inclusive or non-elite group which included all executives on the lists in the public archival database.

2.6.2 Feature 2: Sample Sizes

As revealed in Table 2.3 below, the sample sizes of much of the work in this area are small. This is because gaining access at all to TMTs is extremely difficult. The number of TMTs ranges from 5 to 247 across all methods (public archival data, questionnaire and interview), with the mean being 60 TMTs. Studies which rely on public archival data achieve a slightly higher mean of 62, questionnaire survey methods slightly lower, 51, whilst there is only one interview study which had a sample size of 24 TMTs. It means that the relationships between the variables need to be very strong indeed to show effects. Many of the cited studies find few effects. For example, Glick et al. (1993) found only 3 significant effects out of a possible 49

[3] Operations level in the UK

relationships (sample of 79 teams), and Jackson et al. (1991) in one instance found 4 significant associations from a possible 24 relationships (sample 93 teams).

2.6.3 Feature 3: Use of Indirect Research Methods

Because gaining access to TMTs is quite difficult (Priem et al., 1999), indirect research methods are often used as a means of overcoming this difficulty. Moreover, it is claimed that one of the great strengths of demography research is that it can be undertaken from a distance (Pfeffer, 1983). This is because demographic information on top management teams, together with company performance data, is readily available through public databases held by regulatory institutions such as Companies House in the UK. Companies are under obligation to provide demographic information on their directors and board members (i.e. upper echelons) and must continue to update it along with performance data every year. Furthermore, detailed demographic information is collected by agencies such as Standard & Poors in the USA or Dun & Bradstreet in the UK and USA, who then sell it for marketing and research purposes.

Using public archival data about top management teams is a relatively straightforward way to collect demographic information and draw conclusions about the relationship between demographic variables and performance and team turnover. This is occasionally supplemented by content analysis of corporate documents, media clippings and press releases (see for example, Tetlock, 1979; Murray, 1989; Hambrick et al., 1996). Usually, these sources are scrutinised for information concerning a significant company event or major corporate decision, such as a merger, joint venture or acquisition. Inferences are then made by researchers as to what decision-making processes were involved. Assertions they make about process must then be regarded with caution as these are not directly measured.

Secondary archival data cannot by itself allow one to get close to a top management team, or understand what goes on inside the team. Whether public archival data are used alone to make assertions about demographic variation and performance, or whether they are used in conjunction with media clippings and corporate reports, the resulting analysis can, at best, only provide a remote and highly subjective appraisal of what is happening within top management teams. Yet most of what we know about top management teams is based on public archival data (see for example Finkelstein, 1992; Haleblian & Finkelstein, 1993; Jackson et al., 1991; Michel & Hambrick, 1992; Murray, 1989; Norburn & Birley, 1988; Simons, 1995; Wagner et al., 1984; Wiersma & Bantel, 1992; Wiersma & Bird, 1993).

Pettigrew's (1992, pp75) observation, first noted in chapter 1, is still generalisable across the vast majority of work in this areas: *"...no one has ever been anywhere near a top team in an organizational setting, either to directly observe a team in action, or to interview the members of the links between their characteristics and structure, processes of communication and decision-making and their impact and performance".*

As Table 2.3 shows, in the 24 studies identified earlier in this chapter as having studied TMT demographic variation, 13 rely exclusively on public archival data. Two studies use public archival data and media sources, two further studies use public archival data and a questionnaire survey, two use just questionnaires. There are five studies that claim to have used interviews in combination with other methods, but closer reading reveals that in three of these (Clark et al., 1997; Smith et al., 1994; and Knight et al., 1999), the 'interviews' were simply meetings with the CEO to: (a) explain the research; (b) obtain financial performance data; and (c) obtain authorisation to survey the TMT[4]. A fourth (West & Anderson, 1996) similarly used 'interviews' simply to explain the research to the CEO in order to gain access to survey the TMT. The fifth, (Glick et al., 1993) used telephone interviews as a screening mechanism in assembling their sample which also involved public archival data and a questionnaire survey. They also claim to 'draw upon 120 interviews with CEOs' from earlier work in developing their research model. Three features of this study are worthy of note. First, the number of CEOs did not exceed 30 (interviews were conducted four times over a 24 month period); second, the research did not address whole TMTs; and third, neither the qualitative nor the quantitative data arising from the interviews is reported in the 1993 study. The study however, is represented here, as the source data is that reported by O'Reilly et al. (1993) who interviewed the CEOs of 24 companies about TMT processes in relation to implementing strategic change.

[4] These three cited works are all based on the same study, they report different aspects in the various papers.

Table 2.3 Sample Sizes and Methods in TMT Demographic Variation Studies

AUTHOR/DATE	SAMPLE SIZE	METHOD
Bantel & Jackson (1989)	199 TMTs, Banks	Questionnaire Survey of CEO & HR Director
Boone et al.(2004)	5 TMTs Dutch Newspaper Companies, 5 x 5 year periods (53 Execs)	Public Archival Data
Carpenter (2000)	247 TMTs from Standard &Poor's Industrial Index 1990 - 1997	Public Archival Data
Clark, Smith, Sims, Flood, Moore, Morley & O'Regan (1997)	21 TMTs from Irish High Technology Firms 57 TMTs from US High Technology Firms	Meeting with CEOs & Questionnaire Survey of TMT
Finkelstein & Hambrick (1990)	35 TMTs in Computer Firms 35 TMTs in Chemical Firms 30 TMTs in Natural Gas Distribution Firms	Public Archival Data
Finkelstein (1992)	36 TMTs in Computer Firms 36 TMTs in Chemical Firms 30 TMTs in Natural Gas Distribution Firms	Public Archival Data & Questionnaire Survey of TMT
Glick, Miller & Huber (1993)	79 TMTs of SBUs	Public Archival Data & Questionnaire Survey of TMT
Halebian & Finkelstein (1993)	26 TMTs in Computer Firms 21 TMTs in Natural Gas Distribution Firms	Public Archival Data
Hambrick, Cho & Chen (1996)	32 TMTs US Airlines	Public Archival Data & Media Sources
Hermann & Datta (2005)	122 TMTs large firms	Public Archival Data
Jackson, Brett, Sessa, Cooper, Julin & Peyronnin (1991)	93 TMTs in Bank Holding Firms (625 Executives)	Public Archival Data
Knight et al.(1997)	53 TMTs High Tech Firms in USA (230 Executives) 26 TMTs of US MNC in Ireland (98 Executives)	Meeting with CEO of 53 firms & Quest. Survey of TMT
Krishnan, Miller & Judge (1997)	147 TMTs Merged or Acquired 1986-88	Public Archival Data
Michel & Hambrick (1992)	134 TMTs from 1974 Fortune 500	Public Archival Data
Murray (1989)	26 TMTs of Integrated Oil Firms 58 TMTs from Food Firms	Public Archival Data & Media Sources
Norburn & Birley (1988)	150 TMTs from 5 industries (953 Executives)	Public Archival Data
O'Reilly, Snyder & Boothe (1993)	24 TMTs Electronics Industry	Interview with CEO
Simons, Pelled & Smith (1999)	57 TMTs in Electrical Component M/F Firms	Public Archival Data
Smith, Smith, Olian, Sims, O'Bannon & Scully (1994)	53 TMTs in Technology Based Firms (230 Executives)	Meeting with CEO & Questionnaire Survey of TMT
Wagner, Pfeffer & O'Reilly (1984)	31 TMTs from 1976 Fortune 500 (599 Executives)	Public Archival Data
West & Anderson (1996)	27 TMTs NHS Trusts	Meeting with CEO & Questionnaire Survey of TMT
West & Schwenk (1997)	39 TMTs in Machine Tools Industry 26 TMTs in Electronics Components Firms	Questionnaire Survey of CEO & 2 TMT Reps
Wiersma & Bantel (1992)	87 TMTs from 1980 Fortune 500	Public Archival Data
Wiersma & Bird (1993)	40 TMTs Listed on Tokyo Stock Exchange (220 Executives)	Public Archival Data

This means that there is only one, single study in the published TMT demographic literature that has actually come face to face with CEO's of TMTs in an organizational setting and used TMTs as the unit of analysis, and the number of teams involved was 24[5].

2.6.4 Feature 4: Causal Assumptions

Following the 'upper echelons' tradition, demographic variation research into TMTs is based on the causal assumption that demographic attributes are principally responsible for choices or decisions (Hambrick & Mason, 1984), which in turn, affect financial performance. There are two issues to raise about this. First is the efficacy of using financial performance data, without reference to time lag or environmental considerations. The second is that other mediators, such as strategic choice, group interaction processes or group affect, remain unmeasured in most studies (Priem et al., 1999), giving no challenge to the causal assumption.

As to the first issue, it is eloquently and succinctly expressed in the following quote: *"Strategies are devised and implemented by top management, and remain appropriate so long as the economic imperatives driving them remain in place. The lag between cause (the top management group and the broad strategies they devise) and effect (financial performance) is variable, as is the effect's duration. This poses serious methodological problems for any researcher investigating links between management and performance, and may explain the greater variance between popular beliefs about managers' efficacy and the empirical evidence"* Murray (1989, pp 139). It is of interest to note that such a consideration, despite its seriousness, is rarely if ever, mentioned in TMT demographic variation studies.

A further assumption that characterises research in this genre is that a bad decision, means a bad decision-making process (see Priem et al., 1999). On the other hand, a decision deemed to be good, either by reference by media coverage or financial performance data, is held to indicate a good decision-making process. Yet, exactly how demographic variation affects the implementation of a decision remains virtually unexplored (Priem et al., 1999). At least one study, Glick et al. (1993) surveyed the CEO and other TMT representatives as to their reflective

[5] A further study (Hambrick, 1981) interviewed CEOs of 20 organisations and surveyed other members of the respective TMTs. Hambrick's (1981) study did not include demographic variables, so is not addressed in this chapter, it is however, included in Chapter 3 on TMT processes.

retrospections on a significant company issue on which to base research conclusions as to how demographic variation is linked to process. Unfortunately, retrospective reflection can result in biased recollections of what went on, depending on the individual's role and commitment in the decision-making process, and how they benefited or suffered as a result of the decision outcome (Amason, 1996). Both Glick et al. (1993) and Clark et al. (1997) enquired as to TMT members perceptions of processes such as cohesion and cognitive diversity. However, as noted by O'Reilly et al. (1993), quantitative self report data do not necessarily reflect the reality of TMT processes. None of these studies used an independent measure of TMT process.

Inferring the effects of demographic variation on process is not unusual. In point of fact, the way that demographic variation is given precedence over process is fundamental to the theoretical ontology. Pfeffer (1983) argued that *"demography is an important, causal variable that affects a number of intervening variables and process and, through them, a number of organizational outcomes"* (pp 348).

One of the major criticisms of 'upper echelons' and 'organisational demography' research is that process remains a 'black box' the contents of which are currently unknown (Lawrence, 1997; Pettigrew, 1992). Contrary to popular opinion, neither Lawrence (1997) nor Pettigrew (1992) gave indications as to what should go into the 'black box'. Rather, as will be discussed in more detail in Chapter 11, they were concerned about the theoretical development of 'upper echelons' theory as it relates to methodological advances in empirical research.

Pfeffer (1983) contended that *"It is possible for demographics to do a better job of explaining variance in the dependent variables than measures of the presumed intervening constructs, for the reason that many of the intervening constructs are mental processes (attitude toward various elements of compensation, for example) that are more difficult to access and reliably measure"* (pp 351).

However, in order to address the criticisms of 'upper echelons' and 'organisational demography' research, particularly the inconsistent and contradictory findings, researchers will no doubt need to increasingly tackle the difficulties of accessibility and reliable measurement of process (Lawrence, 1997; Pettigrew, 1992). Only five studies have attempted to do this (Glick et al., 1993; Smith et al., 1994; Clark et al., 1997; Knight et al., 1999 and O'Reilly et al., 1993). These will be reviewed more closely in the following chapter.

There are some tortuous arguments that suggest that whilst demographic variation may affect organisational performance (as discussed above), organisational performance may affect demographic variation. For many years, the basis for these arguments stemmed, in large part from Kanter's (1997) notion of homosocial reproduction. This is the idea that organisations selectively weed out dissimilar individuals and diverse groups of people so that organisations become generic. This has been tempered somewhat with more recent developments in globalisation and the focus in domestic and international law concerning equal opportunities policies and anti-discrimination legislation (Jackson & Joshi, 2001). As discussed earlier in this chapter, the international management team literature is concerned with cultural representation of TMTs particularly in multi-national corporations, as they operate both at home and abroad (Adler, 1997). It is perhaps fair to say that in this latter arena, the notion that organisational performance affects demographic variation is more easily understood.

Generally however, the 'upper echelons' literature, without exception, does not have an issue with causality. In all expositions and modifications of the theory, the overall model is linear and sequential with demographics acting as the starting point and organisational performance being the end point (Hambrick & Mason, 1984; Hambrick et al., 1996; Hodgkinson, 2001; Carpenter et al., 2004). As this research purports to be a test of 'upper echelons' theory, the causality assumed is that demographic variation is a predictor of cognitive variation and team processes. However, in deference to the criticisms made at the beginning of this section, rather than just assume the relationships, they will be defined, tested and measured.

2.7 Other Considerations

Before leaving a discussion of TMT demographic research, it is appropriate to consider other literature that also includes TMTs as a feature of interest and other factors that are generally missing from 'upper echelons' research.

2.7.1 The International Management Team Literature

The international management team literature mentioned in the last section has relevance here. It reflects real world concerns for multi-national organisations to become multicultural to reflect the environments in which they operate, and grew from the recognition of the high costs associated with increasing expatriate failure in overseas postings (Jackson & Joshi, 2001). Some of the complexities within this research genre are that TMTs are often globally distanced and use virtual technology to communicate (Adler, 1997) which means that their specific demographic makeup is less likely to have an effect on TMT dynamics. Essentially, this type of research is concerned with cultural diversity as this is deemed to be a better proxy for underlying attitudes and life experiences than simple demographics as in 'upper echelons'. Whereas 'upper echelons' is concerned primarily with the differences between people, international management research sees the differences and similarities as being equal and is more concerned with 'cultural synergy', that is the extent to which the people involved albeit from different cultures, can work together (Adler, 1997, pp 107).

Whilst 'upper echelons' is a theory of descriptive relevance (Priem et al., 1999), that is to say, it describes processes and outcomes in relation to demographics, international management research regards cultural diversity as something to be 'managed' and harnessed as a resource. In other words, it can be manipulated to achieve certain outcomes. Cultural diversity is deemed to be advantageous when an organisation wants to expand perspectives and in specific circumstances such as new product launches, planning a new operation or assessing emerging trends from cross-national perspectives (Adler, 1997).

Interestingly, the underlying logic appears similar to that of 'upper echelons', that multi-cultural TMTs are more effective than bi-cultural or mono-cultural teams due to the fact that they bring many perspectives on a situation, *"but they frequently experience greater difficulty in integrating and evaluating these perspectives (thus causing losses in productivity due to faulty process)"* (Adler, 1997, pp 131). Ilgen, LePine & Hollenbeck (1999) posited that cultural diversity affects three aspects of decision making in international management teams: the definition of the problem, the sharing of information and conflict or consensus.

The two major differences between the 'upper echelons' and international management traditions are (a) that the benefits of diversity are situation specific and contingent on many factors; and (b) that diversity can be manipulated dependent on the circumstances. As a result, it is likely that international management team research has more practical relevance and prescription for TMTs than 'upper echelons' which is descriptive only.

2.7.2 Antecedents

There is a growing recognition that many contextual features could have a bearing in 'upper echelons' research (Hodgkinson, 2001; Edmondson et al., 2003; Carpenter et al., 2004). This is a topic that will be returned to in the next chapter, and again in Chapter 11.

Salient to the discussion in this chapter is that the studies reviewed in this chapter all tend to assume that ceteris paribus, situation specific factors have no influence on team processes or organisational performance (Papadakis, Lioukas & Chambers, 1998). On the other hand, it is commonsensical to assume that, TMT effectiveness can vary from one situation to another and hence situation specific factors should be included in 'upper echelons' research (Edmondson et al., 2003).

As will be highlighted in the next chapter, it is extremely rare for antecedents to be included in 'upper echelons' research leading one commentator to conclude:

"Researchers have barely scratched the surface in the quest to understand the causal antecedents and consequences of executive cognition" (Hodgkinson, 2001, pp 425). That this approach has hampered 'upper echelons' research to date is further attested by the following quote:

"Many more years of research will be needed to achieve a good understanding of how context shapes diversity's consequences" (Jackson & Joshi, 2001, pp 218). Despite this admonition, very little, even amongst the most recent 'upper echelons' research, includes consideration of antecedent factors (Carpenter et al., 2004).

The next section reviews the propositions arising from the studies reviewed in this chapter.

2.8 Propositions Arising From TMT Demographic Variation Research

From the foregoing it is clear that the challenges of studying top management teams are complex. In order to overcome some of the limitations observed in the extant research literature, an ideal study would:

(1) not rely on secondary archival data;

(2) use whole top management teams, not just the CEO or representatives;

(3) get face to face access to the TMT rather than rely on indirect methods;

(4) not use financial data as the only measure of performance; and

(5) investigate team processes, not make causal inferences based on demographic proxies.

Table 2.3 illustrates that there a pressing need for more field research, particularly that which seeks to get up close and personal to whole top management teams and observe them in their operational settings. It is simply incredible to find that research of such perceived importance to TMT and organisational performance has in most cases, been conducted without reference to the TMTs it concerns. Furthermore, only one study of 24 teams (O'Reilly et al., 1993), has ever gained face to face access in the inner sanctum of TMTs. The present research aims to go some way to addressing this need.

2.9 Conclusion

This chapter began by reviewing some of the seemingly straightforward propositions in the 'upper echelons' and 'organisational demography' theories. At face value, the idea that a range of demographic attributes represented on a team will result in cognitive variation and enhanced capabilities seems reasonable. However, the research tradition these theories have spawned has attracted severe criticism due to results appearing to be very uneven.

The confusion which is apparent in the literature has led some commentators to suggest that it is "uninterpretable" (Priem et al., 1999). The preceding review has attempted to redress some of this confusion and to unlock some of its opacity by distinguishing between demography (average levels of attainment on a particular variable), dissimilarity (individual's distance from others in the team on a particular variable), and diversity (degree of variation in team members' attainment on a particular variable). Notwithstanding the problems caused through conflating the constructs, the review in this chapter discussed additional fundamental areas for concern, namely most of these studies rely upon indirect research methods and assume causal effects of demographic and cognitive variation on team processes, and through them on performance. It was noted that inferences regarding team processes are often derived from demographic proxies, rather than exploiting the fact that team processes have been studied extensively by small group researchers.

The next chapter will investigate the small group literature for theory and research concerning team processes of relevance to TMTs.

CHAPTER 3

Theory and Research: Cognition, TMT Processes & Decision Belief

3.0 Overview

The previous chapter demonstrated that top management team (TMT) studies typically omit or draw causal inferences about team processes based on demographic proxies. To address this issue it is necessary to investigate a wider literature base. This chapter explores complementary literatures about managerial cognition, team processes, behavioural decision-making processes and decision belief with particular reference to TMTs. In particular, the relationships concerning four team processes, procedural rationality, frequency of team meetings, reflexivity and psychological safety are explored. All are espoused in the small group, decision-making or strategic management literatures as being influenced by demographic and cognitive variation, and being important for organisational performance. Like the previous chapter, this chapter finds that relatively little is actually known about TMT processes or decision-making. An important aspect of the latter is decision belief, that is to say, confidence, satisfaction and perceived effectiveness of the decision-makers. The chapter concludes by illustrating how the current study proposes to extend the existing knowledge about these aspects of top management team functioning.

3.1 Introduction

It was established in Chapter 2, that the reason demographic variation is considered to be important for TMTs is that it engenders cognitive variation amongst decision-makers. Furthermore, it is generally assumed that cognitive differences, that is to say, different beliefs, biases, filters, perspectives and opinions held by TMT members, are important for effective team decisions (Hodgkinson, 2001; Mohammed & Ringseis, 2001; Hambrick et al., 1996; Hambrick & Mason, 1984). However, this link has not been clearly established mainly because cognitive variation is typically not measured. Indeed, in a review of the major works on strategic decision processes, Das & Teng (1999) found that none had explicitly incorporated cognitive biases.

Small group researchers, in contrast, believe that cognitive variation assumed to be caused by demographic variation, has potentially deleterious effects on team functioning (Guzzo, 1982; 1996).

It was stated in the overview that this chapter would draw on a wider literature base than that concerned exclusively with 'upper echelons'. However, in keeping with the overarching aim of the thesis to stay true to what is known about TMT functioning, coverage of other literatures is selective rather than extensive. It is of interest to note here that despite over 500 citations of the original paper by Hambrick & Mason since its publication in 1984, there have only been three comprehensive reviews of the 'upper echelons' field (i.e. Jackson, 1992, Finkelstein & Hambrick, 1996 and Carpenter et al., 2004). Only Carpenter et al. (2004), in deference to Hambrick & Mason (1984) focuses exclusively on TMTs. Indeed, they argue that this is absolutely necessary, as with reference to the original paper (see also Chapter 2), 'upper echelons' theory exclusively concerns TMTs, not international teams, not teams per se, and not CEOs and their subordinates. That such an approach is valid, is borne out by Williams & O'Reilly (1998) who argue that many contra-indications as to findings concerning demographic variation, process and outcomes arise due to the fact that there are vast differences in types of organizational workgroups studied. This is further attested to by Flatt (1996) who argues that there are differences in outcomes for the same organisations when using the demographic variables of the board of directors and those of the TMT.

To the extent possible, pointers concerning cognition, decision belief and TMT processes arising from studies included in these three earlier reviews, together with the critique by Hodgkinson & Sparrow (2002), will provide the basis for the rest of the Chapter.

3.2 Managerial Cognition: Individual Executives Thinking Differently

For several decades, behavioural decision researchers investigated the correspondence between subjective and objective probability (Beach, 1997). That is to say, they calculated the expected value of a decision outcome to the decision maker based on subjective and objective assessments, typically presented as 'gain' or 'loss' and using experiments such as "book bags and poker chips". Findings tended to suggest that for repeated predictions in which subjects had the opportunity to revise their judgement, decision makers became increasingly more conservative with their progressive predictions (see for example, Peterson, Schneider & Miller, 1965). The conclusions basically were that decision makers' repeated judgements were neither accurate nor coherent, which in turn gave rise to studies of biases of decision makers and their habitual modes of thinking (heuristics) (Kahneman, Slovic & Tversky, 1982; Tversky & Kahneman, 1974; Tversky & Kahneman, 1982). With rare exception (see for example, Golden, 1992; Hodgkinson & Thomas, 1997), such work has been carried out in laboratory settings. It is rare indeed for heuristics research to be applied to strategic decision-making (see for example, Bateman & Zeitmahl, 1989) and almost unknown at the TMT level of analysis. Generally speaking, managerial cognition is conceptualised as a set of mental models by which managers either make retrospective sense of their environments or project prior events as being their proxy map of future reality (Huff, 1990).

More recently, a very small but growing body of research on shared cognition is beginning to examine how individuals entering a group decision-making context amalgamate their various viewpoints into a cognitive consensus, which does have a bearing on the current discussion concerning TMTs. The most interesting in terms of this discussion is that by Mohammed & Ringseis (2001) who measured cognitive consensus (similarity in viewpoints) pre- and post-discussion. This is an important contribution to the field because it recognises both the individual and team nature of decision making. That is to say, individuals have private opinions prior to a discussion, which need to be melded to form a team consensus, but the individuals within the team may also continue to have private opinions after a discussion. The limitation of this particular study is (a) that it involved 20 year old undergraduate students in role play; and (b), that the underlying assumptions or cognitive biases were confidentially provided to students prior to the cognitive consensus pre-discussion survey. In the context of their study, it was deemed important not to allow participants to develop their own viewpoints, in order to maximize the variability of cognitive diversity across the various groups in the simulation. Whilst Mohammed & Ringseis's (2001) study has considerably extended the boundaries of this type of research by measuring cognitive diversity and consensus, the artificial teams and the simulation (a farmers market) is of limited application to TMTs where individuals are likely to bring divergent viewpoints based on their immutable demographic traits. However, the conceptual issues it raises concerning pre- and post-discussion individual opinions and team consensus are of interest in TMT research, and are built upon in the present research.

With direct reference to TMTs, Glick et al.'s (1993) study singularly attempted to both define and measure cognitive diversity in TMTs. Cognitive diversity was defined at a broad level as *"referring to variation in beliefs about cause-effect relationships and variation in preferences about different goals for the organisation"*. It was measured in terms of preferences in terms of human resource goals, system maintenance goals, and profitability goals. Although not readily transferable to every study, Glick et al. (1993) demonstrate that definition and measurement can be attempted.

The same source data was re-analysed some years later by Chattopadhyay, Glick, Miller & Huber (1999) to determine whether functional conditioning (i.e. similarity in functional background and position) or social influence affected individuals' beliefs. This paper compared the beliefs of each individual with those of his peers and used age, tenure and functional background similarity (inversion of the Euclidean distance measure discussed in Chapter 2) as predictors. Similarity of beliefs was deemed important as a proxy for shared sense-making. Chattopadhyay et al. (1999) reported that age similarity and functional background similarity led to conformity whilst age dissimilarity and functional background dissimilarity manifested itself in disagreement and polarization. Contrary to expectation, the opposite was found concerning tenure similarity. In order to test the effects of social influence, which was defined as "social information processing, shared sense-making and other communication processes" (pp 763) Chattopadhyay et al. (1999) regressed each individuals' questionnaire responses onto the team mean response for the various questions concerning innovation, quality, bottom line etc. Like many other studies in this field (as discussed in Chapter 2 and again in Chapter 11), Chattopadhyay et al. (1999) obtained several null results when testing their hypotheses, and indeed the authors questioned the validity of accepting the supportive results with their tiny effect sizes. Notwithstanding any reservations they may have had, they put forward a plausible model concerning functional conditioning and social influence. However, as the authors did not actually measure 'social information processing, shared sense-making or other communication processes', they were obliged to put forward a possible alternative explanation as to how social influence affected individual executives' beliefs. First, Chattopadhyay et al. (1999) argued that corporate culture and sub-cultures within the TMT may arise due to common exposure to significant life-events or events concerning the organization or industry (they controlled for environmental turbulence amongst other items). Second, they argued that executives may be attracted to the organisation and selected to the TMT based on their similarity of viewpoints. However, as indicated above, and discussed by Chattopadhyay et al. (1999), their results were not consistent with these explanations. For example, if the argument holds, individuals who join the team at a similar time will hold similar beliefs compared to cohorts who join at other times based on their common exposure to organisational life events, however, the results were that tenure similar individuals hold the most disparate opinions.

What Chattopadhyay et al.'s (1999) study really shows is the complexity regarding managerial cognition and how to understand the determinants of individuals' beliefs. Although eight control measures were included (such as environmental turbulence, munificence, autonomy and organization size), as the authors admitted, *"we obviously could not control for all conceivable characteristics of their common experiences"* (Chattopadhyay et al., 1999, pp 783).

Clearly, the dearth of studies available for review regarding TMT cognition, indicates that there is still considerable work to be done in this area. Moreover, there is a growing recognition that cognition is fluid rather than static. That is to say, individuals' judgements are continuously being affected by changing aspects of organizational life and other emergent processes (Huff, 1997), although what form such processes may take remains largely unexplored. Furthermore, there is considerable debate as to whether macro-processes such as market-place conditions, and perceptions of competitor threats are more or less important than so-called micro-processes, such as strategy formulation and implementation (Hodgkinson & Sparrow, 2002). As will be discussed in the next section, although these factors may shape individuals' and teams' perceptions and actions, they are not typically viewed as processes of interest in 'upper echelons' research.

TMT decision-making is also of interest in terms of the strategic management literature. This is because with increased emphasis on transparency, accountability and governance processes, business and civil society are starting to question the role of top management and their legal and moral responsibilities (Petzall, Selvarajah & Willis, 1991). In recent years such a focus has particularly been directed at the multi-national corporations (Adler, 1997) and much work has been done in the international management team arena to understand the TMT decision-making as it relates to devising strategic or corporate plans (Higgins, 1980).

The strategic management literature has typically investigated issues such as competitor categorization processes amongst strategic groups within particular industries in response to competitor behaviour (Hodgkinson & Sparrow, 2002). Superficially, this may seem the same as 'strategic choices' made in the context of 'constructed realities' formed in response to 'strategic situations' (Finkelstein & Hambrick, 1996). The latter being the definitive extension of 'upper echelons' put forward by Finkelstein & Hambrick (1996) concerning TMTs. With regard to the former, the focus of interest is the extent to which competitive structures (i.e. dominant coalitions or clusters of firms) form around shared intra-industry cognitions and the demise of rivals that fail to interpret environmental and organisational stimuli in the same way (Hodgkinson & Sparrow, 2002). In spite of the level of analysis being different, some of the basic concepts are somewhat similar. Both perspectives assume the strategic situation (Finkelstein & Hambrick, 1996) or macro environment (Hodgkinson & Sparrow, 2002) to include all external (e.g. economic factors, industry position, regulatory forces) and internal stimuli (e.g. history, structure, diversification posture etc). In each perspective a strategic decision is required in response to such stimuli. In the 'upper echelons' perspective at the TMT level of analysis, demographic factors are used as proxies for observable experiences and psychological factors such as values, cognitive models and other personality factors (Finkelstein & Hambrick, 1996). These meld into a limited field of vision which affects strategic perception and interpretation of the original stimuli. A perceived reality is constructed of these pieces, such that a strategic choice is made which in turn impacts upon organisational performance. In the strategic management perspective, at the cluster of firms level of analysis, more emphasis is given to how dominant coalitions form, based upon individual firms' perceptions and imitation of industry leaders. Response over time to similar circumstances results in social learning. On the other hand, cognitive inertia can result when organisations are intransigent concerning their established viewpoints concerning the imitation of acknowledged industry leaders who may be losing ground in unfavourable markets (see Hodgkinson & Sparrow, 2002 for a review).

Thus whilst no less important, the 'upper echelons' perspective can be considered to be concerned with the micro-processes of decision-making within particular TMTs, the strategic management perspective is primarily concerned with the macro-processes, or those of dominant players in particular industries. Whilst the similarities of the two perspectives have been outlined above, the level of analysis makes some of the concepts less important than others. Demographic factors are not so important in the strategic management perspective. On the other hand, demographic variation at the TMT level is argued by Hodgkinson & Sparrow (2002) to reduce any tendency towards cognitive inertia. This is not to say that the two perspectives are completely distinct, as the concepts of team mental models, distributed cognition and causal maps have been applied at the team level in organisations (see for example, Markoczy, 1997).

At the micro level, when decisions are studied in real-world situations, it is observed that individuals think a little, and then evaluate the outcomes and think and act some more (Connelly & Wagner, 1988). This is particularly true of individual top managers as they continuously fine tune their understanding of organisational performance in line with industry and environmental considerations (Hrebiniak & Snow, 1982). Without doubt, the interactive experience of engaging in a team decision must of itself influence individuals' cognitive perceptions as an emergent phenomenon (Chattopadhyay et al., 1999). That there needs to be a balance between studies at the individual and team levels is illustrated in the following quote:

"Thus far, very few studies of executive cognition have gone beyond the individual level of analysis. To the extent that executives' individual perceptions and beliefs actually form a key element of the management decision process at the team level, it makes sense to continue exploring the determinants of strategic cognition" (Hodgkinson, 2001, pp 426).

To this end, the next section explores decision-making processes with particular relevance to TMTs.

3.3 Overview of Major Decision-Making Theories Concerning TMTs

Organizational theorists and psychologists tend to derive their understanding of team decision-making from the group problem solving literature. Guzzo (1982) acknowledged the patchy and fragmented understanding of decision-making in groups and gave three reasons for this:

(1) several disciplines (such as economics, management and psychology) have established theories, vocabularies, assumptions and research methods for studying group decision-making, but these studies are not interdisciplinary;

(2) studies are often not comparable because of their limited focus on single aspects of decision-making (e.g. idea generation, leadership, power, or stressfulness of decision-making);

(3) the research settings are so specific (governmental agencies, juries or business organisations) as not to be generalisable and applicable to other settings.

Generally, group problem-solving research has been conducted almost exclusively in laboratory settings and the extrapolation and application of findings to real-world environments is "fraught with danger" (Hoffman, 1979, pp 368). Furthermore, by far the most of this collective research has sought to evaluate single decision choices made by organizations. This focus on simple decision events has been lamented by a new generation of decision researchers (Oransu & Connelly, 1993), who claim that the multiplicity of complexities impacting upon decision-makers (such as ill-defined goals, responsibility for irreversible decisions, time pressures, dilemmas, constantly changing environments and multiple stakeholders) require observational and descriptive research as opposed to prescriptive and rational models. Moreover, the group problem-solving literature is particularly limited for understanding decision-making processes in natural settings, because it unrealistically assumes that all things are equal, such as, the investment of decision-makers, available resources, and the allocation of effort to all parts of the decision-making process.

As a result of these shifts in social science research into decision-making in organisations, there has been a definite outgrowth from economic and behavioural decision research to that which is now known as naturalistic decision-making theory. Indeed, the emphasis by previous studies on problem-solving it is claimed, has been unhelpful, in that the great richness afforded by studying various types of decision situations and processes has yet to be realised (Mintzberg, Raisinghani & Theoret, 1976). Naturalistic decision-making, on the other hand, represents an emerging body of work (begun in the early 1990's), which embraces a variety of observational and behavioural research approaches to describe and define what decision-makers actually do in practice. Unfortunately, none of this literature addresses TMTs specifically. This chapter now addresses three models from the organisational structural approach that are supposed to have direct relevance to decision-making in TMTs.

3.3.1 Satisficing and Structural Compensation

The first description of the decision-making process is firmly grounded in the economic and behavioural decision research paradigm. Indeed, all theories and models developed since can in some ways trace their roots back to a series of seminal works by Simon in the 1950s and 60s (Bazerman, 1998). The influence of time and cost factors on the decision process was perhaps first recognised by Simon (1957; 1961), who proposed what is now sometimes known as 'bounded rationality'. Simon suggested that the marginal cost (time, money etc) of an extended search for alternatives may outweigh the marginal benefit of finding the optimal solution, causing decision makers to take a satisfactory option rather than a superior option. Satisficing is a term which was first coined by Simon (1957) to describe a process of decision-making in which sees decisions made when a sufficient solution as opposed to the optimum solution is found.

Simon argued that while time and cost factors may inhibit the search for information in some situations, in others complete information may be available but not utilised because of limited cognitive information processing capabilities of decision makers. This is borne out by other researchers (Beach, 1997) who also find that decision-makers often have difficulty comprehending the problems facing them, and hence tend to simplify things in order to deal with them.

Satisficing is proposed by Simon as a series of suboptimal yet satisfactory decisions which could still lead to an optimum solution over a period of time. Similarly, Lindblom (1959), argued that decision-makers are more likely to make incremental decisions (those that do not depart significantly from the status quo), owing to the impracticality of assigning attributes to, and evaluating the costs and benefits of greatly differing alternatives.

The reduced quality of decisions obtained by suboptimal and incremental decisions was thought to be better than prolonged inaction due to the search for optimality. Satisficing has intuitive appeal to decision makers who are in danger of suffering information overload, and endorses March & Simon's (1958) observation, that decision makers attend to the immediate, specific, operational and doable and ignore the distant, general and difficult to translate into action.

Perhaps not surprisingly 'Satisficing' is criticised as being too limiting (Etzioni, 1967) because decision makers may become comfortable with routine, incremental decision-making so that they miss or waste opportunities to make innovative decisions. Simon's (1957) classic work ushered in a new era of decision research in organizations and while it departed from microeconomic treatments, it still adhered to some of the rational concepts of the decision process such as those suggested by Dewey (1933). It was a revolutionary approach because it suggested that organizational structures (both formal and informal) are the basis for understanding processes such as decision-making which contribute to organizational effectiveness and act as compensatory mechanisms to cater for the limited information process capabilities of decision makers. How so? It proposed that an organization be structured so that decision problems and decision makers with appropriate capabilities, responsibilities and resources (including time for decision activity) are allocated to decision choices (March & Simon, 1958; Cohen, March & Olsen, 1972; Cohen, March & Olsen, 1976).

3.3.2 The Garbage Can Model

Another model, with a structural orientation is the 'garbage can model' of managerial decision making. It suggests that organizations are 'organized anarchies' of ambiguous choice opportunities, problems, solutions and participants (decision makers), and that decisions are made as an interpretation of several relatively independent factors or garbage "streams" (Cohen et al., 1976, pp 26).

Critical to the 'garbage can model' are the windows of time in which choice opportunities become defined by problems, solutions and the energies of decision makers to achieve a satisfactory outcome. The model suggests that the connections between sets of participants and garbage streams at any given time are unpredictable, and hence it is descriptive of managerial decision making.

Commentators have critiqued the model as too chaotic and unrealistic as a general model (Hickson, Butler, Cray, Mallory & Wilson 1986; Beach, 1997) and have further suggested that because the emotive imagery "of rollicking in a dented dustbin is fun" (Hickson et al., 1986, pp 2), the model has received greater profile and credibility than it deserves (Beach, 1997).

3.3.3 The Top Decisions Model

The 'top decisions' model (Hickson et al., 1986), asserts that an organization is set up and sustained by a dominant coalition of stakeholders who provide the 'rules of the game' for decision making. Like the previous models, 'top decisions' suggests that the organization is the framework for understanding decision making. It is primarily concerned with strategic decision-making (such as new product launches, organizational restructuring and takeover decisions), and defines complexity on four dimensions - rarity (frequency of similar previous decisions), consequentiality (radicalness, seriousness of contingencies, time frame), precursiveness (extent to which this decisions constrains later decisions), and involvement (number of participants in decision activity). Central to this model is politicality, and when linked to complexity, types of decision process can be retrospectively identified (Hickson et al., 1986).

3.3.4 Summary of Relevance of Models Considered

Each of the models just considered in this section is asserted to be a form of managerial or organizational decision-making and as such informs the present discussion concerning TMT decision making. However, each of these models still falls short of describing the complexities involved in most decision events. Moreover, their focus on final decision outcomes fails to account for the dynamic and creative process of exploring, evaluating, generating and modifying alternatives which characterises human decision-making processes (Svenson, 1979; 1992; Montgomery, 1993; and Orasanu & Connolly, 1993).

It is apparent from this literature review that the naturalistic models whilst usefully descriptive, either tend to be context specific, or are not readily observable or measurable. It is perhaps hardly surprising then that researchers still gravitate back to variations of Simon's (1957) bounded rationality as a way of simplifying and explaining the complexity inherent in strategic decision-making (Schwenk, 1984). Just how this is conceptualised as a team process (i.e. procedural rationality) is covered in the section on team processes below. Before moving on to team processes, however, it is appropriate to discuss one final aspect of decision-making, which concerns the accuracy or quality of decisions.

3.4 Decision Quality Vs. Decision Belief: The Measurement Debate

Decision quality is a ubiquitous term in the decision-making literature, but it actually proves to be a slippery one. In many cases, it is used in the context of achieving an a priori solution to a problem pre-determined as the correct one by the researchers (Gigone & Hastie, 1997). Thus, subjects' degree of achievement of the correct solution is scaled as decision quality (see for example, Heath & Gonzalez, 1995). Furthermore, if decision-makers arrive at an incorrect decision, researchers tend to conclude that it is human judgment rather than the 'expert solution' that is flawed (Maule & Hodgkinson, 2003). Setting an a priori solution is a research technique which is suitable for artificial groups using a set problem-solution task with a clear and single correct answer. Obviously, it is virtually impossible to use the same kind of measure at the strategic decision-making level, as there is no correct solution to many of the organisation structure issues studied. Hence a much tighter calibration of human judgement (Maule & Hodgkinson, 2003) as to whether a decision was the 'right' one must be used.

In other cases the term is used to ask decision makers if the quality of the decision they made was good (see for example, Schweiger et al., 1986; Amason, 1996; Dooley & Fryxell, 1999). The problem with 'quality' is that it is a subjective term, and what might be perceived as high or low quality by one person may not be so to another (Swift, Humphrey & Gor, 2000). Quality in terms of goods and services essentially means fitness for purpose (Swift et al., 2000), but this does not translate particularly well to slightly more esoteric concepts such as decisions. Hence the term 'decision quality' as it is currently used in the TMT literature is problematic, yet the intent to understand whether decision-makers believe a decision to be the right one for the organisation is pertinent. Therefore, it is perhaps more apt to talk about decision belief, that is to say, what decision makers believe to be true regarding their decision.

It is widely recognised that the experts in strategic decision-making (that is, those who can vouch for the validity of outcomes), are actually the decision-makers themselves (Dean & Sharfman, 1996) rather than researchers assessing outcomes according to theoretical prescriptions (Mintzberg et al., 1976). Indeed, it is thought that "*the best way to gauge the quality of an individual strategic decision is to ask those who have observed its effects and who understand its context to judge, retrospectively and on several dimensions, how the decision turned out*" (Amason, 1996, pp 134).

However, there are problems associated with asking decision-makers about their decisions. Not the least of these being that managers are sometimes mistaken, and/or have inaccurate recollections based on erroneous facts (Mezias & Starbuck, 2003; Hrebiniak & Snow, 1982). Recollection of decision-making processes may also become distorted over time depending on whether the decision was implemented or whether the consequences were good or bad for the organisation (Nisbett & Wilson, 1977; Hambrick, 1981; Schwenk, 1985; Golden, 1992).

Three factors have been identified that need to be taken into account when asking decision-makers to judge their decisions (Huber & Power, 1985). The first of these is that multiple respondents should be used (Golden, 1992; Dean & Sharfman, 1996). This means not relying on one person such as the CEO for example, who will only give one idiosyncratic perspective. The second feature is to ensure that the respondents are very close to the decision being reviewed. This is because the best informants about managerial decisions are those who are steeped in the process, who understand the substance of the decision, who have no secrets to hide and who do not stand to benefit from lying (Winter, 2003). The third factor is that the judgement of the decision should be measured in its own right and directly after the decision is made in order to mitigate any exogenous effects (Dean & Sharfman, 1996).

In measuring managers' beliefs about their decisions, three aspects are important. These are confidence that the decision was the right one (Heath & Gonzalez, 1995; Sniezek, 1992), satisfaction with the decision process used (Schweiger, Sandberg & Ragan, 1986) and perceived effectiveness of the team in working together (Edmondson, Roberto & Watkins, 2003; Cohen & Bailey, 1997). All of these concepts have been variously tested in laboratory groups. The rationale for their importance is that confidence and satisfaction and perceived effectiveness will engender more commitment amongst participants to follow through and ensure that the decision will be implemented (Sniezek & Henry, 1990; Schweiger et al., 1986). Moreover, it is argued that if TMTs are not confident they have made the right decision, are dissatisfied with the process, and do not believe they are functioning effectively as a group, it is expected that they will not continue working together as a unit for the common good of the organisation (Nadler, Hackman & Lawler, 1979).

Amason (1996), in a study of 48 TMTs, posed three questions concerning decisions that are broadly similar in ideological intent to the three concepts noted above. He found that friction and tension in TMTs concerning a decision (affective conflict) were negatively associated with belief in the decision. On the other hand, individual differences of opinion within the TMT, about the decision were positively associated with decision belief, suggesting that the more alternatives available for consideration meant a positive perception of the decision process.

The works cited above with regard to decision belief do not include links to demographics. Outside the TMT literature, there is some evidence to suggest that men in a minority in mixed gender teams experience less satisfaction than women, but that this is far from conclusive (Wilson, 2003). Within both the TMT literature and relational demography literature more generally, there are many vague assertions that demographic variation (particularly age, gender, racial/ethnic background and tenure) is associated with 'process losses', which in turn leads to less confidence, satisfaction and effectiveness (Milliken & Martin, 1996). This lack particularly of satisfaction, is said to lead to higher rates of team turnover (Jackson et al., 1991) and to the root cause "homosocial reproduction" (Kanter, 1977). That is to say, because demographically dissimilar individuals are the least satisfied, management teams and organizations will seek to drive out dissimilarity and diversity and replace it with people who are more similar, so that they will fit in more (Ely, 1994; Milliken & Martin, 1996).

Clearly, there is still much more work to be done in investigating these aspects of decision belief in the particular contextual domain of TMTs and with particular reference to demographic variation. The present study aims to measure decision belief. The next section investigates what is known about team processes.

3.5 The Empty Black Box: Ignorance of TMT Processes

The 'upper echelons' and 'organisational demography' traditions reviewed in Chapter 2, assume that cognitive variation leads to better decision-making and thereby better organisational performance. However, the lack of investigation of processes has severely hampered 'upper echelons' research (Pettigrew, 1992; Lawrence, 1997; Carpenter et al., 2004). Although recent studies have made some progress in this area, the processes studied within TMTs are (as the review by Hodgkinson & Sparrow, 2002 attests): procedural rationality (also called decision comprehensiveness) (Simons et al., 1999; Dean & Sharfman, 1996); conflict (Amason, 1996); debate (Simons et al., 1999); communication and social integration (O'Reilly et al., 1993; Smith et al., 1994; Knight et al., 1999); innovation (O'Reilly et al., 1993; Bantel & Jackson, 1992) and strategic consensus (Smith et al., 1994; Knight et al., 1999).

In a highly regarded review by Williams & O'Reilly (1998) which covered four decades of demographic variation research at various levels in organisations, it was observed that *"group process is most frequently investigated in terms of three primary dimensions: social integration, communication and conflict"* (pp 91). Clearly there is a very wide area still to cover with regard to process. The 1996 review by Milliken & Martin could only find three TMT studies that had investigated process. Almost a decade later, Carpenter et al. (2004) were similarly struggling to find TMT studies with empirical demographic-process linkages, only being able to review a few more studies.

The assumption that intervening processes account for the relationships between demographics and performance outcomes, whilst these remain typically unmeasured, has proved to be one of the most contentious criticisms of 'upper echelons' research (Priem et al., 1999). For example, Lawrence (1997) refers to this as a 'congruence assumption' (that is to say, because A leads to B, it must be caused by C). She argues that such assumptions are invalid, observing: *'Because the final test of an intervening process explanation requires measuring the subjective concept, it is not possible to provide evaluation criteria for such explanations under the 'congruence assumption''* (pp 10). This topic and the Lawrence (1997) critique in general will be returned to in much greater depth in Chapter 11. Suffice it to say, the links between demographic variation and team processes are feeble and typically unsubstantiated. Furthermore, that demographic proxies are supposed to capture the social processes inherent in real TMTs, is deemed to be one of the "intractable problems" facing 'upper echelons' research (Priem et al., 1999, pp 943).

Whilst some reviews of the field seek to downplay the Lawrence (1997) critique by suggesting that it is "forced" (Williams & O'Reilly, 1998, pp 98), others are sensitive to fact that many of the processes supposedly established under the rather tenuous 'congruence assumption have not actually been tested (Priem et al., 1999). For example, Milliken & Martins' (1996) review adds qualifiers to the reported findings concerning intervening processes between demographics and outcomes, such as, "presumably because of" (pp 410), and "appeared to operate through" (pp 412).

Small group researchers have a lot to say about team processes and decision-making and how some processes are positive and some are negative. Negative processes are those that either impede attainment of consensus such as conflict, or that act as conflict avoidance strategies that cause the team to suspend rational judgment, such as groupthink (Janis, 1972).

On the other hand, some team processes such as communication, social integration and procedural rationality (Collins & Guetzkow, 1964; Murray, 1989; Glick et al., 1993) are thought to act positively as mechanisms for achieving consensus and uniformity of opinion, (Isabella & Waddock, 1994). This is by no means a comprehensive list, and considerable disagreement exists as to which processes are good and which are bad. For example, some argue that cohesion (which they perceive to be similar in manifestation to groupthink) is negative, whereas conflict is argued to be positive because of the opportunity to air divergent viewpoints (Kanter, 1983; Lawrence, 1997). Others cast cohesion as a positive attribute, one which sees individuals in groups actively working together for the strategic good rather than individual competition (Sniezek & Henry, 1989). At one level, the differences between how processes are viewed are purely semantic. At another, understanding of team processes is impeded through such inconsistency. Of interest to this discussion is the tension between variation (demographic and cognitive) and top management team processes, and how these influence decision-making. Moreover, other research suggests that processes may not simply characterise a whole team, but that sub-cultures with conflicting processes may exist within teams (Hodgkinson & Sparrow, 2002). So-called demographic faultlines can exist in teams, which means that demographic factors or job similarities cause individuals to cluster together within a team and take a conflicting perspective to other mini-groups concerning a particular task (Lau & Murnighan, 1998)

The following sub-sections will review four processes deemed to be important (both positive and negative), for TMT decision-making, that will feature in the empirical research to be presented in later chapters.

3.5.1 Procedural Rationality

It is widely held that decisions with important strategic consequences such as those typically made by TMTs, tend to be based on rational/comprehensive procedures (Dean & Sharfman, 1996; Miller, Burke & Glick, 1998; Simons, Pelled & Smith, 1999; Papadakis & Barwise, 2002). In order to learn something of the processes inherent in TMT decision-making, one must turn to the strategic decision-making literature, which is based on two major perspectives. The first, presented by Weick (1979) argues that managers' create organization structures consistent with their perceptions of the environment they want to operate in. Hence, it can be said that they create, that is to say, they make strategic choices about, their operating environments. The second perspective is known as resource dependence theory (Salancik & Pfeffer, 1977). It is extremely similar to the first, but posits that managers create organization structures that react to the environment in which they find themselves. By far the largest body of strategic decision-making literature is actually about strategic planning around organization structure, and has little relevance to TMT decision-making (Dean & Sharfman, 1996). Taking the literature as a whole, there are three recognised theoretical orientations and resulting research streams (Hendry, 2000). The first is the traditional rational perspective which is inherently assumed in the 'upper echelons' tradition (Hodgkinson, 2001). The second is the action perspective which asserts that strategic actions do not always arise as a result of strategic choice. Therefore, strategic actions may precede the decisions by which they are justified. The third perspective on strategic decision-making is the interpretative approach, which is essentially, a sense-making exercise. That is to say, decision-makers structure their images of reality in order to function in their environments (Sparrow, 1994).

The least problematic of these, in terms of being consistent with decision-makers' experience, is the traditional rational perspective (Hendry, 2000) and procedural rationality is the most studied decision-process in 'upper echelons' research (Papadakis & Barwise, 2002). This is because, whether from the traditional rational or the action perspective of strategic decision-making, researchers agree that strategic decisions are likely to involve fairly complete information and this is argued to result from exploring, investigating and disseminating available information (Mintzberg et al., 1976).

Procedural rationality is defined as "*the extent to which the decision process involves the collection of information relevant to the decision and the reliance upon analysis of this information in making the choice*" (Dean & Sharfman, 1996, pp 373). Procedural rationality (Eisenhardt & Bourgeois, 1988) is similar to Fredrickson's (1984) dimension of comprehensiveness, which is defined as "*the extent to which organizations attempt to be exhaustive or inclusive in making and integrating strategic decisions*" (pp 445). Procedural rationality was first conceived of as a feature of TMT decision-making by Eisenhardt & Bourgeois (1988), and was developed into 5 item self response scale (Dean & Sharfman, 1996) in which the items were derived from a vast literature review of the strategic decision-making literature, including studies cited above. Unfortunately, with rare exception (see Papadakis & Barwise, 2002; and Dean & Sharfman, 1996; Simons, Pelled & Smith, 1999) later studies (see for example Kilduff, Angelmar & Mehra, 2000; and Mohammed & Ringseis, 2001) are conducted exclusively with student groups assuming the identity of TMTs in role play.

Those that have been conducted in real TMTs, for example, in a study of 38 Greek firms involving 70 strategic decisions found that TMTs tend to use procedural rationality more when the stakes of the decision for the organisation are high (Papadakis & Barwise, 2002). Moreover, TMTs that had used procedural rationality were found to have been more effective in achieving their decision objectives (when revisited by the researchers 2 years later) than those that used intuitive techniques (Dean & Sharfman, 1996).

Simons et al. (1999) studied demographic variation in relation to procedural rationality (which they termed as decision comprehensiveness) and change in profitability and sales in 57 TMTs. They found that more lively debate was positively associated with decision comprehensiveness, ensuring that TMT decisions considered a wide range of options. Furthermore, decision comprehensiveness was positively associated with increased profitability, and increased sales.

With regard to the link between this intervening process and demographic variation, Papadakis & Barwise (2002) found that TMT education (central tendency towards similarity) was positively associated with procedural rationality. They argued that a *"well-educated TMT may thus be efficient enough to reach an objectively better solution, thus mitigating political processes"* (Papadakis & Barwise, 2002, pp 85). This is an interesting observation, as it is generally argued in the literature that decision comprehensiveness is a mechanism for harnessing and mitigating the negative effects of cognitive variation afforded by demographic variation through debating all the alternatives. Debate in TMTs, exacerbated by educational diversity (team level differences), enhances organisational performance, which in turn is moderated by the process of procedural rationality (Simons et al., 1999). Interestingly, educational diversity by itself was not associated with performance. Thus it is fair to say that demographic variation (diversity and dissimilarity rather than central tendency) is negatively associated with procedural rationality. However, as procedural rationality is supposed to reduce cognitive variation by providing an apolitical mechanism for exploring all alternatives, one would expect greater consensus to occur in teams characterised by procedural rationality. Furthermore, as a tendency towards agreement is associated with higher levels of satisfaction and confidence (Sniezek & Henry, 1990), one could expect a positive relationship between procedural rationality and decision belief.

Clearly, much more work is needed within the domain of real TMTs with regard to procedural rationality, especially with regard to investigating the relationships with demographic and cognitive variation and decision belief as an outcome. However, as this is a process first conceived and devised specifically as being relevant to TMTs, it is an important investigative theme in this study.

The next section explores another feature of TMT dynamics, which is supposed to have similar effects, frequency of team meetings.

3.5.2 Frequency of Team Meetings

In order to make any collective decision TMT members must communicate with one another. Most scholars suggest that communication is essential for high quality decisions (Hirokawa & Pace, 1983; Hirokawa, 1990) as it is the "*means by which group members attempt to meet the requisites for successful group decision-making*" (Gouran & Hirokawa, 1983, pp 170). Johnson & Johnson (1987) found that communication, particularly in the form of verbalisation & re-iteration, increased comprehension, understanding and retention of the content of discussion, which they suggest promotes effective performance. Much of the work about communication is conducted in laboratory settings with synthetic or ad hoc groups so that researchers can count types of verbal exchange such as 'interruption', 'repetition', and 'seeks clarification', using coding guides such as Bales (1949).

There can be no doubt that communication has to be central to a process where a group of individuals come together to reach a collective decision. Findings of previous research suggest that it is how divergent viewpoints are assimilated into a consensus, through information sharing and collective commitment to decision-outcomes (Heath & Gonzalez, 1995). However, the efficacy of using counts of verbal exchanges to define and measure communication is dubious. What researchers are really trying to get at is an understanding of "*how members interact during the strategic decision-making process, because it is expected that this has a potent effect on the quality of their decisions, and on how well and how quickly those decisions are implemented*" (Dooley & Fryxell, 1999, pp 389).

Glick et al. (1993) in a survey of 79 TMTs counted frequency of informal and formal meetings amongst team members as a proxy for communication, as did O'Reilly et al. (1993). Their reasoning was that frequency of meetings tends to socialise the team into procedural norms, so that the members tend to make decisions in a consistent manner. Over time, a TMT establishes a regular pattern of making decisions together, and members behave in a regular and consistent fashion when engaged in the strategic decision-making process. Other researchers too have studied variations of the frequency of interaction in TMTs as a means of measuring how TMTs work together (see for example, Papadakis & Barwise, 2002; Eisenhardt & Schoonhoven, 1990; and Smith et al., 1994).

Eisenhardt & Schoonhoven (1990) found that the more members of a startup TMT had worked with each other previously, the more focused and speedily decisions were made. Their conclusion was that prior experience of working together, and being comfortable with each others' ways of working, was beneficial for TMTs making strategic decisions in uncertain environments. Papadakis & Barwise (2002) measured frequency of participation by TMT members in decision-making and found that the more important the decision, the more members that would be involved at all stages of the decision. More meetings amongst TMTs were found by Glick et al. (1993) to be associated with more comprehensive decision-making. Smith et al. (1994) were surprised to find a negative relationship between frequency of meetings and organisational performance. They concluded that team meetings are a mixed blessing, vital for team members to share information and hence make more comprehensive decisions, but meetings absorb valuable time that could be spent instead on task oriented activities.

Notwithstanding the one adverse finding from Smith et al. (1994), the thrust of the research to date is that frequency of team meetings is a good index of TMT communication and should predict decision belief.

With regard to the supposed link between frequency of team meetings and demographic variation, Smith et al. (1994) found that tenure diversity was negatively related to the amount of informal communication within TMTs. O'Reilly et al. (1993) found that more tenure diversity meant less communication amongst TMTs, from which the authors argued that tenure homogeneous teams may have more open communication and less distortion of messages. With regard to functional diversity, Glick et al. (1993) found that functional diversity had positive effect on frequency of communication within TMTs, whilst Smith et al. (1994) found no effects of functional diversity on communication. Across the wider literature, functional diversity is widely held to be beneficial for bringing different perspectives, knowledge and skill sets, but detrimental for within group functioning (Williams & O'Reilly, 1998). Also with regard to the non TMT literature, it has been found that age diversity (team level) is negatively associated with frequency of meetings, but that employees similar in age tend to communicate more frequently (Zenger & Lawrence, 1989). The non TMT literature suggests that gender diversity may be negatively associated with frequency of team meetings although this has not been consistently proved (Williams & O'Reilly, 1998). This latter argument stems from the observation that women can be excluded from male communication networks in some instances (Ibarra, 1992), but that such findings are not always replicated, and that the proportion of men and women present in a team may influence such results (Williams & O'Reilly, 1998). It is probably fair to say that the common thread from these reported studies is that demographic variation is likely to be negatively associated with frequency of team meetings, and with cognitive variation, that is to say, positively associated with consensus. Furthermore, as consensus is associated with greater satisfaction, confidence and effectiveness (Sniezek & Henry, 1990) one might conclude that frequency of team meetings will be positively associated with decision belief.

The next section turns to a discussion of a process that is argued to have contra effects to those just discussed.

3.5.3 Reflexivity

The presence of conflicting viewpoints is generally thought to be beneficial for generating alternatives for consideration. However, teams need a mechanism for articulating all the alternatives and for deciding whether they have chosen the right one. This has generally been referred to as functional conflict (Amason, 1996), constructive controversy (Tjosvold, 1992) or reflexivity (West, 1996).

Tjosvold (1992) first described a naturalistic concept embedded in teamwork which he called 'constructive controversy' - a process by which opposing viewpoints (conflicting interests) of members are explored. Moreover, he argued for a direct causal relationship with effectiveness. In teams which exhibit constructive controversy, he observed, there is a climate of mutual co-operation and trust as opposed to a climate of competition and mutual distrust, and critical review of ideas is viewed as a healthy constructive process rather than a destructive and aggressive process. The task orientation of such teams is to achieve an excellent outcome rather than simply achieving consensus and avoiding conflict that may result in a suboptimal task outcome. Tjosvold goes on to describe the dynamics of constructive controversy and its relationship as a construct to team decision-making in this way:

"Controversy, when discussed in a co-operative context, promotes elaboration of views, the search for new information and ideas, and the integration of apparently opposing positions. This controversy copes with the biases of closed-mindedness, inadequate evaluation of new information, simplifying the problem, and unwarranted confidence in initial positions. These processes in turn result in understanding opposing positions and the problem, development of alternatives, adoption of and commitment to high-quality solutions". Tjosvold, 1992, pp 172.

Further theory development and research suggests that constructive controversy, as evidenced by groups fully exploring opposing opinions, analysing task related objectives, strategies and processes in advance improves the quality of decision-making (Tjosvold, 1995; Tjosvold, Yu & Hui, 2004).

Constructive controversy, with particular reference to top management teams, has been conceptualised and referred to as reflexivity (West, 1996; West, Carletta & Garrod, 1997). It is defined as *"the extent to which group members overtly reflect upon the group's objectives, strategies and processes, and adapt them to current or anticipated endogenous or environmental circumstances"* (West, 1996, pp 559).

Reflexivity is said also to be heavily based upon Schön's (1983) work concerning the way in which practitioners reflect upon what they are doing while they are doing it ('reflection-in-action') or retrospectively reflect upon what they have done ('reflection-on-action'). Practitioner reflection is primarily concerned with the ability to react flexibly to exogenous circumstances and to learn by experience for future practice.

Schön (1983) believed that technical rational prescriptions were very apparent in the problem-solving literature, which generally asserts logical, sequential and linear approaches to solving problems, and therefore allows no flexibility for dealing with uncertain, dynamic and challenging environments. Such a viewpoint finds resonance with contemporary decision-making theorists who note that the inflexibility of technical rational problem-solving models, tends to label decision-makers (practitioners) as inconsistent and ineffective problem solvers if they don't perform in the prescribed way (Rasmussen, 1993). Schön (1983) asserted that this omission was due to the emphasis of the technical rational perspective on problem-solving to the neglect of problem setting, thereby negating choice, because in the real-world problems are not givens - they emerge from puzzling, complex uncertain circumstances and factors. Like other decision-making theorists (e.g. Orasanu & Connelly, 1993), Schön (1983) recognised that, in practice, decision-makers 'frame and re-frame problems' in response to constantly changing circumstances. In fact, he asserted that practitioners come to understand unique and uncertain situations through their attempts to change them (pp 132).

With regard to the links between demographic variation and reflexivity, one needs to consider effects on substantive or task conflict as this is the closest conceptually to reflexivity. The non TMT literature suggests that age diversity leads to less rather than more conflict, in organisations low in the organizational hierarchy, but higher up, there are no reported effects of age diversity and substantive conflict in managerial networks (Williams & O'Reilly, 1998). The TMT literature suggests that tenure diversity may be associated with higher levels of conflict (O'Reilly et al., 1993), an argument that is consistently supported by findings from a range of different types of non management teams (Williams & O'Reilly, 1998). With regard to gender diversity, again from the non TMT literature, there are consistently no effects with reference to substantive conflict (Williams & O'Reilly, 1998; Wilson, 2003).

Drawing together the common threads with regard to demographics and processes similar in intent to reflexivity, it appears that demographic variation is very likely to be negatively associated with reflexivity. There are convincing arguments that reflexivity exacerbates constructive debate by drawing out conflicting viewpoints and opinions into open discussion. Thus one could expect that reflexivity would be a facilitator of cognitive variation. Moreover, as differences of opinions are linked to reduced levels of confidence, satisfaction and effectiveness (Sniezek & Henry, 1990), it is fair to assume that reflexivity would be negatively associated with decision belief.

The chapter now turns to a discussion of psychological safety.

3.5.4 Psychological Safety

Reflexivity involves being able to verbalise and deal with conflicting viewpoints. For this to operate, individuals need to feel psychologically safe. Thus, psychological safety is defined as the team being safe for interpersonal risk taking (Edmondson, 1998). It is not the same as group cohesiveness or 'groupthink' (Janis, 1972), which concept refers to an unwillingness to challenges others' viewpoints.

Psychological safety is characterised by all team members sharing in influence over decisions, frequent interaction, information sharing and listening to all opinions even minority viewpoints (West & Anderson, 1996; Anderson & West, 1994). It is argued that psychological safety, that is shared trustworthiness, within TMTs is a key element in synthesizing dissent and consensus (Dooley & Fryxell, 1999). It is said to lead to greater commitment on behalf of individuals to decision-outcomes (Lawler & Hackman, 1969), greater idea generation (Kanter, 1983; King, Anderson & West, 1992), greater consensus amongst team members (Sniezek & Henry, 1990) and a belief in shared group goals (Dooley & Fryxell, 1999). Edmondson (1998) argues that for psychological safety to be a team-level construct, most members must perceive that it is acceptable to make a mistake in the team, based on previous experiences for themselves or others where this has been the case.

Psychological safety is traditionally studied in workgroups at various levels in the organisational hierarchy well below TMT level (see Pelled, Eisenhardt & Xin, 1999). This is because it is assumed that groups assembled for the purposes of say, product redesign or the development of marketing strategies, need a psychologically safe environment to spawn creativity and innovation (West & Anderson, 1996). However, it has been suggested that psychological safety is also important to CEOs and TMTs, because they are interdependent, and the issues facing them are complex and ambiguous (Forbes & Milliken, 1999). Thus it is argued that executives are likely to differ considerably regarding what they think are appropriate courses of action to deal with uncertain situations (Forbes & Milliken, 1999). What is needed is a psychologically safe climate for managing conflicting opinions, so that a consensus decision can be reached (Janssen, Van De Vliert & Veenstra, 1999).

Studies that have investigated psychological safety and similar concepts in TMTs have found that when psychological safety is low TMT members do not believe the decision to be right one (Dooley & Fryxell, 1999). That is to say, fears of opportunistic behaviour by others causes some executives to withhold vital information regarding the decision under consideration, which subsequently results in a collective lack of confidence in the decision (Dooley & Fryxell, 1999). Such a situation can degenerate into person-oriented animosity (Janssen et al., 1999). On the other hand, members of TMTs that are characterised by high levels of psychological safety and participation: feel more in control of the decision process and the outcome (Eisenhardt, 1989); generate more innovative ideas (West &

Anderson, 1996); and use more varied information (freely elicited by members) in arriving at a collective decision (Thomas & McDaniel, 1990). The more complete and accurate the information, the more effective the implementation of the subsequent decision (Dean & Sharfman, 1996). Psychological safety is clearly an important factor in relation to understanding decision-making in top management teams (Edmondson, Roberto & Watkins, 2003), and is an important consideration in this study.

With regard to the links between TMT demographic variation and psychological safety, one can extrapolate from the literature with regard to social integration (the closest conceptually of the processes studied at this level). Smith et al. (1994) found no effects of functional diversity on social integration in a study of TMTs. From an exhaustive review of the non TMT literature, it appears that gender diversity is consistently linked to social integration, but differs depending on the proportional representation of men and women. Some research suggests that men are more socially integrated in female dominated groups, whilst other studies have more negative psychological outcomes than women in the minority (Williams & O'Reilly, 1998; see also Tsui et al., 1992 and Wilson, 2003).

Despite there being so few definitive indications from the literature, it is still widely held that demographic variation is negatively associated with psychological safety, and that psychological safety is low in TMTs (Edmondson, 2002). This, it is argued is not simply due to demographic factors. Indeed, several factors are put forward including fear of the hiring and firing power of the CEO, and the fact that disputes within the TMT tend to become common knowledge in the organisation, alienating those who appear to challenge the status quo (Edmondson et al., 2003). TMTs characterised by high levels of psychological safety, enjoy high levels of interaction (West & Anderson, 1996), and individuals tend to feel more comfortable sharing ideas and viewpoints, even going so far as to present information that contradicts prevailing views held by the TMT (Edmondson et al., 2003).

Synthesising these arguments suggests that whilst psychological safety is low in demographically diverse teams, psychological safety is a facilitator of cognitive variation, meaning that in a psychologically safe environment it is acceptable to have a different opinion. However, as conflicting opinions are associated with reduced levels of satisfaction (as discussed above), it is possible that psychological safety is negative predictor of decision belief.

3.6 Conclusion & Propositions Arising from Literature Reviewed in this Chapter

From the preceding sections, it is clear that cognitive variation in TMTs is rarely studied and this needs to be remedied. Central to an ideal study would be the definition and reliable measurement of cognitive variation at both the individual and team level. Moreover, a rigorous study of cognitive variation would attempt to discover the effect of TMT processes on individuals' decisions and the team's consensus.

This chapter also provided an overview of the vast and disparate decision-making literature and found that two streams may have relevance to TMTs. These are the group problem-solving literature and the strategic decision-making literature. The first can be broadly discounted because it deals with unrealistic situations in specific groups, usually assembled for the purpose of research. The second is more pertinent because it is supposed to be about managerial elites at the strategic apex of organisations.

However, even within this literature, selectivity is required as the models and perspectives tend to be after-the-fact descriptive explanations of particular decisions about organisation structure. The quality of strategic decisions is difficult to define and measure. However, three aspects of decision belief seem to be necessary, these are confidence, satisfaction, and effectiveness. Moreover, it is decision-makers themselves who are the best judges of decision belief. From the foregoing it is clear that the study of decision-making processes and decision belief in TMTs is rare indeed.

This chapter finds that there are significant reasons for studying top management processes as they relate to demographic and cognitive variation and decision-making. One of the most studied concepts in TMT decision-making, albeit that the literature exclusively about TMT decision-making is very small, is procedural rationality, or decision comprehensiveness. All decision researchers regardless of their ontological perspective agree that there is a role for information search, evaluation of alternatives and selection of an option for a decision to happen. Moreover, decision comprehensiveness is consistently found to be a factor in strategic decisions with important consequences. Frequency of team meetings has been found to be associated with the component parts of decision making, and has been studied at TMT level (Smith et al., 1994). Reflexivity is argued to enable a team to critically appraise its objectives and processes, whilst psychological or participative safety facilitates the expression of divergent viewpoints without expectation of censure.

In order to overcome some of the limitations noted in previous studies cited in this chapter, research is required that:

1. uses real TMTs as opposed to artificially contrived groups;
2. measures individual and team level differences in cognition;
3. measures decision belief as perceived by the decision-makers themselves;
4. investigates the relationship between demographic variation and team processes;
5. investigates the relationship between cognitive variation and team processes (rather than rely on a congruence assumption);
6. investigates the relationship between team processes and decision belief.

The next chapter synthesises the concepts and propositions from Chapters 2 and 3 into a guiding model to steer the research for this thesis.

C H A P T E R 4

Objectives of the Study in Detail

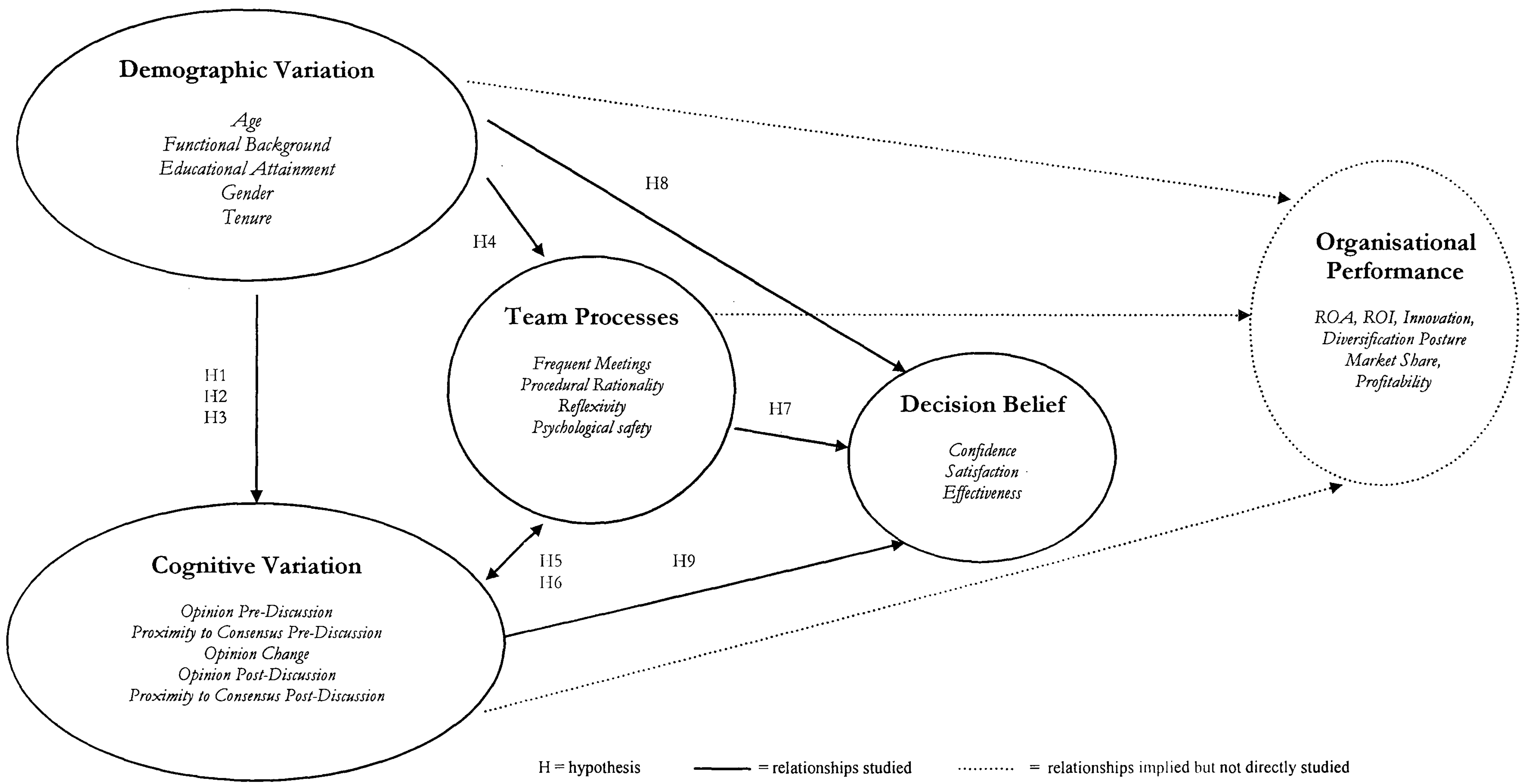
4.0 Overview

This chapter draws together the conclusions from the previous two chapters. The four constituent parts of the study, demographic variation, cognitive variation, team processes and decision belief are brought together. Nine specific hypotheses are generated and these are presented in a way that builds on the process mediator model first introduced in Chapter 2.

4.1 Introduction

Scholars have critical concerns regarding the nature of much top management team research, particularly that which does not define cognitive variation, and which uses demographic variables as proxies for decision-making processes (Priem, Lyon & Dess, 1999). In order to address this concern, research which seeks to provide a detailed specification of the relationships between demographic variables and decision-making processes is needed. In order to provide a response to these challenges, and provide a basis for this research, themes from the previous two chapters are presented in a descriptive framework (see figure 4.1 below). It also shows the predicted effects of demographic variation, cognitive variation and team processes, and of these three sets of variables on decision belief.

Figure 4.1 Research Model for this Thesis



The direction of the arrows suggests the links between the component parts. Hypotheses 1, 2 and 3 relate to the link between demographic variation (at team and individual levels) and cognitive variation, hypothesis 4 relates to the links between demographic diversity and team processes and hypotheses 5 and 6 relate to the link between cognitive variation and team processes. Hypothesis 7 refers to team processes and decision belief, whereas hypotheses 8 and 9 refer to the links between demographic diversity and cognitive diversity respectively and decision belief.

The red dotted lines show the wider context of the research, that is to say, the traditional congruence assumption underpinning 'upper echelons' theory (see for example, Hodgkinson, 2001). As explained in Chapters 2 and 3, TMT demographic variation research typically assumes a link between demographic variation and organisational performance (one of the red dotted lines), and further assumes that process mediates this relationship, although the latter is rarely studied directly.

This research, by way of contrast, assumes that if demographic variation (at the team and individual levels) does not contribute to a collective decision and team processes (as shown by the solid blue lines and consistent with the demographic differences/TMT process model), then it cannot automatically be assumed to contribute to organisational performance (as in the 'upper echelons' model). Therefore, this research aims to contribute significantly to the field by directly investigating the casual assumptions underpinning 'upper echelons' theory.

4.2 Testable Propositions

In the research literature it is argued that decisions made by top management teams depend on important demographic characteristics such as age, organizational tenure, functional background, and education (Hambrick & Mason, 1984; Smith et al., 1996). As explained in Chapter 2, many of the studies of this genre have conflated diversity (proportion of difference represented at the team level), and dissimilarity (the degree to which each individuals differs from his or her peers). Each of the propositions for this study will now be explained, and explicit reference will be made to either diversity or dissimilarity depending on the research question under investigation.

4.2.1 Demographic and Cognitive Variation: Hypotheses 1 -3

Hypothesis 1: 'Upper echelons' is predicated on evidence (albeit limited) to suggest that demographic differences influence cognitive variation, which is generally held in the literature to mean multiple idea generation, or differences of opinion. In Chapter 2, it was explained that many of the posited demographic diversity effects in top management teams, for example turnover rates (Jackson et al., 1991), are directly attributable to dissimilar individuals within the team.

Hypothesis 1 begins with the individual and proposes that it will be those persons demographically most dissimilar to their top management team colleagues in terms of gender, educational background, etc., who will show the greatest cognitive dissimilarity from their colleagues. The proposition is stated as:

Demographic dissimilarity will be positively associated with cognitive dissimilarity.

Hypothesis 2: Diversity refers to the spread of variation or proportion of difference within a team. There is consistent support in the demographic diversity literature for arguing that diversity (rather than the central tendency of the team on isolated demographic traits) is important for decision-making. Previous research in this area makes distinctions between attributes of demographic diversity arguing that certain types of team diversity will lead to different outcomes. For example, Jackson (1996) argues that diversity of professional backgrounds may result in team creativity through members bringing varied expertise to bear. Just as demographic diversity is a team level phenomenon, cognitive diversity for the purpose of this study is defined as variation in opinion held by the team. The second hypothesis predicts that it is those teams characterised by greater variation in gender, educational background, etc., that will show the greatest cognitive variation across members (i.e. the least agreement). Therefore, the second testable proposition can be written thus:

Demographic diversity will be positively associated with cognitive diversity.

Hypothesis 3: A research question naturally arises from the previous two hypotheses, that is, do team level demographic factors (i.e. diversity) influence the way individuals think (i.e. dissimilarity? This is a concept not previously investigated in the research literature, although it is a logical juxtaposition arising from 'upper echelons' theory.

The prediction is that individuals in teams with greater variation in gender, educational attainment, etc., will show greater cognitive difference relative to those in more homogenous teams. The proposition is stated as:

Team demographic diversity will be positively associated with individual cognitive dissimilarity.

4.2.2 Demographic Diversity and Team Processes: Hypothesis 4

Hypothesis 4: The research literature has long suggested that demographic differences are not conducive to well functioning teams (Agrell & Gustafson, 1996). Much of this work has been conducted with groups brought together specifically because their members are different, such as multi-disciplinary teams, in which it seems very difficult to communicate owing to differences in vocabulary and approaches to problem solving.

Specifically this hypothesis posits that demographic diversity will be negatively associated with the team processes, namely: procedural rationality, frequency of meetings, reflexivity and psychological safety. As discussed in Chapters 2 and 3, this hypothesis finds support from the TMT literature. TMT research has associated aspects of demographic diversity with more disputation and debate in team decision-making (Simons et al., 1999) and less frequent communication (Glick et al., 1993), for which frequency of team meetings is often used as a proxy. With regard to the expected effects of demographic diversity on reflexivity and psychological safety, the literature again is informative. Diversity has been associated with less agreement and more conflict (with reference to reflexivity) (Knight et al., 1999; Amason, 1996; Finkelstein & Hambrick, 1990); and more disaffection amongst individuals (Amason, 1996) (i.e. less psychological safety).

Hypothesis 4 can be stated as:

Demographic diversity will be negatively associated with team processes.

4.2.3 Cognitive Variation and Team Processes: Hypothesis 5 and 6

Hypothesis 5: Although hypothesis 4 takes a blanket approach to the effect of demographic diversity and team processes based upon established views concerning small group research, it is of interest to investigate the complexities of team processes with regard to cognitive diversity. For example, it is generally accepted that demographic diversity causes conflict and communication problems in teams (Jackson, 1996), that is to say, diversity is bad for decision-making. On the other hand, there is a growing body of work that suggests that conflict and minority dissent is actually good for small group decision-making (Tjosvold, 1996; Nemeth & Owens, 1996) and particularly for TMTs (West et al., 1997). This is because it is expected that the variety of opinions including dissent, better covers the available options. Two processes were put forward in Chapter 4 as being particularly important for assuaging dissenting views and two for facilitating and encouraging dissenting views, these are discussed next.

As the purpose of team decision-making is to arrive at a consensus decision, a solution that can be embraced and implemented by all members of the team, it is logical to assume that some team processes will smooth out the cognitive variation in teams caused by demographic differences, enabling problem solving, cohesion and consensus (Sniezek & Henry, 1989). Procedural rationality (Dean & Sharfman, 1996) is one such process, whereby teams systematically and analytically assess the information for decision-making, honing in on important information and ignoring irrelevant facts. The second is not strictly a process, but is considered with the processes here, that is, frequency of team meetings. It stands to reason that the more often a team meet, the more familiar the group norms become in terms of communication, political behaviour, and approach to decision-making etc. Therefore, frequency of team meetings is included along with procedural rationality as an expected facilitator of less variation, that is to say, more agreement in personal opinions.

Two processes are believed to facilitate the effective management of conflicting viewpoints. The first of these is team reflexivity which, it is argued, assists the objective identification of problems without expectation of censure (West, 1996). If teams have a number of diverse viewpoints, that is, they experience cognitive diversity, reflexivity is a beneficial process by which they can clarify opinions. In order to be reflexive, a second process variable, psychological safety (Anderson & West, 1996), is necessary. This refers to a shared belief amongst team members that ideas (however radical) and well-intentioned actions will not be punished, rejected, ridiculed or attacked by the rest of the team (Edmondson, 1999). The degree to which team members are willing to exchange information, contribute ideas and explore creative solutions is a function of psychological safety (West, 1984).

Therefore, it is logical to assume that teams which enjoy higher levels of psychological safety will show the greatest cognitive variation. Similarly, teams which feel free to objectively identify problems without expectation of censure are likely to explore more radical ideas than teams which are neither reflexive nor psychologically safe.

Based on this argument, hypothesis 5 is:

Procedural rationality and frequency of team meetings will be negatively associated with cognitive diversity, whereas, reflexivity and psychological safety will be positively associated with cognitive diversity.

Hypothesis 6: Hypothesis 6 is similar in form to hypothesis 3 and asks the question: Do team processes predict individual cognitive dissimilarity? That is, if a team uses procedural rationality and meets often will there be a corresponding reduction in dissimilar views? Conversely, if a team is characterised by reflexivity and psychological safety, will it make the individuals in those teams more radical in their opinions? One would expect that psychological safety and reflexivity will increase cognitive dissimilarity, whilst procedural rationality and frequency of team meetings will reduce it.

Procedural rationality and frequency of team meetings will reduce cognitive dissimilarity, whilst, reflexivity and psychological safety will increase individual cognitive dissimilarity.

4.2.4 Team Process and Decision Belief: Hypothesis 7

Hypothesis 7: As discussed in Chapter 4, there is mixed evidence in the research literature as to the efficacy of team processes for perceived decision belief. Analytical decision-making strategies such as procedural rationality (Dean & Sharfman, 1996) are supposed to make a team effective, and should make a team confident that the decision is correct. Moreover, teams which meet more often are likely to have a history of decisions that have worked well and those that have not, so that they will have more of a feel for what is a good decision and what is not. Hence is it reasonable to assume that frequency of team meetings (as a proxy for familiarity with group norms and communication), will increase perceived effectiveness, satisfaction and confidence.

The contribution of reflexivity to effective decision-making is to foment conflicting and diverse viewpoints amongst team members (West 1996; West, Garrod & Carletta, 1997). Therefore, one would expect more discord, disagreement and affective conflict amongst reflexive teams resulting in reduced levels of satisfaction, confidence and effectiveness due to competing opinions. As psychological safety is supposed to be almost symbiotic with reflexivity, it would be reasonable to assume that persons who feel psychologically safe in a team are able to actively engage in discussions where dissenting views are aired (West & Anderson, 1996; Edmondson, 1999) and therefore, report reduced levels of decision belief. Hypothesis 7 is stated as:

Procedural rationality and frequency of team meetings will be positively associated with decision belief whilst reflexivity and psychological safety will be negatively associated with decision belief.

4.2.5 Demographic Diversity and Decision Belief: Hypothesis 8

Hypothesis 8: The literature reviewed in Chapters 2 and 3 consistently supposes a negative relationship between demographic diversity and decision belief. Although the underlying premises in 'upper echelons' and value-in-diversity traditions, suggests that greater variation in age, functional background, educational attainment, gender and tenure, results in better team decisions, it is widely held that such teams are less satisfied, confident and effective at decision-making than more homogeneous teams. Hypothesis 8 is stated as:

Demographic diversity will be negatively associated with decision belief.

4.2.6 Cognitive Diversity and Decision Belief: Hypothesis 9

Hypothesis 9: Chapter 3 identified three factors inherent in perceived decision belief, namely decision satisfaction, confidence, and perceived effectiveness. Teams that enjoy true consensus (i.e. all members agree with one another), generally report greater satisfaction and confidence (Sniezek & Henry, 1990). Yet this study proposes to investigate variation, that is, the opposite of harmony, so one could expect greater variation to lead to less satisfaction and confidence. Therefore Hypothesis 9 is stated as:

Cognitive diversity will be negatively associated with decision belief.

4.3 Some Implications of the Model for Empirical Research

At a general level the model against which the above propositions have been mapped is an important starting point for research into effects of top management team variation and decision-making. Intact existing teams would need to be used to increase its ecological validity. This is a significant departure from either the top management team research, or most small group research.

The small group research literature favours two typical approaches, either, observational studies of behaviour or task performance (observer coding), or self-report measurement of performance using questionnaires, interviews and case studies. Traditionally, observational studies of teams in general, top management teams, group dynamics, group problem solving are conducted with ad-hoc or synthetic groups (Orasanu & Salas, 1993). Ad-hoc groups are typically made up of randomly selected students. Sometimes, each student is assigned a role play for the top management team and the pseudo team is disbanded after the task.

Synthetic groups may similarly comprise students, but they have a history beyond the research, for example, they work together throughout their course, undertaking various tasks during that period. Other examples of synthetic groups are juries, or senior executives from various organizations on training courses, who are assigned to groups for specific tasks. Belbin's (1983) now famous work on team role models was conducted with synthetic managerial teams as was Isabella & Waddock's (1994) investigation of innovation in banking top management teams.

The advantage of the control afforded by this approach is that the experimenter can dictate the size of the team, usually between 2 and 4 members (McGrath, 1991). The huge disadvantage is the lack of generalizability to existing teams, who have a 'past and a future as a team, and are embedded in a larger social context,' the organization of which the team is a part (McGrath, 1991). In order to study real team processes, an observational study of processes in real teams is appropriate.

4.4 Conclusion

This chapter began by presenting a conceptual overview of the broad themes arising from the literature concerning demographic and cognitive variation, procedural rationality, frequency of team meetings, reflexivity, psychological safety and decision belief. More detail as to how the research seeks to measure the various constructs will be presented in Chapters 5 and 6.

The nine hypotheses presented in this chapter, which draw upon research directions from the literatures reviewed, form the basis for the empirical study which follows in the remaining chapters. The propositions seek to address the inadequacies in current research as they seek to form a basis for understanding the relationships that have traditionally only been indirectly measured, or omitted by previous researchers. The next chapter addresses the recommendations made in the review chapters for conduct of the ideal study into these constructs.

CHAPTER 5

Research Design and Methodology

5.0 Overview

This chapter describes the research design and the methodological decisions made to address criticisms of previous research in this area, as highlighted in Chapters 2, 3 and 4. It explains how a state-of-the-art simulation designed specifically for empirical TMT decision-making research, was used with 23 real teams to address the hypotheses. The chapter concludes with an evaluation of the research design against known validity threats, demonstrating how such threats were addressed.

5.1 Introduction

Earlier chapters have criticised the approaches used in previous studies of top management teams. These have included: (1) the use of synthetic teams; (2) the reliance on public archival data for identifying TMT members based on particular titles and thereby potentially omitting key senior executives involved in particular decisions and who make up the TMT in practice; (3) artificial manipulation of the size of teams; (4) inference drawn from corporate material and news reportage rather than the direct measurement of team processes during decision making; and (5) the use of measures of decision belief open to the biases of selective recall of interviewees based on their commitment to the decision being investigated, status and past experience within the team. The limitations of such approaches concern their external validity, how well one can assume they generalise to actual top management teams.

This chapter explains how the current study sought to overcome some of these methodological flaws by: (1) using real teams; (2) getting companies to identify the members of their top management teams so no one would be missed; (3)

allowing the size of team to vary based on the number of people who made up the TMT in each company; (4) observing teams in the process of decision making so as to understand team processes in situ; (5) using independent raters to observe those processes; and (6) the use of a realistic simulated decision task by all teams so as to hold the task constant for each individual and team.

The first part of this chapter is divided into 8 sections which describe the decisions that had to be made concerning research design. The second, substantial part of the chapter is devoted to describing how The PEAK Selection Simulation© (a sophisticated decision-making simulation designed for use in empirical TMT research) was operationalised to test the hypotheses. The third and final part of the chapter identifies and explores potential criticisms of the research that would, in a conventional quasi-experimental research design, be considered as threats to internal and external validity.

5.2 Research Design Considerations

5.2.1 Use of Real Teams

Traditionally, observational studies of teams are conducted with ad-hoc or synthetic groups (Orasanu & Salas, 1993). It was strongly argued in the preceding chapters that artificially contrived groups of synthetic top management teams cannot generally reflect naturally occurring situations or processes.

In order to overcome some of these methodological flaws, it was decided to conduct this study with actual top management teams in their own environments[1]. The complexity of using such real teams required decisions regarding the selection of appropriate research techniques and valid measuring tools.

5.2.2 Self-Identification of Teams

Definition of TMTs based on title alone has been an important feature in 'upper echelons' research which is chiefly conducted by scanning public archival data. Yet, most executive managers would report that the board does not generally actively engage in the process of decision-making per se, merely it ratifies the decisions which have been made and presented to them by the Chief Executive Officer (CEO) and

---

1 As will be explained later, a very small number of teams chose an off-site location rather than their

his or her group of senior executives. In real world settings, it is this latter group that makes up the top management team.

Following Jackson (1992) this study defined top management team to be the group of senior executives which represented the highest reporting level in the organization, and allowed the CEO or MD of each company to nominate the team members. No designation was prescribed to participating organizations as to who should comprise the top management team. This proved to be a non-issue as far as the companies in this study were concerned, as those at the highest reporting level in all cases proved to be the top management team in practice, which gave comparability across the sample.

5.2.3 Variability in Team Size

It has been argued that real-world dimensions of team functioning such as irregular team size, or a member being absent for a meeting, are features which should be included in empirical research (McGrath, 1991). On two occasions during this research, a nominated member of the top management team could not participate due to illness or other unforeseen circumstance. In all cases however, there was a quorum for decision-making, and this was in keeping with the naturalistic orientation of the research.

5.2.4 Direct Measurement of Team Processes

Very few studies of teams have directly examined dynamic processes because it is so difficult to do (Weingart, 1997). Video recording has been suggested as a means of capturing verbal and non-verbal behaviour in teams that subsequently can be carefully analysed which, it is argued, is better than relying solely on coding in real time (Weingart, 1997). It was decided to use this method.

Permission was sought from all teams to audiotape and videotape their interaction and technical assistance in learning how to use the equipment (i.e. how to place the device to include all members, how to be unobtrusive in placing the audio equipment etc) was provided by technicians in the Psychology department.

5.2.5 Simulated Decision-Making Task

It was necessary to devise a way in which the decision-making process could be observed with a rigour which would control for some of the vagaries of teams decision-making identified by Cohen & Bailey (1997), such as different amounts of time taken by teams for decisions and the type of decision being considered. The alternative to considering then-current business decisions was a simulated decision-making task.

The rationale for using a simulated decision-making task is that the investment of time, energy and commitment to the decision outcome is the same for all individuals and teams, as is the nature of the decision itself.

The usefulness of simulations for research purposes is well documented (Fripp, 1993). However, in terms of observing team processes, this depends on the extent to which they mirror real life, and many, such as those that require making double helix models from drinking straws, or building a white water raft (Hurst, Rush and White, 1989; Fripp, 1993), do not. Indeed, these tend to force individuals to role play behaviour that is inconsistent with their normal behaviour, and to adopt group norms that do not reflect the real-life interaction of top management teams (Hurst et al., 1989).

For the above reasons, highly sophisticated simulations specifically designed to replicate real top management team decision-making tasks, such as those designed by the Center for Creative Leadership (CCL), North Carolina, USA, have been expressly recommended for use in top management team research (Hurst et al., 1989). One such simulation, The PEAK Selection Simulation© (PSS) from CCL, is employed in the current study. Attesting to the plausibility of the PSS to real world situations, it was not uncommon for participants in this study to ask if the simulation had been based on their particular company, as the content and process were so realistic and applicable to their business.

The objective of the PSS is for the top management team to select an incumbent from a list of four candidates for the position of President of the Advanced Product Division of a fictitious company (Looking Glass Inc.).

5.3 Previous Applications of the PSS

Until the current study, the PSS had been used exclusively with synthetic [illegible] CCL in North America. These were teams of senior executives from blue chip American companies on residential leadership courses, with teams made up solely for training purposes. Informational cues within the PSS had been changed so that the candidate pool included either an Afro-American male or a white female. This was due to the focus of research at CCL conducted at that time being on biases in selection and recruitment.

Coding guides designed and supplied by CCL for use with the PSS divided the discussion into 3 discrete timeperiods (of 15 minutes each). The purpose of this coding was to identify which items of information or aspects of candidates' characters were discussed by the team and for how long. These coding guides were not relevant to the hypotheses being investigated by the current study and were discarded. Moreover, CCL does not record the final 15 minutes the group spends together documenting their consensus decision. Yet, typically this study found these final minutes of team discussion to be robust and rich in interaction, important in terms of understanding the team process under investigation.

Owing to its superior emulation of real world decision-making tasks, the PSS is quite involved both in its execution and in its potential for data collection. A later and substantial part of this chapter is devoted to an in-depth explanation of how the PSS is operationalised to address the hypotheses in the current study. Before effusing further describing the PSS, other research design considerations, such as the size and scope of the study, and the process of gaining access to the teams, are addressed in the next four sections.

5.4 Gaining Access: Size and Scope of Study

In order to overcome the difficulty experienced by most researchers in gaining access to top management teams, the means of access and the potential value of participation to the teams required careful consideration. The specific aim of this study was to observe as many whole teams as possible whilst they were actually engaged in the process of decision-making.

Some studies, like that by Amason (1996), with 53 teams incorporates two sub-samples. In Amason's (1996) study, the first sample contained 48 teams (122 respondents) and the second sample contained 5 teams (21 questionnaire survey respondents). Fifteen interviews were subsequently conducted with a selection of survey respondents in the second sample. Both instruments required retrospective responses regarding the executives' participation in a strategic decision nominated by the CEO.

Although Amason (1996) achieved a highly respectable number of teams (53 in total) he did so by restricting the number of members from each team (3 or fewer per team) making them less representative of teams as a whole. Moreover, his research relied on the subjective recollection of the respondents, whose level of participation in the decision-making process and commitment to the outcome undoubtedly influenced their responses (see Cohen & Bailey, 1997 for a critique of such methods). Such an approach would not allow for the current study to rigorously test the hypotheses generated in Chapter 4.

In order for it to be comparable in number with other empirical studies of top management teams (see for example, Hambrick (1981) with 20 teams, O'Reilly et al. (1993) with 24 teams), this study aimed for a sample size of between 20 and 30.

A single study was deemed necessary in order to achieve the desired target of 20+ whole teams. This would make the study between 4 and 10 times larger than the in-depth analysis of sub-sets common in empirical case study research, yet directly comparable in terms of individual participant numbers to some studies that essentially gathered their data via questionnaire survey. In all, 23 teams were obtained.

The next section outlines how the research project was marketed to attract potential participants.

5.5 Attracting the Participants

As an experienced executive trainer, the author designed a one-day training workshop around the PSS, adding discussion, feedback and action planning tasks. It was entitled *Effective Decision-Making in Top Management Teams: A Top Management Team Development Workshop*. Next, a marketing flyer was devised which outlined the following benefits:

- Personalised, individual workshop - only one top management team at a time could participate.
- Immediate feedback from experienced facilitators on observed team performance.
- Hard copy report on the team's responses to a questionnaire, which included comparison data for 650 teams and 83 top management teams.

- Presentation by facilitators on research into teamwork and methods of decision-making.
- Facilitator guidance for critical reflection and action-planning, if appropriate.

The flyer also outlined the cost, which was set at a nominal £50.00 per head, to supplement funds from ESRC. The author's prior experience of working as a consultant and executive trainer strongly suggested that companies would be more likely to accept and value the invitation if they had to pay for participation. Although it is not remarkable now for companies to pay for involvement in research projects, in early 1997 it was not as common. The workshop was still appropriately advertised as not for profit, as the costs in purchasing laptop computers and other technical equipment needed to run the simulation, together with the hiring of vehicles to transport it all, along with the necessary overnight accommodation costs, far outweighed any financial contribution made by the companies.

Setting up the equipment at every site and cueing up each computer for the simulation task to have the same identification code as the paper questionnaires assigned to the individual team members, required careful planning. The equipment needed for the workshop included:

- Laptop Computer for each participant
- Simulation CD ROM – identification code cued in advance
- Sony Professional Walkman
- 2 x Pressure Zone Microphones positioned on the discussion table
- 2 to 1 jack stereo connector
- 90 minute audio tapes
- Headphones for each participant
- VHS Video Camera
- 3 hour VHS Videotapes
- Laptop Computer and Portable Printer for Author to analyse company data using SPSS, compile and print out feedback report.
- LCD Data Projector
- VHS VCR for playing PSS Video to participants
- Spares of everything, batteries and backup recording equipment.

Each participant was supplied with a laptop computer (already programmed and coded for their use) and a set of headphones. The author provided a demonstration of the PSS multimedia programme using an LCD data projector. Team members then conducted an individual information search pertinent to Looking Glass Inc. and the candidates, using the multimedia package. Because the workshop was conducted in one room, the individual computer search took place with all the team members physically being present. Headphones were supplied, enabling team members to access the interactive parts of the programme without disturbing or distracting the others.

5.6 Programme of Activities Conducted with Participants

During the workshop programme, 1 hour was dedicated to housekeeping and facilitation (i.e. explaining tasks and giving instructions), 1½ hours were spent on presentations and feedback to the team by the author or, if appropriate, a companion presenter. Up to four hours were dedicated to research and data collection, of which two and a half hours (decision-making tasks) were video and audiotaped. Most of the tasks were supplemented by self report questionnaires at either group or individual level, and were further supported by in vivo observational coding guides. Only the PSS tasks (individual computer search and team discussion) were used for this research study. The afternoon programme was made up of filler activities designed to engage the interest of the teams and make the workshop programme appealing. The morning activities concerning the PSS are detailed in Figure 5.1 below.

Figure 5.1 Workshop Programme (Morning Activities)

9.00 – 9.15	Welcome and overview of Activities
9:15 – 9:30	Completion of TMT Questionnaire
9:30 – 9:45	PEAK Selection Simulation Introduction and Tutorial
9:45 – 11:00	Individual Computer Searches
11:00 – 11:15	Morning Coffee
11:15 – 11:25	PEAK Selection Simulation Video & Facilitator Overview to Team
11:30 – 12:30	Team Discussion
12:30 – 1:30	Lunch

5.7 Securing the Participation of Top Management Teams

In the first instance the author wrote to the CEOs or MDs of 120 organizations in the Institute of Work Psychology (IWP) database, inviting them to participate in the workshop (see the original letter in the appendix). An initial follow up phone call was made within a few days to ensure that the letter had been received and to get a reaction as to acceptance or rejection.

Positive reactions (i.e companies that initially said yes) were received from 58 companies. Literally dozens more phone calls and follow up letters were made to the companies to secure their participation. In all, 30 companies nominated a date on which they could participate and arrangements were made for delivery of the workshop to these 30.

Typically, the CEO's secretary was the point of contact and she negotiated with all the other directors on the top management team as to their diary dates. The challenges of assembling the whole top management team together on the same day, and then getting them to devote the whole day exclusively to participating in the research, was difficult. One company in the West Midlands conducted the workshop on a Saturday so it would not interfere with the working week. The challenge proved too much for 6 of the companies, and they eventually opted out of the programme. For the final company, the workshop was scheduled to include the whole team of 6 senior executives. On arrival, the author discovered that only two of the executives were available as the other 4 had declined to be involved. It was inappropriate to include this sub-set of the team, so a short presentation on the work of IWP was given to the two directors followed by a factory tour and pub lunch.

The average top management team size was 6, with the largest having 8 members and the smallest, 3. The TMTs that participated were at the strategic apex of companies that ranged in size from 60 to 1000 employees with the average having 253 employees. The turnover of the companies involved (in 1997) was £1 to £7.5 million p.a. with the average turnover being £2.66 million. All were in the UK manufacturing sector. The companies that did not participate had the same characteristics as those involved (this was a pre-requisite for them being in the IWP database for the longitudinal study). Hence there was no significant difference between the sample and the population from which it was drawn.

5.8 Intensive Nature of the Study

The current study was extremely intensive regarding the amount of research and senior manager time required. It took more than 8 months simply to process the total sample of 23 teams through the workshop.

In terms of direct research work, 920 hours (115 days) of author time were required (23 teams x 3 days for the author, plus 23 teams x ½ a day x 2 raters) and 1430 hours (179 days) of senior managers' time (i.e. 1 x 8 hour + day, plus an average of 3 hours each for advance preparation).

All top management teams were offered the options of either undertaking the task at the Institute of Work Psychology in Sheffield, on-site at their premises or at an agreed off-site venue. Nineteen of the teams used the boardroom or other suitable rooms in their offices. Two teams regularly used a local hotel for top management team meetings in order to minimise interruptions and phone calls from staff and colleagues, and took this option in participating in this research. Two of the teams chose the option of attending at Sheffield, where a room in the Psychology department was used as a mock board room for the day.

The participating top management teams were located all over the United Kingdom (Scotland, England, the Isle of Wight, Wales and Northern Ireland), which involved a significant amount of travel for the author and anyone else that accompanied me in order to transport the considerable number of pieces of audio-visual and electronic equipment from site to site. Three times visiting Professors interested in observing the PSS attended, on two occasions doctoral students were in attendance and for 12 of the workshops an MSc student came along.

5.9 The PEAK Selection Simulation©

Generalizability was argued in previous chapters as being an appropriate aim for the current study. The PSS meets this requirement as the selection decision it simulates, with its replication of actual events, information and circumstances encountered by organisations, is typical for any top management team regardless of company or industry.

The current study is predicated on the assumption that team members will bring an array of information and personal preferences to bear on decisions based on their different demographic attributes. The PSS allows for this assumption to be tested. It makes a vast amount of information available to individual team members in text, audio and video formats. Yet, there is too much information for a single individual team member to access and assimilate before they discuss it with their fellow team members. This means that each person has to be selective about the information they choose to review based on their personal preference. The PSS is made up of 3 parts, advance preparation, a computer simulated interview at an individual level, followed by a team discussion.

5.9.1 Advance Preparation by Participants

In order to help executives prepare for the simulated task, a pack containing information on the company, the vacancy, and the candidates is sent about a week prior to the simulation. Each pack contained the following 11 items:

1. Looking Glass Inc. Company Profile
2. Consolidated Balance Sheet
3. Consolidated Statement of Earnings
4. Historical Overview of the Company
5. Draft of the Ten-Year Strategic Plan
6. Memo from a management consultant firm
7. Memo from a financial institution
8. Description of the Advanced Products Division
9. Description of the position of President, Advanced Products Division
10. Organizational Chart
11. Résumés for each of the four short-listed candidates

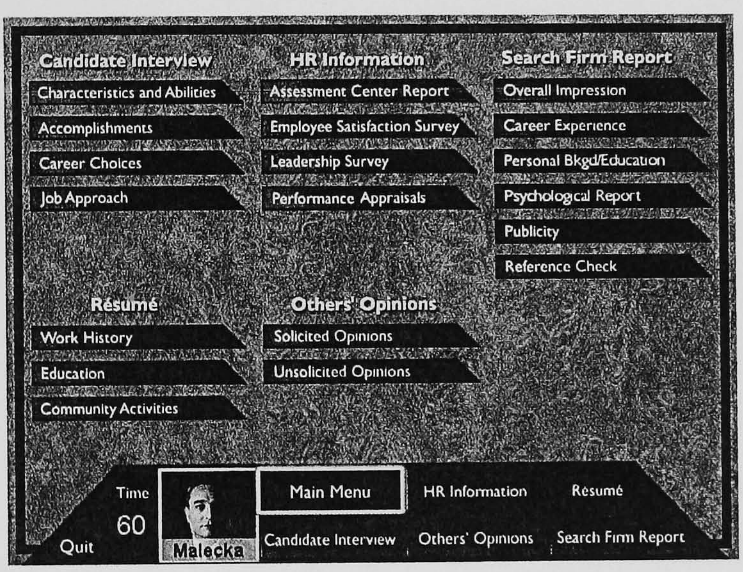
Each of the top management team members was instructed to read and become familiar with this material and to bring the pack with them in order to participate fully in the simulation. All participants in the study reported that they had read the material in advance. Furthermore, they reported that it typically took between one and three hours in order to read and assimilate the information.

5.9.2 The Individual Computer Search

Simulating a real-world type selection scenario, the PSS also allows the top management team members to mock interview the candidates via videoclips and to take up references by phone, hearing the candidate's referees evaluate their performance. This is done by means of an interactive CD ROM accessed via laptop computer. Positive and negative attributes are presented for each candidate, and all are equally represented in the simulation. As with the overload of information in the pack, the balanced yet vast spread in the interactive section allows team members to access what they wish based on personal preference.

The CD ROM has 5 major areas of information (described below) for participants to base their selection decision upon. Icons for each of these sections appear on the screen simultaneously (see fig 5.2 below), no direction or external guidance is given to lead the participants to any particular aspects of the simulation. By clicking on the candidate picture icon, participants could then choose which of the candidates they wished to pull up on screen.

Figure 5.2 Screen Shot of the PEAK Selection Simulation



The 'candidate interview' section allows the participants to select from a drop down menu of pre-set questions to put to the candidates and see and hear their responses. As close as possible to a real executive selection interview, these questions specifically focus on the candidates' characteristics and abilities, career choices, accomplishments and job approach.

Example questions posed are:

"Why are you interested in becoming President of the Advanced Products Division?"

Following the candidate's response, the follow up question is:
"If selected, what would be your strategy for the Division?"

For each question posed, a videoclip of the candidate answering the recorded question is played.

Human Resources (HR) information is provided on three of the four candidates, as they are seeking internal promotion to the position within the company. The 'HR Information' section contains an assessment centre report, an employee satisfaction survey by subordinates, a leadership survey and performance appraisals.

The 'Search Firm Report' section, explains why the candidate was shortlisted and how they performed in the initial round of interviews with the search firm. This section contains subsections on overall impression, career experience, education, personal background, publicity, reference check and psychological reports. Participants are able to hear the candidates' referees talk about them.

The 'Résumé' section contains written information on the candidates' work history, education and community activities.

The fifth section is called 'Other Opinions', and contains tape recordings of three interviews with both male and female referees nominated by the candidate. Two of the references for each candidate are very detailed concerning particular products or plants, or specific instances within the fictitious company or the candidates' lives, whilst the third is non-specific. The participant hears the interviewer pose a question to the referee, and then hears the referee's response. The example below is of a non product, or plant specific, question posed to one of the candidate's (Hank Cooper) referees.

Interviewer: *"Mr Linley, you succeeded Hank Cooper in the TQM position. What would you say were his strengths and weaknesses?"*

Mr Linley: *"All manner of strengths really. I think his strongest is his ability to win loyalty, to win involvement in one project or another. He knows what's right and succeeds in convincing others. He therefore summons the best energies to the best causes. I always have the image of Hank being on the outside though. As soon as he gets the group onto a project, Hank's off to the next project. It's as if he's looking over my shoulder, not right at me".*

Interviewer: *"Use three words to describe Hank Cooper".*

Mr Linley: *"Committed, sincere, ambitious. Also fun! He knows how to party!"*

This part of the CD ROM also contains a simulated phone-in hot-line for two additional references from people not nominated by the candidate. Consistent with the balance inherent in the rest of the simulation, there are two phone-ins for each candidate, one positive and one negative. It is up to the participants to focus on what is important to them personally.

It is almost impossible to capture the reality of this simulation by transcribing the text to paper. The accents, inflexions, pauses, laughter and other nuances of human speech are genuinely conveyed through the actors' voices in the sound clips. The interactive dimension of hearing the voices, and watching the body language of the candidates during the interview, all had a bearing on the selection decisions of the top management teams. Several times during the subsequent team discussions, references would be made to the video clips, particularly the way in which candidates maintained or avoided eye contact, looked confident, or shifted uncomfortably in their seats in response to certain questions. Whilst the written material provided data for objective analysis, the sound and video clips provided cues for subjective judgements by the participants, authentically imitating real world experience.

The final screen of the computer programme requests the participants to rank order the four candidates in order of preference for selection, from 1 = first choice, to 4 = least preferred.

The four candidates can be selected in any one of 24 different rank orders by the participants. To illustrate how the ranking works, there are six alternative rank orders for each candidate of choice, for example:

Malecka, Thompson, Davidson, Cooper

Malecka, Thompson, Cooper, Davidson,

Malecka, Cooper, Davidson, Thompson

Malecka, Cooper, Thompson, Davidson

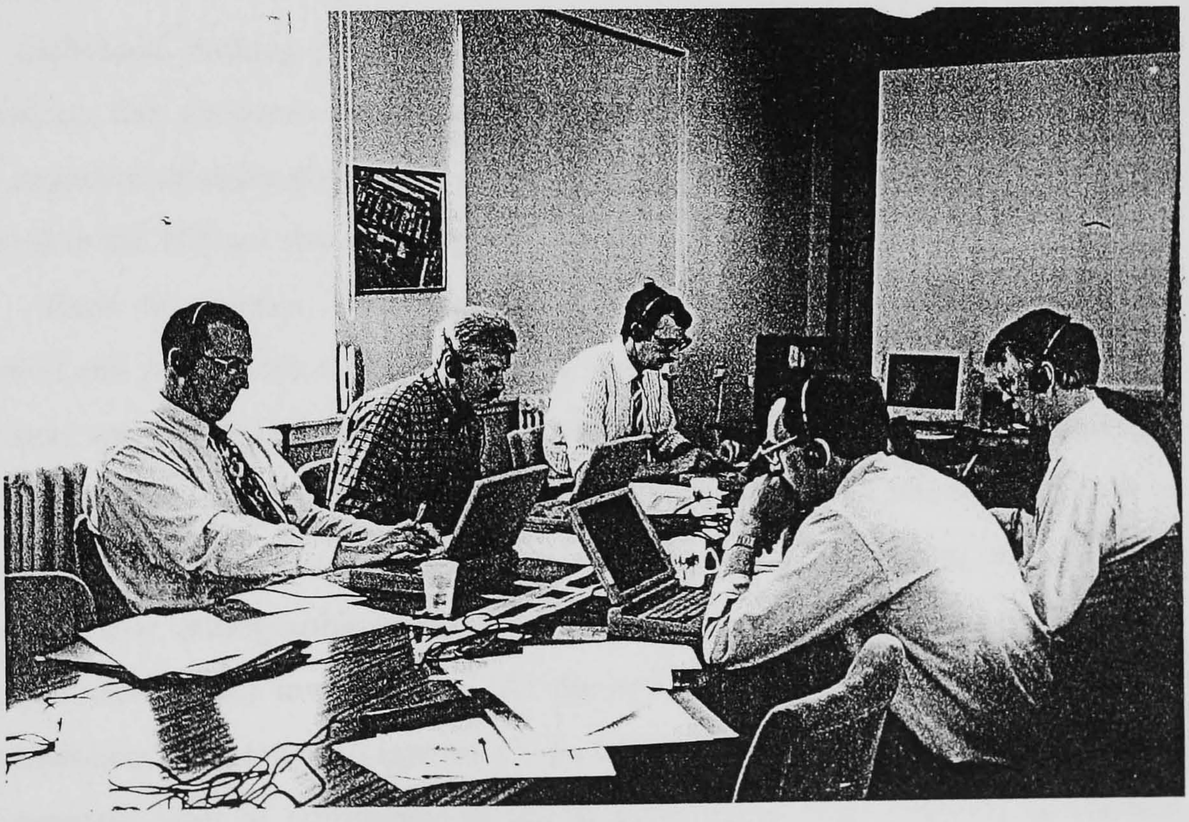
Malecka, Davidson, Thompson, Cooper

Malecka, Davidson, Cooper Thompson.

A disadvantage of any simulation is that in order to demonstrate difference of opinion, the simulated task requires a bounded number of decision choices. However, the 24 choices did not unduly limit the number of different opinions available as 22 were exhibited by the individuals in this study.

In order to test the hypotheses thoroughly, participants were asked to record their personal ranking of the candidates immediately after the computer search (pre-discussion) and again after they had discussed the candidates with their colleagues. As will be explained in Chapter 6, these two data points (together with the team ranking described in the next section) gave rise to 5 individual level variables and 5 team level variables which were used to test Hypotheses 1-3.

Figure 5.3 Photograph of the First TMT in the UK to use the PSS



5.9.3 The PEAK Selection Simulation© Task Part #2

Subsequent to the individual computer search, the team was shown a video-taped message from the CEO of Looking Glass Inc. as an orientation to this session. They were told to discuss the candidates and the position for a period of 45 minutes, during which they are to reach a consensus decision as to the rank ordering of the candidates. At the end of the 45 minute discussion, the team is given a Team Candidate Ranking Sheet on which to record the agreed team ranking of the candidates. Typically this final part of the discussion took another 15 minutes making the overall discussion 60 minutes long.

Immediately subsequent to the team discussion, the individuals in the team were asked to record their private, individual ranking of the candidates. This was to ascertain whether they had changed their minds during the discussion owing to influence from others or information being brought to bear by other team members. Moreover, it allowed for insight into whether the individual team members concurred with the team consensus (see Chapter 6, pp 118 to 127 for a detailed explanation of how the cognitive variation variables were computed). From these three (individual ranking pre-discussion, team ranking and individual ranking post-discussion), five measures of individual cognitive dissimilarity and five measures of team cognitive diversity were computed. The measurement points and the activities involved in the PSS are shown in Figure 5.4 below.

Each 60 minute TMT discussion was videotaped. Independent raters observed and coded each videotape for the team processes of procedural rationality, reflexivity and psychological safety. The raters used a behavioural coding guide, the development and application of which is explained in a section below. Suffice it to say at this point, the observation of the videotapes enabled the testing of hypotheses 4, 5 and 6, that demographic and cognitive variation is positively associated with the team processes under investigation. At the end of the team discussion, individual team members were asked to rate on a 5-point Likert type scale, their perceptions of: their personal level of confidence in the decision made; their personal satisfaction with the decision-making process; and the team's effectiveness in the decision making process.

These 3 items measure decision belief and allow for the testing of hypotheses 7, 8 and 9 which posit that team processes will be associated with decision belief, whereas demographic and cognitive diversity will be negatively associated with decision belief. In summary, it can be concluded that the PSS is a sophisticated simulation that replicates real-life decision-making at senior executive levels, suitable for rigorously testing the hypotheses generated in Chapter 5.

The next section examines how the research design and methodology described thus far in this chapter overcomes typical threats to validity.

5.10 Factors Potentially Jeopardizing Internal and External Validity

It is appropriate for research designs where there could be some ambiguity over the direction of the effects, and/or studies which include an experimental intervention to review factors which are known to jeopardise internal validity. In this case, there is no question as to the direction of the relationships, in as much as, whilst demographic variation may predict team processes and decision belief, the latter cannot predict the former. Hence, threats to internal validity that may typically compromise other studies are not as critical in the current research. However, as the primary research instrument was an artificial simulation, and the principal measures were taken before and after a discussion, some of the concepts about internal validity are of interest. These will be addressed in turn below.

A further feature of the research design was that it involved a single study with intact TMTs, the author was present with the teams, and the teams were audio and videotaped. Hence, issues concerning external validity, that is to say, the degree to which the research design allows for generalisability of results arising, are appropriate to consider. Campbell & Stanley (1963) is still the definitive work on the subject of threats to validity (Trochim, 2002), and their 12 factors will be addressed in turn.

5.10.1 Internal Validity

Campbell & Stanley (1963) identify 8 classes of extraneous variables, which if not controlled for in the research design, may produce effects attributed to the experiment or intervention. Although the current research did not involve an intervention, some critics may wonder how much the simulation, albeit realistic, affected the participants and their responses.

ure 5.4 Design of the Study: Overview of Activities

Subjects	Activity	Amount of Time	Measure
Individual	Read the information pack	½ to 4 hours, up to week before the Workshop	
Individual	Complete demographic survey	Immediately prior to engaging in PSS	Demographic Data
Individual	Conduct computer search for information about candidates	55 minutes	
Individual	Rank candidates in order of preference	5 minutes	Individuals' pre-discussion Ranking
Team	Team Discussion	45 minutes	Observation of Processes
Team	Rank candidates in order of preference	15 minutes	Team Ranking
Individual	Rank candidates in order of preference	5 minutes	Individuals' post-discussion Ranking

5.10.1.1 History

The history threat to internal validity refers to extraneous events or processes occurring between the first and second measurement which may affect outcome measures. In this study, the time from the first (pre-discussion), to the second (post-discussion) measurement was approximately 1¼ hours with only the team discussion and a short coffee-break taking place in-between. Therefore, any change in individuals' personal rankings could be reasonably attributed to the discussion in which a team consensus ranking was achieved as this was the only significant activity that took place between measurements.

5.10.1.2 Maturation

The maturation threat (cf. Campbell & Stanley, 1963; Cook, Campbell & Peracchio, 1990) refers to biological changes or environmental pressures that affect individuals as a result of the passage of time. In this study, each team experienced the simulation in exactly the same way, at the same time of day. Teams did not have the opportunity to mature at different rates in respect to the outcome measures.

5.10.1.3 Testing

The testing threat (cf. Campbell & Stanley, 1963; Cook et al., 1990; Trochim, 2002) refers to the change in individuals' test scores that can occur due to familiarity with the questions, or altered responses on the second test after having discussed the items with others. With regard to the PSS, it could be possible for individuals to change their preferential ranking order of candidates post-discussion to suggest an opinion that could be deemed to be more socially acceptable. So for example, if an individual had ranked the candidates pre-discussion in order of preference as 4123 and the team through discussion came to a consensus ranking of 3214, the individual might also change their private post-discussion ranking to 3214. This change in rankings could be due to simply changing one's mind and being influenced by team processes during the discussion, which is being investigated by the current study. On the other hand, the change in ranking pre- and post-discussion could be due to the individual wanting to present himself better, by recording what he now perceives, (post-discussion with his peers), to be a more socially acceptable personal ranking. If this latter scenario was true, then the so-called testing threat would have confounded the internal validity of the pre- and post-discussion ranking. This potential threat was addressed in two ways by the research design used in the current study.

First, both measurements were completely anonymous. The pre-discussion measure was entered direct to the computer by each individual. The post-discussion measure was gathered by means of a hard-copy self-report questionnaire. Anonymity encourages honest responses because there is no pressure for individuals to achieve a socially approved answer.

Second, a control question was introduced as a means of checking whether individuals were seeking social desirability in their post-discussion ranking. The control question asked individuals to record their pre-discussion ranking on the hard copy questionnaire. The assumption was that if individuals had genuinely changed their minds due to increased awareness of information due to the team discussion, then they would be honest about their pre-discussion ranking. However, if they were seeking social approval, then it is likely they would lie about their pre-discussion rank.

A comparison between the actual pre-discussion ranking and the control recollected ranking was computed as a variable called Liar in which a match attracted value of 0 and mismatch (lie) = 1. One hundred and eleven (86%) of respondents told the truth about their pre-discussion ranking and 19 (14%) appeared to have lied (or made a mistake). Further investigation found that in 12 cases the ranking order was similar with two ranks transposed (e.g. 3142 recorded and 3124 recollected). In all of these cases the first candidate was correctly recalled. In 3 cases individuals could not recall the ranking that they had made to the computer. This left only 4 (3%) actual liars whose recorded and recollected rankings were entirely different (4213/2314; 3412/2341; 3241/4213; 2341/3142 respectively). All 4 liars (3 males and 1 female) were in different teams. This comparison of actual pre-discussion rankings and recollected pre-discussion rankings served as a proxy for assessing the testing threat. The overwhelming majority (95%) of respondents correctly stated the candidate of preference. Therefore, one can conclude that the internal validity of this dataset was not compromised by the testing threat.

5.10.1.4 Instrumentation

Instrumentation decay refers to differences between measurements which are attributable to increased familiarity with, or change in, the instrument over time and across interventions. With respect to the current study, this threat could have been a problem for the independent observers used to rate team processes as they had to learn to use the coding guide. As will be explained in the relevant section of Chapter 6, they were taught to use the instrument on a series of training tapes from CCL so that there was no need for them to learn on the actual dataset. Moreover, no information was provided to the raters as to the rankings by individuals pre- or post-discussion, so their judgement was not clouded by expectations as to how a team might interact, nor did they become familiar with the teams, as each videotape was watched only once by each rater.

5.10.1.5 Non-applicable threats

Further threats to internal validity identified by Campbell & Stanley (1963; see also Cook et al., 1990 and Trochim, 2002), apply particularly to multiple group or pre-post test measurement. These are: statistical regression which applies when the sample has been selected simply because it is an extreme group; biases attributable to non-random selection of comparison groups; experimental mortality, which refers to loss of respondents between measurement timepoints; and selection-maturation interaction which as the term implies, refers to biases due to selection factors interacting with other threats. In the current study, these factors did not apply. There was only one group of 23 teams, therefore no comparison group biases could infect the dataset and the timeperiod between the measurements and discussion were so condensed that maturation interaction effects were not an issue.

5.10.2 External Validity

The chief threat to the external validity of the research design described in this chapter is called the reactive arrangements effect (Campbell & Stanley, 1963; Trochim, 2002). That is, did the simulation, the videotaping and the researchers being present, affect the process so as to negate the generalisability of the findings from the study? Each of these aspects will be addressed in turn.

Although the simulation makes it easy to measure decision-making by reducing the concept to a comparison of rankings, it should be remembered that the TMTs in this study are making the selection appointment of a very senior individual. In real life, such a choice is likely to have far reaching effects on the company and the TMT itself. The decision has to be right. Moreover, the simulation replicated a decision that every one of the teams would have been involved in before and likely to be in the future. Therefore one would expect that it would provide a reasonably accurate insight into how TMTs behave in real life selection situations, without unduly affecting generalizability.

As respects the videotaping, the potential threat is that people act differently for the camera. It may be, for example, that they are shy or nervous and so do not participate as much as they would normally, or they play act so as to present themselves in a better light. Although this is a real threat, there are two factors that significantly reduce its effect in this study. First, videotaping of team meetings is becoming increasingly common in management development (Weingart, 1997) so that it is not particularly unusual at this level in organisations. Second, the TMTs in this study are fully functioning, intact teams. This means that it is difficult for individuals to adopt and keep up a persona which is substantially different to how they regularly behave, whilst they are in a team meeting interacting with their peers in their own environment. Moreover, observation of and discussion with the teams suggested that they were essentially unaware of being videotaped.

Closely linked to the latter threat, is the reactive effect of the researcher being present. Following recommendations by Weingart (1997), this was done as unobtrusively as possible. The presence of the researcher sitting in a corner out of sight, did not intrude on the teams' normal seating arrangements or on the way members interacted with each other. Evidence from the videotapes bears out that the researcher was not included in the discussions, nor was the author addressed by the teams during the team discussion.

The external validity of studies is often challenged on the selection biases of the subjects involved. Indeed, as was revealed in Chapter 2, many studies claim that their results are germane to TMTs, but they actually study teams or dyads at lower levels in organisations, negating their generalizability to TMTs. This study sought to focus exclusively on TMTs and makes no assertions about other types of teams.

5.11 Conclusion

This chapter has presented a novel yet robust research design for studying TMTs in the process of decision-making. The design is novel in that it is a single study of 23 whole, intact teams, of 130 top managers, using a simulated, realistic decision-making task.

The simulation enables the data capture of individuals' personal decisions (rankings) pre- and post-discussion, together with their teams' consensus decision. Thus it offers a rare opportunity to examine the degree to which individuals' opinions differ or converge with each other and with the team consensus. This goes right to the heart of 'upper echelons' theory, which posits that demographically diverse teams will have more variation in the individual ideas brought to bear on a team decision. None of the studies reviewed in Chapters 2 and 3 measured the conclusions reached by individual team members and the extent to which they differ *prior* to a team decision. To the contrary, most research in this arena relies upon post hoc, retrospective recollections of a sub-set of individual team members as to how they differed from their peers. Rather than rely on biased, or subjective memory recall, the recording of the individuals' pre- and post-discussion decisions is immediate, providing little opportunity if any for the expedient massaging of decisions to suit team influences or to impress the researcher. Hence, using the PSS for research purposes provides an important and reliable robustness not available in traditional methods.

In order to properly understand the dynamics of procedural rationality, reflexivity and participative safety, this study relied on observing the teams in the process of decision-making. Although the presence of the researcher and the videotaping equipment may have influenced the process to some degree, it was reasoned that any reactive effect' was negligible due to the individuals being surrounded by their peers in a typical team environment, engaged in a familiar decision process. Other threats to validity, such as history and maturation, were addressed by the restricted timeframe in which the PSS was conducted. All teams experienced the simulation in the same way, at the same time of day. The timespan between the measurement of individuals pre- and post-discussion rankings was 1 ¼ hours, punctuated only by the team discussion.

The research design and methodology outlined in this chapter presents an exciting yet rigorous way to get inside the black box of TMT functioning. The next chapter describes the measures used.

CHAPTER 6

Measures and Descriptive Statistics

6.0 Overview

This chapter describes the measures used to test the hypotheses generated in Chapter 4 and operationalised in Chapter 5 concerning four discrete aspects of investigation: demographic variation, cognitive variation, team processes and decision belief. The chapter also presents the descriptive statistics for these four groups of variables.

6.1 Introduction

Of the nine hypotheses proposed in Chapter 5, one was at the individual level (dissimilarity) six were specifically at the team level of analysis (diversity), and two transcended both levels. This chapter describes how standard computational measures were applied to the demographic attribute data at the individual (dissimilarity) levels and team (diversity) levels. The chapter then goes on to show how these same measures were applied to the cognitive (ranking) data to derive cognitive dissimilarity and cognitive diversity variables.

In the appropriate sections, the chapter explains how dispersion at the team level can be measured in different ways, using indexes such as Blau (1977) for categorical variables, and a mean or median coefficient of variation (standard deviation divided by the mean or median) for continuous variables that have a theoretically and practically fixed 0 point. The median coefficient is preferred in circumstances where the mean is artificially inflated by a skewed distribution. In circumstances where the mean is badly affected by non-normal distributions, and at the same time, where there is no theoretically fixed zero point (as for age and tenure diversity), a new index (Stride) is introduced which works on the same principle as Blau (1977). These indices (Blau, median coefficient of variation and Stride) are all suitable for parametric analysis. Blau's (1977) and Stride's index (Stride, Swift & Wall, 2000; see appendix A) are used for the demographic diversity variables, and Blau's (1977) and the median coefficient of variation are used for the cognitive variation data (unlike the demographic variables of age and tenure, the cognitive variables do have a theoretically fixed zero point, making a median coefficient of variation statistic appropriate).

Team processes were measured using a bespoke coding guide, and a substantial part of the chapter is devoted to explaining the rationale for its development, the training of raters, the application, and the validity and reliability of the guide during the execution of the PEAK Selection Simulation©. The final part of the chapter describes the self report instrument used to measure decision belief.

6.2 Measuring Demographic Variation

Demographic data were collected regarding individuals' age, functional background, highest educational attainment, gender and top management team tenure. Three further items were also collected, namely industry tenure, company tenure, and ethnicity. The tenure variables are all closely linked conceptually, and highly statistically significantly correlated in practice, so that it is appropriate to use one as a proxy for all three. As this study is primarily concerned with teams, the team tenure variable was selected and forms the basis for the analysis reported in the next chapter. Individuals were also asked to state their ethnic origin. Only 3 recorded an ethnic origin other than White UK, and all participants were White. The non-UK individuals were Irish, Austrian and New Zealander males. As the proportional representation of ethnic minorities was negligible, this variable was discounted in analysis.

Data on the five demographic attributes (age, functional background, educational attainment, gender and tenure) were used to compute dissimilarity indices at the individual level and diversity indices at the team level. Although superficially similar, the measures of dissimilarity and diversity are poles apart. Dissimilarity measures assign a discrete value to each individual based on his or her distance from others within the team. Diversity indices on the other hand, apply only at the team level, and assign a single value per team representing the proportion of difference across team members as a whole. In practice, as the examples in this chapter will illustrate, an individual may be very dissimilar from his or her team colleagues (for example, a finance director in a team with three engineers), but the team level diversity may be relatively low (because three of the four have exactly the same functional background). Each attribute will be taken in turn, and the relevant measures explained.

6.2.1 Age

Individuals' ages ranged from 25 years to 66 years, the mean was 44.9 and the S.D. was 9.3.

6.2.1.1 Age Dissimilarity

A standard Euclidean distance measure was used to compute the extent to which each individual differed from the others on their team with respect to age. This is in keeping with other TMT research which uses this computation (see for example, Jackson et al., 1991; Wagner et al., 1984). The Euclidean distance measure takes the individual as the absolute starting point and progressively calculates the difference between the individual and every other individual on the team. Age dissimilarity ranged from 3 years to 57.9 years with a mean of 22.0 and an S.D. of 10.6. In order to illustrate how the Euclidean distance measure works, the ages of a seven member team are shown in the table below. In this case it was the youngest member, aged 29, who was most dissimilar to his peers, taking a distance value of 57.97. The next most dissimilar individual was 66, who had a distance value of 47.77.

Table 6.1 Age, Age Dissimilarity and Age Diversity Values for Team 15

	Age (years)	Age Dissimilarity (Euclidean Distance to Peers)	Age Diversity (Coefficient of Variation)	Age Diversity (Stride Index)
Person 1	46	23.27	.23	.91
Person 2	47	22.05	.23	.91
Person 3	59	33.48	.23	.91
Person 4	66	47.77	.23	.91
Person 5	29	57.97	.23	.91
Person 6	50	22.45	.23	.91
Person 7	48	21.64	.23	.91

6.2.1.2 Age Diversity

The most widely used method of computing team level diversity on continuous variables is the mean coefficient of variation (Allison, 1978). Allison (1978) was primarily concerned with how sociologists measured dimensions of social welfare and inequity between nation states, cities and what he called other social units. The measure that he proposed was appropriate for measuring variables such as income inequity and was first used with extremely large datasets. This measure has since been widely adopted by social psychologists to assess differences across organizations and work teams (see for example, Wagner et al., 1984; Bantel & Jackson, 1989; Michel & Hambrick, 1992; Jackson et al., 1991; Pelled, Eisenhardt & Xin, 1999).

The coefficient of variation statistic most commonly used is the standard deviation divided by the mean. As a measure of inequality, it has the advantages of being easy to compute and of providing a scale invariant measure of dispersion. However, it is sensitive to outliers, inappropriate for non-normally distributed data and unstable when the sample size is small (Stride et al., 2000). Yet, despite these negative considerations, it is the statistic of choice for small group researchers investigating demographic diversity. In a review of fifteen demographic diversity studies, Tsui et al. (1995) found that all those that included a demographic attribute measured on a continuous scale employed the mean coefficient of variation. The study of work teams at any level of the organisational hierarchy (top management to shop floor) is likely to include small teams or tiny teams and one extreme value within a small team can inflate the coefficient of variation disproportionately. The coefficient of variation also requires the variable to have a theoretically fixed zero point (i.e. to be a ratio scale), which is clearly not applicable to age diversity in this study. Moreover, the mean coefficient of variation does not have an upper limit and hence cannot indicate how close any sub-sample is to the maximum achievable level of diversity within the whole sample. Although there are alternatives (Gini & Simpson indexes for example, see Tsui et al., 1995) they are actually based on the same computation used by Allison (Agresti & Agresti, 1978).

Appropriate statistical help was sought as to how to address this issue. Dr Chris Stride from IWP devised an alternative measure, based on conventional statistical ranking theory. Stride's index ranks all ages recorded across the dataset, then for each team, the difference between the ranks is computed and divided by the number in the team. The Stride index is bounded by maximum and minimum levels of diversity achievable within a subset of the whole sample. The resulting team values for both (Allison & Stride) were compared. Where the Allison index typically assigned an artificially high diversity value to a small team that contained one considerably older person, and then assigned the same value (i.e. the same level of heterogeneity) to a larger team with a greater spread of age (i.e. more diverse), the Stride index was better equipped to assess diversity in the particular team as compared to the rest of the dataset. A full explanation of the Stride index is provided in the Appendix (p 298).

Suffice it to say here, the Allison index severely restricted the range of diversity across the 23 teams from 0.3 to 0.42, making any inferences about low, moderate, high diversity very difficult. The Stride index is superior, ranging from 0.53 to 0.93. A further advantage of the Stride index is that because it takes values of between 0.00 and 1.00 ranging from total homogeneity to total diversity) per team, it can more readily be compared with indices such as Blau's (1977) for dichotomous variables (such as gender, functional background etc) which take the same values. It is the Stride index that is used for age diversity in the analysis that follows in Chapters 7 through 10.

6.2.2 Functional Background

The 130 executives came from seven functional backgrounds. Engineering (n = 52), production (n = 13) and scientific (n = 12) backgrounds accounted for 60% of team members, whilst management (n = 22), finance (n = 17), marketing/sales (n = 12) and human resource management (n = 2) backgrounds made up the remaining 40% of the sample.

6.2.2.1 Functional Background Dissimilarity

Typically, the TMT literature does not measure dissimilarity on functional background, probably because standard distance measures (such as Euclidean distance, and squared Euclidean distance) are not appropriate for non-ordered categorical data with more than two categories. This is because the distances between the categories (e.g. finance and engineering) are arbitrary and not ordered.

However, in order to test hypothesis 1, posited in Chapter 5 concerning demographic dissimilarity and cognitive dissimilarity, an appropriate measure of distance for functional background (and educational attainment) is necessary. A review of past and present textbooks on statistical methods overwhelmingly confirmed the absence of a standard distance measure for non-ordered interval data. Information on similarity proximities (Downie & Heath, 1970; SPSS, 1993), however, demonstrated how simple matching measures, done manually on a match = 1, mismatch = 0 basis, could yield a similarity index on non-ordered categorical data. This basic principle was adopted for this study, but mismatches attracted the heavier weighting (i.e. 1) as this was the feature of interest. The match/mismatch needed to be honed slightly to acknowledge the similarities within what are termed elsewhere (Murray, 1989) the hard management functions (i.e. scientific, engineering, production) and those within the so-called soft management functions (i.e. finance, personnel, marketing). In practice this meant that a mismatch between say, finance and engineering attracted a weighting of 1, whereas a mismatch between finance and marketing attracted a mismatch of 0.5. An example drawn from the dataset, team 2 which has 4 members, is tabled below.

Table 6.2 Match/Mismatch Distance Index for Functional Background[1]

	1 Engineering	2 Production	3 Finance	4 Engineering	Distance Index
Person 1	-	.5	1	0	1.5
Person 2	.5	-	1	.5	2
Person 3	1	1	-	1	3
Person 4	0	.5	1	-	1.5

In this case it is the individual from the finance background that is most dissimilar to his peers from engineering and production, with the individual from the production background being more similar to but not the same as his engineering colleagues.

Functional background dissimilarity ranged from 0.0 (where each person was no different from the other) to 6.00. The mean was 3.1 and the S.D. was 1.54.

[1] Full matrix included to show summary by rows

6.2.2.2 Functional Background Diversity

Blau's index (1977) is the most widely used for computing degrees of diversity within a defined sample set such as teams for categorical variables with any number of categories. It essentially takes a position of total homogeneity as the absolute value 0.00, and measures the aggregated proportion of difference added by each individual. The upper value limit of 1.0 is determined by each possible category being represented.

Functional background diversity was measured using Blau's (1977) index of heterogeneity and took values from 0.00 to 0.80, indicating that there were teams in the sample that were totally homogeneous and some that were highly diverse.

The team (number 12) which took a 0 diversity value was made up of three individuals who all had an engineering background. The other 22 teams took values of 0.28 upwards.

6.2.3 Educational Attainment

Individual team members were asked to state the highest educational qualification they had attained. These were then collated into 10 categories. At the lowest end of the scale were O Level/GCSE exams (n = 11), and at the highest was PhD (n = 6). Fifty nine executives had some form of higher education at undergraduate level ranging from HNC to BSc. Thirty two had postgraduate or professional qualifications.

6.2.3.1 Educational Attainment Dissimilarity

As with functional background dissimilarity, the TMT literature offers little direction for computing educational dissimilarity. The vast majority of studies do not refer to it, or use prestige of university attended to differentiate amongst individuals (e.g. Jackson et al., 1991). This latter proxy was inappropriate in this dataset as less than half the individuals had attended University, some stated other professional qualifications (such as accountancy and NVQs) as their highest educational level, whereas others had not been educated beyond secondary level.

The match/mismatch index established for functional background dissimilarity and reported in the last section, was used as a template for computing educational attainment dissimilarity. That is, the same principles were applied to determine distances between educational attainment, where a mismatch is weighted as 1, a match as 0 and a mismatch within a type of education such as tertiary (i.e. BSc and MSc), 0.5. An example drawn from the dataset, team 9 which has 6 members is tabled below:

In this case, it is the individual whose highest qualification is an A Level, and who has not engaged in continuing professional development or tertiary education, that is most dissimilar to his colleagues.

Table 6.3 Match/Mismatch Distance Index for Educational Attainment

	1 PhD	2 NVQ	3 Professional	4 ONC	5 A Level	6 MSc	Distance Index
Person 1	-	1	1	1	1	.5	4.5
Person 2	1	-	0	.5	1	1	3.5
Person 3	1	0	-	.5	1	1	3.5
Person 4	1	.5	.5	-	1	1	4
Person 5	1	1	1	1	-	1	5
Person 6	.5	1	1	1	1	-	4.5

Although the match/mismatch index devised for the dataset may be crude, it serves as useful proxy the accuracy of which can be confirmed by reference to the raw data in the tables above. The spread of qualifications in team 9 (illustrated in Table 6.3 above), from A Level through to PhD, demonstrates the reason that using prestige of University attended (Jackson, 1991) is an inappropriate proxy for educational dissimilarity.

Educational attainment dissimilarity for the 130 individuals ranged from 0.5 to 7.0, with the mean being 3.31 and the S.D. 1.39.

6.2.3.2 Educational Attainment Diversity

Blau's index of heterogeneity was used and took team values for educational attainment diversity of between 0.44 and 0.83 indicating that all 23 teams had a reasonable level of difference in educational attainment.

6.2.4 Gender

Twelve participating executives were female, which represented 9.2% of the overall sample of 130. Although this is a relatively small number of women per se, gender was still included for two reasons. First, the 9.2% female representation is higher than other top management team studies, which typically drop gender diversity owing to non or minimal female representation (e.g. Jackson et al., 1991). Second, female team members were distributed across 9 of the 23 teams meaning that 39% of the teams in the study had some female representation. This meant that a diversity index could meaningfully be used to differentiate between all male and mixed gender teams. No previous study of TMTs has reported gender, or achieved a 39% representation across teams. As this is a naturalistic study of team demographics, to omit gender would potentially ignore an important part of team functioning and process, and one of the more salient demographic variables.

6.2.4.1 Gender Dissimilarity

Gender dissimilarity was computed using the Binary Euclidean Distance measure, which is derived from a fourfold table as the square root (b+c) where b and c represent the diagonal cells corresponding to cases present on one item (i.e. male) but absent on the other (i.e. female). The distance values for gender were then standardised, so that in team 4, for example, which had 5 male members and 1 female member, the males took values of 0.17 and the female, 0.83. In team 11, which had 8 members, the three females took values of 0.63 and the 5 males took values of 0.38. Gender dissimilarity across the sample of 130 individuals ranged from 0.00 to 0.87, with a mean of 0.14 and an S.D. of 0.22

6.2.4.2 Gender Diversity

Blau's (1977) index was used to measure gender diversity and took values of between 0.00 and 0.47 for the 23 teams. The mean was 0.13 and the S.D. 0.17. The 0.00 value represents teams that were 100% male. In the current sample, team 4 which had 1 woman with 5 men, took a value of 0.28. In team 11 where 3 of the 8 members were women, the team diversity value was 0.47.

6.2.5 Tenure

Individuals' top management team tenure ranged from 2 months to 32 years with a mean of 6 years and an S.D. of 6.2 years.

6.2.5.1 Tenure Dissimilarity

The Euclidean distance measure was used to compute individuals' dissimilarity to other fellow team members on top management team tenure. The values obtained ranged from 5 months to 23.4 years, with a mean of 12 years, and an S.D. of 8.7 years.

6.2.5.2 Tenure Diversity

As reported with regard to the other continuous demographic variable in this analysis (age), the coefficient of variation (Allison, 1978) was unsuitable for all the reasons previously mentioned. The Stride index was once again employed, so that each individual's tenure was ranked across the dataset, then the difference computed for individuals' ranks relative to the rest of the team. The Stride index for tenure diversity ranged from 0.41 to 0.95.

6.2.6 Summary of Demographic Variation Variables

This section has explained how distance measures were applied to the demographic attribute data to determine individual demographic dissimilarity (Euclidean Distance, Binary Euclidean Distance and match/mismatch), and diversity indices were applied at the team level (Blau, 1977, for categorical variables; and Stride et al., 2000 for age and tenure). In summary, there are 5 demographic dissimilarity variables at the individual level, and 5 demographic diversity variables at the team level.

Table 6.4 Descriptive Statistics for Demographic Variables

Variable	Min	Max	Mean	S.D.
Dissimilarity				
Age dissimilarity	3.0 yrs	57.9 yrs	22 yrs	10.6 yrs
Gender dissimilarity	0.00	0.87	0.14	0.22
Functional background dissimilarity	0.00	6.00	3.10	1.54
Educational attainment dissimilarity	0.50	7.00	3.31	1.39
Tenure dissimilarity	5 mths	23.5 yrs	12 yrs	8.7 yrs
Diversity				
Age diversity	0.53	0.93	0.83	0.92
Gender diversity	0.00	0.47	0.13	0.17
Functional background diversity	0.00	0.80	0.61	0.18
Educational attainment diversity	0.44	0.83	0.65	0.10
Tenure diversity	0.41	0.95	0.78	0.15

n = 120 - 130 for dissimilarity variables depending on missing values
n = 23 for diversity variables

The next section describes how the same statistical principles were applied to the cognitive variation data in order to measure individual level cognitive dissimilarity and team level cognitive diversity.

6.3 Measuring Cognitive Variation

In Chapter 3 it was established that cognitive differences refer to the different knowledge bases, perspectives, attitudes, beliefs, biases, filters and heuristics brought to bear on decision-making as a result of demographic differences. As discussed in Chapter 3, decision-makers are often oblivious to many of these psychographic factors, and their inherent complexity makes them virtually impossible to measure directly. Fundamentally, however, what is of interest in the current study is how cognitive variation manifests itself, at the individual and team levels, and especially how individual contributions (ideas) are made to the "cognitive gene pool" (Huff, 1990) of the team decision. As was explained in Chapter 5, the PEAK Selection Simulation was employed to measure individuals' judgements, that is to say, their rank order of candidates before and after a team discussion. These rankings are the manifestation of managerial cognition because participants assign a ranking to each of the candidates based on their personal priorities, biases and filters et cetera, as to which characteristics are important. For example, a team comprising solely 50 year old male engineers is likely to value particular candidates' attributes in a similar way owing to the parity of their cognitive filters, biases, attitudes and so on. By way of contrast, a highly diverse TMT is likely to experience more cognitive variation amongst its members, as the individuals' mental models of what is an important candidate characteristic are different. Hence, one would expect that the latter team would come up with greater diversity of rankings than the first. This section has established that individuals' rankings are an appropriate proxy for cognitive variation. As further discussed in Chapter 5, the team as a unit also ranked the candidates during the consensus discussion.

From these three discrete sets of rankings, five individual level dissimilarity variables and five team level diversity variables can be computed. At the individual level these are:

(1) cognitive dissimilarity pre-discussion (distance between the pre-discussion ranking of the individual and other team members);

(2) cognitive cohesion dissimilarity pre-discussion (the proximity of the individuals' pre-discussion ranking to the eventual team consensus);
(3) cognitive change dissimilarity (the correlation between individuals' personal rank order pre- and post-discussion);
(4) cognitive dissimilarity post-discussion (distance between the post-discussion ranking of the individual and other team members);
(5) cognitive cohesion dissimilarity post-discussion (the proximity of the individuals' post-discussion ranking to the team consensus).

For each of the five individual dissimilarity variables outlined above, a corresponding team level diversity index was computed. So at the team level, the dependent variables are:

(1) cognitive diversity pre-discussion (variation in pre-discussion rankings across the team);
(2) cognitive cohesion diversity pre-discussion (variation in pre-discussion rankings achieving proximity to the eventual team consensus);
(3) cognitive change diversity (the dispersion of ranking changes across the team between pre- and post-discussion);
(4) cognitive diversity post-discussion (variation in post-discussion rankings across the team);
(5) cognitive cohesion diversity post-discussion (proportion of post-discussion rankings achieving proximity to the team consensus).

The reason for computing these two sets of variables was to test hypotheses 1 and 2, which broadly assert that differences in demographic attributes will be positively associated with cognitive differences.

The rationale for each computed variable, how each was derived and the descriptive frequencies will be addressed in turn in the sections below. Furthermore, the hypotheses posited in Chapter 4 will be refined as appropriate to harmonise with the operational detail afforded by the PEAK Selection Simulation. This means that where appropriate, an indication as to the direction of the expected relationships will be provided. This may seem a little unusual in a chapter devoted to measures. However, it is fundamental given the complexity of the variables under consideration, to grasp the computational issues concerning the measures in order to understand the expected direction. Hence, these asides are included in this rather than the preceding chapter.

6.3.1 Cognitive Variation: Individual and Team Level

This section reports on four measures, two of which measure the variation in rankings between individuals and two of which measure the variation across the team, pre- and post-discussion.

6.3.1.1 Cognitive Dissimilarity Pre-Discussion

Each individual was asked to rank order the four candidates in order of preference immediately after the individual computer search. There are 24 possible combinations of rank ordering, 22 of which were exhibited.

Across the whole of the sample, Malecka (candidate number 3) was ranked 1st by 46% of individuals. However, within that 46%, the rank order of the other candidates varied across all six remaining possible combinations (i.e. 3124 = 5.6%, 3142 = 7.2%, 3214 = 2.1%, 3241 = 6.4%, 3412 = 15.2%, 3421 = 9.6%).

Although several team members may have selected a particular candidate as their first choice, never was the rank order of all four candidates exactly the same for all members of a team. Of interest to hypotheses 1 and 3 is the degree to which individuals differed from each other. For example, team 12[2] contained three members. Their choices were as follows: 3124, 3421 and 1342. Clearly, the last member is the most dissimilar to the other two, as none of the preferences for candidates matched any of those of his fellow team members. However, although the first two team members each selected Malecka as their first choice, their opinions of the rank order of the remaining three candidates was not the same. The Euclidean distance measure computed standardized differences between all three members' rank order, then the row entries per individual were summed to give a single dissimilarity value for each team member. For the present case these (team 12), these were 2.29, 2.62, 4.26 respectively.

---

2 Teams used as illustrations in this section were all taken from the dataset at random.

A further example comes from team 2, which had six members. Their rankings and respective Euclidean distances are shown below:

Table 6.5 Pre-Discussion Rank-Ordering in Team 2 and Corresponding Distances per Individual

Individual	Rank Order	Euclidean Distance
1	1432	10.09
2	3412	6.99
3	4312	8.99
4	4312	8.99
5	2134	6.99
6	2134	6.99

The most cognitively dissimilar individual pre-discussion in this team was individual number 1, with a Euclidean Distance value of 10.09.

Cognitive dissimilarity pre-discussion took a minimum value of 0.02 and a maximum of 11.20. The mean was 5.06 and the S.D. was 2.32.

6.3.1.2 Cognitive Dissimilarity Post-Discussion

TMT demographic variation literature (as reviewed in Chapter 2) focuses its attention on the concept of cognitive variation pre-discussion. Yet, commonsense dictates that individuals' opinions will change after further information comes to light in a discussion, or opinions converge through exploration of the options. The PEAK Selection Simulation allowed for the testing of cognitive variation after a discussion. This has not been addressed in the literature before, so the overall predicted direction for H1 (i.e. that demographic and cognitive dissimilarity are positively associated) is assumed to extend to post-discussion ranking.

Subsequent to the team discussion, the Euclidean distance measure was used to compute the difference in individuals' ranking. This variable took values of 0.00 to 11.08. The mean was 4.24 and the S.D. was 2.82. The individuals who took values of 0.00 were the 3 men in team 12, described a few paragraphs earlier, who after the discussion each claimed that their personal ranking was now 1342. The individual who had this rank order pre-discussion happened to be the CEO (Managing Director) of the company. He may have been very influential in persuading his two peers to radically change their rankings to kowtow to his minority opinion, as prior to the discussion their rankings were more similar to each other than to his. It is outside the scope of this study to determine whether certain individuals are more or less influential than others within TMTs. What should be stressed here is that each individual was asked to privately record in writing their post-discussion ranking, without reference to other persons. There was no need for anyone to be less than truthful as to their private ranking as only the researcher had access to them.

That cognitive variation post-discussion is a valid and different concept to study can be illustrated by the change in team number 2 referred to in the preceding table. Table 6.6 shows that there is still variation in individuals' rankings, but that it is now individual number 5 who is the most dissimilar to his peers post-discussion, whereas it was individual number 1 pre-discussion.

Table 6.6 Post-Discussion Rank-Ordering in Team 2 and Corresponding Distances per Individual

Individual	Rank Order	Euclidean Distance
1	1432	8.11
2	3412	8.11
3	4132	9.65
4	4132	9.65
5	1234	9.96
6	1432	8.11

6.3.1.3 Cognitive Diversity Pre-Discussion

In order to achieve compatibility with the demographic diversity data, it was desirable to compute cognitive diversity using Blau's (1977) index of heterogeneity, which measures the proportional representation on non-ordered categorical data. In order to use this index for cognitive diversity, each of the 24 possible rank orders was assigned to a category (e.g. 1234 = 1, 1243 = 2 etc) then the proportional variation of ranks across each team was computed. Blau's index takes values between 0.00 (complete homogeneity amongst the team) and 1.0 (total diversity). This variable is referred to throughout this and succeeding chapters as cognitive diversity pre-discussion, and took values between 0.44 and 0.86, with a team mean of 0.73 and an S.D. of 0.11. Following the 'value-in-diversity' mantra, greater demographic diversity is expected to lead to greater ranking diversity.

6.3.1.4 Cognitive Diversity Post-Discussion

Cognitive diversity post-discussion was computed in the same way as cognitive diversity pre-discussion. Cognitive diversity post-discussion took values ranging from 0.00 (team 12) to 0.82, the mean was 0.65 and the S.D. was 0.17.

6.3.2 Cognitive Cohesion: Proximity to the Team Consensus

An interesting avenue for investigation is the proximity of the individuals' rankings to the team consensus ranking. Team 2 (whose individual rankings pre- and post-discussion are illustrated in the preceding tables) came to an agreed consensus decision ranking during its team discussion of 4132. What is demonstrated by Table 6.5 is that prior to the discussion, 2 of the 6 members (individuals 3 and 4) were closest to the ranking the team would choose, with a ranking of 4312, whilst 4 of the members had less proximity.

The measurement of proximity to the team consensus in this study is called cognitive cohesion. At the individual level, cognitive cohesion dissimilarity, was measured at two timepoints, pre- and post-discussion. At the team level, cognitive cohesion diversity was also measured pre- and post-discussion.

6.3.2.1 Cognitive Cohesion Dissimilarity Pre-Discussion

Cognitive cohesion dissimilarity pre-discussion is the Pearson correlation coefficient between the individuals' pre-discussion rank order and the team consensus ranking, and ranged from -1.00 to 1.00 (where 1.00 = complete agreement) The mean was 0.41 and the S.D. was 0.46. Following the 'upper echelons' rubric, one would expect a negative relationship between demographic dissimilarity and individual proximity to the team consensus prior to discussion. This is because dissimilar individuals are supposed to have the least conformist ideas.

6.3.2.2 Cognitive Cohesion Dissimilarity Post-Discussion

The proximity of individuals' rank order post-discussion was correlated with the team consensus ranking. What is demonstrated by Table 6.6 above is that although

Team 2 ostensibly made a consensus decision to which the whole team publicly signed up, only two of the team members privately held this same opinion.

The degree to which individuals' rankings match that of their team consensus ranking is referred to as cognitive cohesion dissimilarity post-discussion. The index ranged from -0.80 to 1.00. The mean was 0.62 and the S.D. was 0.46.

As has already been argued in preceding paragraphs, demographically dissimilar persons are unlikely to change their personal opinions simply to match those of their peers. One might logically assume then, that dissimilar individuals, despite having been part of a team discussion out of which a team consensus has emerged, may still hold divergent views. Therefore, a negative relationship between demographic dissimilarity and cognitive cohesion dissimilarity post-discussion is expected. This is not an additional opposing, hypothesis, merely a clarification of the direction that can be expected with this particular variable in order to support hypothesis 1 that demographic dissimilarity will be positively associated with cognitive dissimilarity. Moreover, it is expected that the association between cognitive cohesion dissimilarity pre- and post-discussion and procedural rationality and frequency of meetings will be positive, whilst the association with reflexivity and psychological safety and this response variable will be negative. That is to say, the former processes will facilitate more agreement (that is, more individuals tending towards a complete match of 1.0) whilst the latter processes will mean more individuals tending towards the least conforming (-1.0) rankings.

6.3.2.3 Cognitive Cohesion Diversity Pre-Discussion

Of special interest at the team level is whether or not demographic diversity predicts diversity in cognitive cohesion. In other words, are teams which are characterised by greater demographic variation also characterised by a larger proportional spread of individuals' proximity to the team consensus? In line with 'upper echelons' theory (Hambrick & Mason, 1984), and as hypothesised in Chapter 5, one would expect a positive relationship.

As the base measure (cognitive cohesion dissimilarity pre-discussion described in section 6.3.2) was already a correlation of individuals' pre-discussion ranking and the team consensus ranking (on a scale of -1.0 to 1.0), it can be treated as a continuous variable, for which a coefficient of variation is the most appropriate measure of diversity within teams. As previously described, the median coefficient of variation of cognitive cohesion dissimilarity pre-discussion was used to measure cognitive cohesion diversity pre-discussion (the proportional spread of the proximity of members' ranks to the team's rank, computed as the S.D./median). For the 23 teams this index took values of 0.52 to 1.95.

6.3.2.4 Cognitive Cohesion Diversity Post-Discussion

As per the previous section, the median coefficient of variation of cognitive cohesion dissimilarity post-discussion was used to measure cognitive cohesion diversity post-discussion. The index ranged from 0.63 to 1.11.

6.3.3 Cognitive Change: Differences Pre- and Post-Discussion

By referring to the rank order column in both of the preceding tables, it can be seen that between pre- and post-discussion ranking, five of the six individuals in team 2 changed their minds. That is, their rank order before discussion is not the same as that after. This variation in cognition was captured by creating a set of dummy variables for the ranked position of each candidate at each time point. The proximity between the two sets of dummy variables was computed using the proximity (Pearson) correlation between individuals pre- and post-discussion ranking. The scores were then standardized on a range of -1 to 1, where -1 = complete change of ranking and 1 = no change (i.e. complete agreement between individuals' pre-and post-discussion ranks). The resultant variable is referred to throughout this chapter and the analysis in Chapters 7 through 10 as cognitive change dissimilarity.

As the link between demographic dissimilarity and cognitive change dissimilarity has not previously been directly addressed in the literature, it is not straightforward to predict one way or the other the expected direction of the relationship. One might, for example, argue that demographically dissimilar individuals are likely to bow to peer pressure and change their rankings in order to blend in more, meaning that the direction would be negative. The rationale is that the odd one out demographically will not want to be the odd one out in terms of opinion or, in this case, ranking, and so will change to conform. However, anyone with experience at this level in organizational settings will appreciate that TMT members often have to push through decisions in their respective domains that may or may not be popular. Indeed, one might reasonably expect a positive relationship between the demographic dissimilarity variables and cognitive change dissimilarity (as no change attracts higher values), based on the premise that those who had the most dissimilar rankings prior to the discussion are not likely to shift radically, thereby holding on to their opinion. The latter is more in keeping with commonsense, and hence is ascribed to here. Moreover, inferential support can be drawn from Jackson et al. (1991), who found that dissimilar individuals were more likely to leave the team than to try to blend in.

6.3.3.1 Cognitive Change Dissimilarity

Cognitive change dissimilarity ranged from -0.80 to 1.00, the mean was 0.79 and the S.D 0.42.

6.3.3.2 Cognitive Change Diversity

The median coefficient of variation was used to compute cognitive change diversity. Cognitive change diversity (the quotient of changed ranks within the team and computed as the S.D./median), for the 23 teams took values of 0.79 to 1.94.

Extrapolating from individual level arguments concerning cognitive change dissimilarity and demographic dissimilarity in the previous section, the posited direction for cognitive change diversity and demographic diversity is negative. This is because one would expect that a group of demographically diverse people would be less likely to change the spread of their ranking dramatically.

6.3.4 Summary of Cognitive Variation Variables

This section has explained how distance measures were applied to the cognitive variation data to determine five individual cognitive dissimilarity variables (using Euclidean Distance and Pearson proximity correlations), and the application of diversity indices at the team level (Blau (1977) for two categorical variables, and the median coefficient of variation for three continuous variables). Table 6.7 presents the descriptive statistics for the cognitive variation variables. The next section describes how team processes were measured.

Table 6.7 Descriptive Statistics: Cognitive Variables

Variable	Min	Max	Mean	Median	S.D.
Dissimilarity					
Cognitive dissimilarity pre-discussion	0.02	11.20	5.06	4.60	2.33
Cognitive cohesion dissimilarity pre-discussion	-1.00	1.00	0.41	0.60	0.55
Cognitive change dissimilarity	-0.80	1.00	0.79	1.00	0.42
Cognitive dissimilarity post-discussion	0.00	11.08	4.24	3.96	2.83
Cognitive cohesion dissimilarity post-discussion	-0.80	1.00	0.62	0.80	0.45
Diversity					
Cognitive diversity pre-discussion	0.44	0.86	0.73	0.75	0.11
Cognitive cohesion diversity pre-discussion	0.52	1.95	0.94	0.93	0.26
Cognitive change diversity	0.79	1.94	1.02	0.98	0.22
Cognitive diversity post-discussion	0.00	0.82	0.65	0.72	0.17
Cognitive cohesion diversity post-discussion	0.63	1.11	0.93	0.99	0.14

n = 120 - 130 for dissimilarity variables depending on missing values
n = 23 for diversity variables

6.4 Measuring Team Processes

Four team processes were measured: procedural rationality, frequency of team meetings (as a proxy measure for communication), reflexivity and psychological safety. Frequency of team meetings was measured by self report. The other three team processes under consideration, procedural rationality, psychological safety and reflexivity, were measured by independent observers watching the videotapes of the team discussion. This section is split into eight sub-sections. The first describes the self-report measure, frequency of team meetings. The next section describes the development of the observational coding guide as this was used to measure the other three team processes. Addressed third is the training of independent observers to use the guide. The fourth section gives illustrations of text and behaviour that would be scored using the guide, and the fifth section reports on the reliability of the raters. The final three sections individually address the observed processes, procedural rationality, reflexivity and psychological safety.

6.4.1 Frequency of Team Meetings

Frequency of team meetings was selected as a proxy measure for communication. This is in keeping with Smith et al. (1994), Clark et al. (1997) and Knight et al. (1999). One of the questions on the self-report questionnaire that individuals were asked to complete regarding their demographic data, included the question:

How many times per month does your entire team formally meet?

In order to aggregate the self report questionnaire responses of individual team members to the team level of analysis, one has to be able to demonstrate agreement amongst team members in their responses (James, Demaree & Wolf, 1984; George, 1990; Kozlowski & Hattrup, 1992). To determine the extent to which there is high perceptual agreement within teams with respect to the number of formal meetings per month, the inter-rater reliability for groups ($^{r}wg_{(1)}$) was calculated following James, Demaree & Wolf (1984). This calculation involves averaging the response values across the twenty-three teams and then dividing by the mean for each team. The distribution of the frequency of team meetings item revealed no significant skew, meaning that parametric analysis was appropriate. The within group inter-rater reliability value ($^{r}wg_{(1)}$) was .73, which is considered to be acceptable, as a value of .70 is satisfactory.

A second test was applied in order to determine the justifiability of aggregation of individual scores into team measures. This was a one way ANOVA of the scales, for which the F ratio must exceed a value of 1.0 in order to effectively discriminate between teams according to Hays (1981). With respect to frequency of team meetings, the F ratio was 3.70 $p<.01$, considerably higher than the required 1.0. The conditions justifying aggregation of the data to team level are satisfied on both counts and means that comparison within and between teams is appropriate.

Once aggregated to the team level, frequency of team meetings ranged from 2 to 6 times per month, the mean was 3.50 and the S.D 1.10.

The three remaining team processes were measured by observer ratings. The discussion now turns to describing the instrument and its reliability and validity.

6.4.2 Development of the Observational Coding Guide

Interaction coding systems are designed to count the frequencies of specific behaviours whilst observing people in specific situations. Bales (1950) for example, has nine categories for measuring interactions between team members to do with the task at hand. Other similar systems (see for example, Morris, 1966) categorise communication between group members – i.e. "repeats", or "seeks evaluation". Still others count speaking turns and specific words or mentions of particular items (Weingart, 1997).

By way of contrast, Hackman & Morris (1975) argued that speech pattern guides are useless in research aimed at understanding group processes and effectiveness. Moreover, they argued that:

".. coding systems are needed that derive directly from conceptual propositions about those aspects of group interaction that are crucial in determining group effectiveness for various kinds of group tasks. The content of such theory-based systems, it is argued, would be substantially different from that of most existing systems and would more clearly reveal just what goes on in groups to sometimes facilitate group effectiveness and sometimes impair it" (Hackman & Morris, 1975, p 13).

Following Hackman and Morris (1975) and Weingart (1997), the observational coding guide used in this study sought to understand what was going on in the teams in terms of procedural rationality, reflexivity and psychological safety. The small number of items offset the complexity inherent in trying to observe the verbal and non-verbal cues of between 3 and 8 people interacting with one another. As there were fewer items than on other coding guides, more time could be spent actually observing and listening rather than making notes.

6.4.3 Training the Raters to use the Guide

Two post-graduate students, each studying for an MSc in Occupational Psychology at the Institute of Work Psychology, agreed to be blind raters for the study. In order not to compromise or prejudice the raters' judgements of the actual teams while they were becoming familiar with the behavioural coding guide, the Center for Creative Leadership (CCL) supplied 4 videotapes for training purposes. The CCL tapes were of synthetic teams engaged in the PSS discussions. Although not directly comparable to the 23 teams in the present sample, the content of the tapes was sufficiently similar to provide full training in using the behavioural guide, a copy of which is in Appendix B.

The procedure for training was as follows: the author gave an overview of the PSS, presented the behavioural coding guide and explained the concepts and ontology underpinning the process scales to the two raters. A range of example cues (see the excerpt transcripts in a following section) derived from experience of working with the teams were suggested to the raters by the author.

The raters and the author then watched the first training video tape together, coding and discussing the team's behaviour. The next 3 video tapes were coded independently by the two raters. Subsequently, the author and the 2 raters met to compare results, and watch the training tapes again together to obtain agreement as to what constituted procedural rationality, reflexivity and psychological safety. The training described here was very much in line with suggested best practice for training people to use coding guides for observational studies of teams (Weingart, 1997).

The layers of conversation are so interwoven in the team discussions that an example of reflexive behaviour can also embody psychological safety, as will be shown in the excerpts below. Non verbal cues were also important to the raters in judging the behaviour of the teams. However, the list of positive and negative non-verbal behaviour would be too long to document and code, and could represent a whole study in its own right. The raters were instructed to record their sense of the extent to which these real top management teams were reflexive, and the extent to which the members appeared to enjoy psychological safety. The discussion being observed and rated went on for 60 minutes and the raters were looking for supporting evidence throughout the entire discussion, not specific instances of pre-determined textual references or specific non-verbal cues.

With practice on the training tapes, it became easier to spot who was the most dominant person in the group, how people reacted to one another, whether or not a team was reflexive, and to sense the tension or lack of it that would indicate levels of psychological safety.

Subsequent to the training described above, each of the videotapes of the 23 teams in the sample was viewed and coded by the two raters independently of each other. It took 3 hours for each rater to code each video tape, involving 6 person hours of coding per tape, approximately 180 hours in total[3]. It is reported that many researchers avoid the development of observational coding guides due to the excessive amount of time necessary to develop and refine their use and then apply them in practice (Weingart, 1997).

6.4.4 The Guide in Practice

Indicative statements of a team engaging in procedural rationality and reflexively planning and challenging its approach to making its decision for the PSS might be observed at the beginning of the team discussion as the following example shows:

CEO: *"How are we going to do this?"*

QUALITY DIRECTOR: *"I suggest that we go round the table individually and say who our first choice was for the job. Then if we all agree, we don't need to discuss it anymore. (Laughs all around table).*

CEO to HR DIRECTOR: *"That's one alternative. Is that what you would recommend Paula?"*

HR DIRECTOR: *"I think we need to go back to the Job Description and Person Specification, and remind ourselves of the key criteria we are appointing for, then we can objectively evaluate the candidates against the criteria, ensuring we get the right person for the job".*

CEO: *"Does everyone else agree with Paula's assessment of what to do?" (Nods of assent). OK John, would you please summarise the key skills and abilities that we are looking for on the whiteboard?"*

OPS DIRECTOR: *"Sure thing. Before I do that, would it also be a good idea for us to weight the criteria too? We could divide them between "must have" and "desirable". Then we'll have a baseline to judge the candidates by" (other team members agree and discussion follows this pattern for a while)*

HR DIRECTOR: *"If you put the candidates' initials down the left column and the criteria along the top, we'll end up with a matrix that should easily differentiates between candidates and criteria".*

The series of exchanges outlined above were typical of teams engaging in procedural rationality. The items on the procedural rationality scale were assessing

---

[3] Including the 4 training tapes

the extent to which the TMT looked for, and systematically analysed the information before making the decision. So for example, in the transcript above, the team decided on a procedure to follow. They then went on compare the knowledge, skills and abilities of the candidates against their criteria in the job description and to use a matrix system to quantitatively analyse the candidates' suitability for the job.

Closely related to these aspects of procedural rationality are feature of teams that were more reflexive. That is, before making the decision, they devised a plan through discussion as to the process by which they would reach a decision. By way of contrast, less-reflexive team members engaged with each other in way that was reminiscent of family discussions around the dinner table. They simply launched into discussing their individual preferences for the candidates, perhaps with the CEO or another director sending the conversational ball into the scrum with a statement like: *"Well, Malecka is the man as far as I'm concerned"*. The rest of the team members would all then pitch in, freely offering their opinions, jesting, disagreeing or arguing in a polite, often pleasant and enjoyable way but which lacked the task strategy or purposeful direction displayed by the reflexive teams.

However, for a team to score highly on the reflexivity scale it was not enough to work out a strategy at the beginning of the task. The observational raters were also looking for instances of teams monitoring their progress and process during the allotted time, or coming up with a new approach if they felt that the initial one was no longer appropriate for any reason. The following example illustrates how a team might monitor itself during the discussion:

CEO: *"OK, now we have the matrix, with the must haves and the desirable requisites. Do we all agree that the requisites are the right ones and that they are correctly labelled as must haves or desirables? (Nods of agreement and assent all round). Right, moving on….., let's take each of the candidates in turn, starting with Davidson. Any comments on Davidson?*

QUALITY DIRECTOR: *'My sense is that we should still continue to be as objective as possible, comparing the candidates fairly across the information provided. (others agree). I propose that we discard the unsolicited opinions". (Pause as the rest of the team consider this.)*

OPS. DIRECTOR: *'No, I disagree with that, we normally get as many informal referrals as we can for a top management team appointment. I do agree that some of the more sensational of the unsolicited opinions should not be privileged, but I would not be happy discarding them all together". (Shakes his head for emphasis and looks at each one of them in turn as if anticipating what they will say in response).*

HR DIRECTOR: *"Why don't we start with the search firm report, after all, we have paid them a considerable sum to do the initial sifting on the pre-requisites for us? Then we can move on to our assessment of the interviews, supplementing our comments with what we know from the other sources of information at our disposal. I would suggest that the external opinions will only serve to complement what we have already decided".*

CEO: *"Good idea. Right. Davidson. Anything significant in the Search Firm Report on Davidson?" (Rustle of papers as team members find the relevant notes they made about Davidson whilst accessing the information during their individual PSS computer search).*

Another indicator of reflexivity is a team double checking at the end of the discussion that it has made the right decision, perhaps asking itself whether there was a better way to approach the task, or critically appraising the choice, e.g.

"What if our first choice candidate refuses the offer, are we sufficiently confident that our second choice meets the criteria adequately, or do we need to go back to the search firm to assemble another candidate pool?"

The transcription above, which highlights behaviour indicative of procedural rationality and reflexive behaviour, also reveals much about psychological safety. One of the items in the observational coding guide referring to psychological safety is "Team members seemed to accept and understand one another" and a second is "Everyone listened to each other's ideas, even if it was a minority opinion".

In the transcription, evidence of behaviour matching these items is revealed. That is, the Quality Director who must feel psychologically safe in the top management team environment, appears to contradict the direction given by the CEO, and then goes on to make a bold assertion that the team discard a whole piece of information. Indicating that the rest of the team similarly feel just as safe, the Operations Director makes a counter suggestion by saying "I disagree". The other team members consider both suggestions, then the HR director makes a third suggestion that they all agree to. What cannot be conveyed through the bald transcription of the text is the tone in which the words are spoken, the subtle nuances of body language, eye contact with some or all members of the team, individual gestures, grimaces and asides that occur in real time whilst these exchanges are taking place. If the transcribed text had for example taken place in a team where hostile glances had been exchanged; or after being rebuffed by the Operations Director, the Quality Director had crossed his arms and tuned out of the conversation; at which point the HR Director had tentatively and persuasively put forward her suggestion as a way of diffusing the tension in the room; the CEO sighed and said "Moving right along"; that same exchange could have been interpreted as demonstrating low levels of psychological safety.

6.4.5 Reliability of the Guide

The face validity of the observational coding guide is evident from the above, but it is important to establish that it can be applied reliably. For this purpose, intraclass correlations were examined to determine the level of correspondence between the independent raters as to their observations of reflexivity and psychological safety. These are reported in Table 6.8 below.

Intraclass correlations can be looked at in two ways, the first (ICC1) is used to determine how well the raters scores covary. The second (ICC2) is used to determine the degree of absolute agreement between the raters. The values were identical for ICC1 and ICC2 (this is common when there are only 2 raters), being 0.85 for procedural rationality, 0.90 for reflexivity and 0.93 for psychological safety. This is well above the acceptable level for such intraclass correlations, which is 0.70 (Howell, 2001).

6.4.6 Procedural Rationality

The five items used to measure procedural rationality were taken from the self report scale published by Dean & Sharfman (1996). The items were kept in their entirety, as they were already cast as past tense items in relation to a specific decision by a TMT. The five items were:

1. How extensively did the TMT look for information in making this decision?
2. How extensively did the TMT analyse relevant information before making a decision?
3. How important were analytical techniques in making the decision?
4. How would you describe the process that had the most influence on the team's decision?
5. In general, how effective was the TMT at focusing its attention on crucial information and ignoring irrelevant information?

A seven-point Likert-type scale for the behavioural coding guide, with anchors of 1 "to a very little extent" to 7 = "To a very great extent" was used to code each of the items. For item 4, the anchors were 1 "completely analytical" to 7 "completely intuitive", reverse coded. As reported in Table 6.8, the scale alpha coefficient was 0.84, higher than the acceptable 0.70 (Howell, 2001).

Procedural rationality ranged from 2 to 5.50, the mean was 3.70 and the S.D 1.08.

6.4.7 Reflexivity

Reflexivity was measured using a derivation of items from the Team Climate Inventory (Anderson & West, 1996). The reliability and validity of the original scale of eight self report items has been well attested in several studies (see; Swift & West, 1998; Carter, 2000). However, the items are broad, universal statements about which team members are asked to give a general opinion, for example, *"This team often reviews its approach to getting the job done".* Such statements were not suitable for the purposes of observation of a discrete decision. The eight items in the original self report questionnaire were condensed to four for the observational coding guide, consistent with a more definitive understanding of how reflexivity would manifest itself in a team discussion (see Swift & West, 1998). The four observational items used in this study were:

1. The TMT challenged the rationale for the task.
2. The TMT challenged its approach for task accomplishment.
3. There was disagreement over what the objectives for the task should be.
4. The TMT critically appraised weakness in the proposed solutions.

A seven-point Likert-type scale for the behavioural coding guide, with anchors of 1 "to a very little extent" to 7 = "To a very great extent" was used to code each of the items. The alpha coefficient for this scale was 0.78. Reflexivity ranged from 1 to 2.25, the mean was 1.40 and the S.D was 0.43. The range for this variable appeared to be quite restricted, and one might consider, that the task is too constraining for there to be any demonstration of reflexivity. However, the distribution was normal for the purposes of parametric analysis and as will be shown in later chapters, one of the largest effect sizes achieved in the study concerned reflexivity, suggesting that there was sufficient variation between teams.

6.4.8 Psychological Safety

Four items were used to measure psychological safety. These were adapted from 12 items used to measure processes in TMTs in 27 hospitals (Anderson & West, 1996). The 12 original items contained broad statements to which the teams in Anderson & West's (1996) study were asked to reflect in general about their team. For this study, the items needed to be specific to the discussion being observed. The 12 self report items were condensed to the following four observational items:

1. Team members seemed willing to freely share ideas with one another.
2. Team members seemed to accept and understand one another.
3. Everyone listened to each other's ideas, even if it was a minority.
4. Team members freely shared information about the decision with each other.

A seven-point Likert-type scale for the behavioural coding guide, with anchors of 1 "to a very little extent" to 7 = "To a very great extent" was used to code each of the items. As reported in Table 6.8, the alpha coefficient attesting to the reliability of the 4 item scale used above was 0.92. Psychological Safety ranged from 2.50 to 5.00, the mean was 3.46 and the S.D was 0.66.

6.4.9 Summary of Team Process Variables

This section has explained how frequency of team meetings was measured using a self-report question that was then aggregated to the team level of analysis. Three further processes were measured using a behavioural coding guide which was both reliable and internally consistent. Table 6.8 presents the descriptive statistics for the team process variables.

The next section describes how decision belief was measured.

Table 6.8 Descriptive Statistics Team Process Variables

Variable	Min	Max	α	Mean	S.D.	ICC	$r_{w(g)j}$
Frequency of Team Meetings	2.00	6.00	n/a	3.50	1.10	n/a	.73
Procedural Rationality	2.00	5.50	0.84	3.70	1.08	.85	n/a
Reflexivity	1.00	2.25	0.78	1.40	0.43	.90	n/a
Psychological Safety	2.00	5.00	0.92	3.46	0.66	.93	n/a

n = 23

6.5 Measurement of Decision Belief

In order to judge the efficacy of decision making during the simulation, decision belief was measured on 3 dimensions. Using a self report questionnaire at the end of the PSS team discussion, participants were asked to rate on a 5-point Likert type scale: their personal satisfaction with the decision-making process; their personal level of confidence in the decision made; and their perception as to how effective their team would be in future decision-making tasks.

Gathering these responses immediately after the team had come to a collective decision meant that the current study did not suffer from the methodological flaws inherent in retrospective studies such as poor or subjective recall mentioned earlier in this chapter.

As none of the correlation coefficients between the items was above (or even approaching) .70 which would indicate multicollinearity, it was deemed appropriate to keep all variables in the analysis as individual items.

As noted earlier, in order to aggregate the self report questionnaire responses of individual team members to the team level of analysis, one has to be able to demonstrate agreement amongst team members in their responses (James, Demaree & Wolf, 1984; George, 1990; Kozlowski & Hattrup, 1992). The inter-rater reliability for groups (r_{wg_1}) were calculated following James, Demaree & Wolf (1984) and one way ANOVAs were performed to determine F ratios, which must exceed a value of 1.0 in order to effectively discriminate between teams (Hays (1981). These are reported in Table 6.9 and discussed below.

6.5.1 Satisfaction

As reported in Table 6.9, the conditions justifying aggregation of the data to team level were satisfied on both counts, $^{r}wg_1$ was .83 and the F ratio was 1.70, p <0.05. This means that comparison within and between teams is appropriate. Once aggregated to the team level, satisfaction ranged from 2.67 to 4.67, the mean was 3.75 and the S.D .51.

6.5.2 Confidence

As reported in Table 6.9, the conditions justifying aggregation of the data to team level were satisfied on both counts, $^{r}wg_1$ was .78 and the F ratio was 1.78, p <0.05. Once aggregated to the team level, confidence ranged from 3.00 to 4.50, the mean was 3.86 and the S.D .57.

6.5.2 Effectiveness

As reported in Table 6.9, the conditions justifying aggregation of the data to team level were satisfied on both counts, $^{r}wg_1$ was .79 and the F ratio was 1.83, p <0.05. Once aggregated to the team level, confidence ranged from 3.63 to 4.75, the mean was 4.09 and the S.D .31.

6.5.3 Summary of Decision Belief Variables

This section has explained three items were used to measure discrete aspects of decision belief. The data were captured using a self-report questionnaire that was then aggregated to the team level of analysis after establishing inter-rater reliability amongst team members. Table 6.9 presents the descriptive statistics for the decision belief variables.

Table 6.9 Descriptive Statistics Decision Belief Variables

Variable	Min	Max	Mean	S.D.	$^{r}wg_{(1)}$
Satisfaction	2.67	4.67	3.50	1.10	.83
Confidence	3.00	4.50	3.70	1.08	.78
Effectiveness	3.63	4.75	1.40	0.43	.79

n = 23

6.6 Conclusion

This study measured demographic variation and cognitive variation at the individual (dissimilarity) and team (diversity) levels. Although standard dissimilarity and diversity indices have been used for decades with regard to demographic attributes, there is no accepted standard measurement of cognitive variation in the extant literature. Hence this study has applied the same dissimilarity measures and diversity indices to cognitive variation. Thus demographic and cognitive variation can be compared in this study with a precision previously lacking in other studies.

Team processes were measured using a bespoke observational coding guide, which proved both valid and reliable when used with the 23 videotapes by two independent raters. Decision belief was measured by self-report questionnaire administered to the teams immediately after the discussion. Table 6.10 below summarises each of the measurements, type of instrument and the time-points within the research activity when each measurement was taken.

Table 6.10 Study Variables and Measurement Points

What Was Measured	When Measured & How	Component Variables
1. Demographic Variation (a) Individual Dissimilarity (b) Team Diversity	Self report via questionnaire immediately prior to participating in the research.	Age Educational Attainment Functional Background Gender Tenure – Team Euclidean distances for dissimilarity and Blau's index and Stride index for Diversity
2. Cognitive Variation (a) Individual Dissimilarity (b) Team Diversity	Immediately after the individual computer search, ranking captured on screen. Subsequent to team decision-making task via self report questionnaire Team decision recorded by team appointed scribe on self-report form.	Individuals' pre-discussion ranking. Individuals' post-discussion ranking Team consensus ranking Euclidean distances for dissimilarity and Blau's index and the median coefficient of variation for diversity
3. Team Processes	(a) Self report questionnaire (b) Observed during team decision-making by raters via videotape	(a) Frequency of team meetings (b) Procedural Rationality 5 items (b) Reflexivity 4 items (b) Psychological Safety 4 items
4. Decision Belief	Immediately subsequent to team decision-making task via self report questionnaire.	Decision Satisfaction Decision Confidence Perceived Effectiveness

CHAPTER 7

Findings: Hypotheses 1 - 2

7.0 Overview

This chapter presents the findings for the first two hypotheses, dealing with the relationship between demographic and cognitive variables at the individual and team levels (as shown in the guiding conceptual model, p 76).

7.1 Introduction

Hypothesis 1 (H1) focuses on the individual level of analysis, predicting a positive relationship between demographic dissimilarity and cognitive dissimilarity. That is to say, it proposes that it will be those persons demographically most dissimilar to their top management team colleagues in terms of age, functional background, educational attainment, gender and tenure, who will show the greatest dissimilarity from those colleagues in their rank ordering of the candidates.

Hypothesis 2 (H2) is concerned solely with the group level of analysis, proposing that team demographic diversity will be positively related to team cognitive diversity. In other words it is predicted that those teams characterised by greater variation in age, functional background, educational attainment, gender and tenure, will also show the greatest variation across members (i.e. the least agreement) in how they rank-order candidates.

For each of the hypotheses, statistical considerations are presented first, followed by zero-order correlations and multiple regression analyses. Within each relevant section, the findings for two pre-discussion measures of cognitive variation are reported first, followed by those for cognitive change, and then those for the two post-discussion measures.

7.2 The Relationship Between Individual Level Demographic and Cognitive Dissimilarity (H1)

7.2.1 Statistical Considerations

For H1, for which the individual is the unit of analysis, the sample comprises 130 senior executives. Examination was made of the underlying distributions of the individual dissimilarity measures in order to determine appropriate analytic methods (Hayduk, 1987). All of the measures except one were sufficiently normally distributed to allow the use of parametric statistics (i.e. skew < 1.0, Kurtosis < 3.0, see Howell, 2001). The exception was gender dissimilarity, which was significantly positively skewed (S = 1.91) with the kurtosis (K = 2.96) close to the acceptable limit. Thus transformations of the measure were investigated. Given the mean was proportional to the variance rather than the SD, the most appropriate normalisation method was a square root transformation. This brought the skew and kurtosis down to acceptable levels (S = .866, K = .502) whilst maintaining the overall integrity of the data. In the analysis that follows the transformed variable is used.

Initial testing of hypothesis 1 involved determining the extent to which each of the five demographic dissimilarity variables is related to the five indices of cognitive dissimilarity through zero-order correlations.

7.2.2 Zero-order Correlations

The findings for the zero-order correlations are presented in Table 7.1. The first pre-discussion variable measures the relative distances between individuals' rankings within the team. With respect to cognitive dissimilarity pre-discussion, 3 of the 5 demographic dissimilarity variables show a positive relationship as predicted. These are dissimilarity in functional background ($r = .21, p<.05$), educational attainment ($r = .21, p<.05$) and team tenure ($r = .15, p<.05$).

These findings mean that it is those individuals who differ most on functional background, educational attainment and team tenure from their team colleagues who are most likely to differ from them in their initial rank-ordering of the candidates. In more general terms, it was the most demographically dissimilar individuals on these traits that were the most cognitively dissimilar to the rest of their team prior to the team discussion.

Table 7.1 Zero-order Correlation Coefficients for Demographic Dissimilarity and Cognitive Dissimilarity

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Age dissimilarity	-									
2. Functional dissimilarity	.09	-								
3. Educational dissimilarity	.36**	.40**	-							
4. Gender dissimilarity[a]	.13	.23**	.14	-						
5. Tenure dissimilarity	.16*	-.01	-.01	-.11	-					
6. Cognitive dissimilarity pre-discussion	.08	.21*	.21*	-.01	.15*	-				
7. Cognitive cohesion dissimilarity pre-discussion	-.16*	-.05	.03	-.02	-.07	-.36**	-			
8. Cognitive change dissimilarity	-.05	.08	.15*	.23**	-.12	-.13	-.32**	-		
9. Cognitive dissimilarity post-discussion	.11	.25**	.34**	.07	-.04	.73**	.05	.51**	-	
10. Cognitive cohesion dissimilarity post-discussion	-.10	-.12	-.15*	-.16*	.03	-.29**	-.42**	-.19*	.67**	-

N = 120-130; * *p* <0.05; ** *p* <0.01 (1 tailed).
[a] this variable was subjected to square root transformation
shaded areas show relationships of interest

Table 7.2 Multiple Regression Analyses for the Effects of Demographic Dissimilarity on Cognitive Dissimilarity

Outcome Variable	Predictor Variable	β	$R^2 \Delta$	*F*	*df*
Cognitive dissimilarity pre-discussion	Tenure Dissimilarity	.13	.02	3.0*	3,113
	Functional Background Dissimilarity	.11	.03		
	Educational Attainment Dissimilarity	.15	.02		
Cognitive change dissimilarity	Gender Dissimilarity	.20**	.07	4.6*	2,115
	Educational Attainment Dissimilarity	.12	.00		
Cognitive dissimilarity post-discussion	Functional Background Dissimilarity	.13	.06	9.3**	2,114
	Educational Attainment Dissimilarity	.29**	.08		
Cognitive cohesion dissimilarity post-discussion	Gender Dissimilarity	-.12	.01	3.5*	2,118
	Educational Attainment Dissimilarity	-.13	.05		

N varies from 117 – 130 depending on missing values; * *p* <0.05; ** *p* <0.01. One-tailed tests of significance levels for standardised β coefficients in final equation. $R^2 \Delta$ refers to discrete steps.
No entry for cognitive cohesion dissimilarity pre-discussion as there was only one predictor variable.

The second pre-discussion measure, cognitive cohesion dissimilarity pre-discussion computes the proximity of individuals' pre-discussion rankings to the eventual team consensus decision. As explained in chapter 6 (p 123), in order to support the hypothesis a negative relationship between demographic dissimilarity and cognitive cohesion dissimilarity pre-discussion is expected. That is to say, demographically dissimilar individuals will not select the same ranking as the team consensus.

The proposition received some support in that age dissimilarity was negatively correlated with cognitive cohesion dissimilarity pre-discussion (r = -.16, $p<0.05$). That is, those who are most dissimilar in age to the rest of their team are the least likely to achieve the same pre-discussion rank order as that ultimately selected by the team. In other words, age dissimilar individuals are the most likely to hold a disparate pre-discussion rank order to the team consensus ranking. None of the other demographic dissimilarity variables was significantly associated.

The third cognitive variation measure assesses the match between individuals pre- and post-discussion ranking on a scale of -1.0 to 1.0 where 1.0 = a complete match. As explained in Chapter 6, it is expected that dissimilar persons will hold on to their rankings, which means that a positive relationship between demographic dissimilarity and cognitive change dissimilarity is expected.

A positive relationship, as predicted, was observed between gender and educational dissimilarity and cognitive change dissimilarity (r = .23, $p<0.01$, and r = .15, $p<0.05$, respectively). This finding suggests that women executives (as they attracted higher dissimilarity values), and those who are dissimilar to their peers in terms of educational attainment, are the least likely to change their rankings after team discussion.

With regard to cognitive dissimilarity post-discussion, two of the five demographic dissimilarity variables show the predicted positive relationship. These are dissimilarity on functional background (r = .25, $p<0.01$) and educational attainment (r = .34, $p<0.01$). This shows that the most dissimilar individuals with respect to function and education are the most cognitively dissimilar to their peers after a team discussion.

With respect to cognitive cohesion dissimilarity post-discussion (the extent to which individuals' rankings achieve proximity subsequent to the team consensus), two of the five demographic dissimilarity variables were significantly associated in the predicted negative direction. These are gender dissimilarity (r = -.16, p<0.05) and educational attainment dissimilarity (r = -.15, p<0.05). The finding for gender means that women (as they took higher values) are less likely than their male counterparts to privately hold the same post-discussion rank-order as the team consensus rank-order. The most dissimilar individual in terms of educational attainment is also less likely to privately hold the same post-discussion rank-order as the team consensus rank-order. Age dissimilarity and functional background dissimilarity were also found to be associated in the expected direction, but not statistically significantly so.

Multiple regression analyses were conducted to test these findings more stringently, by determining the effects of salient individual demographic dissimilarity predictor variables whilst holding the others constant. These are reported in the next section.

7.2.3 Multiple Regression Analyses (Individual Level) for H1

The results of the multiple regression analyses for H1 are shown in Table 7.2. For each of the cognitive dissimilarity dependent variables, the analyses were conducted using those demographic dissimilarity variables found to correlate with them statistically significantly in the zero-order analyses. All variables were entered together into the regression equation. Because of the directional nature of the hypotheses, the zero-order correlations used one-tailed tests of statistical significance, and to ensure equivalence, the p values for the beta weights in the regression analyses (which by convention are two-tailed), are halved (see Howell, 2001). Precedents for this include Parker (1998) and Rogelberg, Leach, Warr & Burnfield (2005).

Three dissimilarity variables (functional background, educational attainment and team tenure) were regressed onto cognitive dissimilarity pre-discussion and together explain a non-significant proportion of the variance (7%). The beta coefficients reveal that none of the predictors are any longer significant.

These findings mean that those individuals who are demographically dissimilar to their colleagues are not more likely choose a dissimilar ranking to their peers prior to discussion.

A further proposition concerning hypothesis 1 was that there would be a negative relationship between demographic dissimilarity and cognitive cohesion dissimilarity pre-discussion. Such a relationship was observed solely for age dissimilarity as shown by the zero-order correlation (Table 7.1). Given a single predictor, multiple regression was not appropriate. Age dissimilarity explains 2% ($p<0.05$) of the variance in cognitive cohesion dissimilarity pre-discussion.

In summary, the finding for the pre-discussion measures is that age dissimilarity predicts the furthest distance from the team consensus.

The results of the multiple regression analysis, shown in Table 7.2 also provide support for H1 that demographic dissimilarity would be positively associated with cognitive change dissimilarity (where higher values means zero change). Gender dissimilarity and educational attainment dissimilarity jointly had a significant association with cognitive change dissimilarity (8%, $p<0.01$), but this was due to the unique effect of gender dissimilarity ($\beta = 0.23$, $p<0.01$) which singly explained 7% of the variance, as educational attainment dissimilarity is no longer significant.

In summary, the finding for cognitive change dissimilarity is that it is women who are less likely to change their rankings after a team discussion.

With consideration to cognitive dissimilarity post-discussion, functional background dissimilarity and educational attainment dissimilarity together explained a significant proportion of the variance (14%, $p<0.01$). The Beta coefficients revealed a significant unique effect for educational attainment dissimilarity ($\beta = 0.29$, $p<0.01$) explaining 8%, but not for functional background dissimilarity, which was no longer significant.

With regard to the final expected negative relationship with cognitive cohesion dissimilarity post-discussion, gender and educational attainment dissimilarity explain a small proportion of the variance (6%, n/s), but the beta coefficients reveal that the effects of educational attainment dissimilarity and gender dissimilarity are not strong enough to achieve unique significance.

In summary, the findings for the post-discussion measures are that educational dissimilarity predicts cognitive dissimilarity post-discussion, demographic dissimilarity is not a predictor of cognitive cohesion dissimilarity.

The next section summarises the findings for hypothesis 1.

7.2.4 Conclusions Regarding Hypothesis 1

Predictions concerning the implications of demographic dissimilarity for cognitive dissimilarity receive qualified support. Significant zero-order correlations were found between demographic dissimilarity indices and the cognitive dissimilarity measures in 10 of the 25 instances examined (five demographic dissimilarity predictors x five cognitive dissimilarity outcomes). Of these, only three maintained statistical significance when subjected to more stringent testing. That so many relationships are no longer statistically significant with the application of one-tailed tests, suggests that the increase in the degrees of freedom nudges the observed relationships out of the tolerances to achieve statistical significance.

Multiple regression analyses demonstrate age dissimilarity explains 2% of the variance in cognitive cohesion dissimilarity pre-discussion[1], gender dissimilarity explains 1% of the variance in cognitive change dissimilarity, and educational dissimilarity explains 8% of the variance cognitive dissimilarity post-discussion.

The number (which is marginally greater than one would expect by chance alone), and strength of findings in this study are not dissimilar to other studies in this field. The restricted number of statistically significant results may simply be a question of power due to having a range of demographic dissimilarity predictors, an observation previously made by others (see for example, Jackson et al., 1991), who also record comparable levels of support for hypotheses. For example, Jackson et al., (1991) achieved four significant results from a possible 24 findings for a hypothesis concerning demographic dissimilarity and individuals' team turnover[2]. Despite the complexity of simultaneously examining a range of demographic predictors, they argue for the inclusion of an even wider array of demographic factors in order to authenticate and empirically test theoretical models such as 'upper echelons' theory. The present study adds to the TMT demographic research literature in significant ways. Substantial methodological improvements were made: cognitive dissimilarity was precisely defined and measured. Individuals' personal decisions (rankings) were captured in real time rather than retrospectively some time

---

1 As age dissimilarity was the only predictor, no regression analysis was conducted
2 Other hypotheses in Jackson's study observed 29% and 50% of possible associations.

after the fact. A comprehensive range of demographic dissimilarity attributes were simultaneously examined. Hence, the findings in support of H1 are meaningful.

7.3 The relationship between team level demographic and cognitive diversity (H2)

7.3.1 Statistical Considerations

H2 concerns the impact of demographic diversity on cognitive diversity, so the team is the unit of analysis. The sample comprises 23 teams. Cognisant that a reasonable effect size in such a relatively small dataset may only be marginally significant, whilst the same effect size across a larger N could attract higher significance levels, it was decided to extend the significance level reported for the 23 teams to $p<0.10$.

It seems obvious when studying demographic differences at the team level that the internal homogeneity of the data should be considered. Yet, with rare exception (see for example Jackson et al., 1991), studies of TMT diversity do not attempt to discover whether top management teams are clustered into groups that are more homogeneous on personal attributes than can be expected by chance. Nevertheless, this is important because it is well known that top managers select to their ranks people like themselves (Jackson et al., 1991; Boone et al., 2004).

Following Lord & Novick (1968) and James (1982), intraclass correlations (ICCs) were computed to estimate the proportion of variance in individuals' demographic traits which are accounted for by the differences in teams. The cited authors use the ICC to determine if responses to self-report scales are accounted for by the difference in teams, but the principle is exactly the same as that employed here. The ICC represents the extent to which the demographic variable values of individuals are more similar to members of their team compared to individuals in other teams across the sample. The higher the ICC value, the more support for the argument that personal attributes in top management teams are clustered together. Values above 0.10 suggest significant concentration within teams (James, 1982; Lord & Novick, 1968).

Intraclass correlation values shown in the table below demonstrate that teams are statistically significantly concentrated in terms of tenure (0.32, $p<0.05$), but not in the case of age, functional background, educational attainment or gender, which all take values 0.05 (n/s).

This means, with respect to the former, that there is no evidence to suggest that individuals are more similar to their team-mates than to others. Even in the case of tenure, which suggests that persons of similar tenure are in the same teams, almost 70% of the variance is still spread across various teams in the sample. The ICCs here demonstrate that there is sufficient variation in the dataset with regard to demographic attributes to test the hypothesis.

Table 7.3 Intraclass Correlation Coefficients for Demographic Traits

	Variance between individuals within teams		Variance between individuals across different teams		ICC	
Demographic Traits	β	(SEβ)	β	SEβ		p
Age	3.37	(5.71)	82.69	(11.45)	0.04	-
Functional Background	0.17	(3.45)	3.45	(0.47)	0.05	-
Educational Attainment	0.85	(25.54)	25.54	(3.73)	0.03	-
Gender	0.00	(0.01)	0.01	(0.01)	0.03	-
Tenure	127.47	(62.68)	445.85	(61.18)	0.32	0.05

N = 126 – 130 depending on missing values
N.B. tenure was measured in months.

A potential criticism that could be levied against the current research is that the sample size is too small for the statistical tests applied. This issue will now be addressed so as to establish the validity of the application of regression analyses at the team level in this and subsequent chapters.

The key question given the number of teams is how many cases are needed to robustly conduct a multiple regression analysis and demonstrate a large effect between variables? In order to demonstrate a small effect, more cases are usually needed. A rule of thumb promoted by Green (1991) and endorsed by Tabachnick & Fidell (2001) is expressed as $N \geq (8/f^2) + (m - 1)$ where f^2 = .01, .15, and .35 small, medium and large effect sizes respectively. If this rule of thumb was used, then the n of cases needed was 26 (8/.35) + (5-1). The n of 23 teams was still shy of this figure. Hence, these were entered as separate independent variables (together with a control variable), for which the equation results in 23 cases being required (8/.35) + (2-1). Therefore, regression analysis is an acceptable method of analysis for this dataset of 23 teams. In order to maximise the possibility of achieving statistically significant effects, it was decided to extend the significance level to a more lenient $p<0.10$ for the team level analysis. Because the hypotheses were directional and the zero-order correlations used one-tailed tests of statistical significance, the p values of the beta weights in the regression analyses which follow in this section are also one-tailed (Howell, 2001; see also p 144).

As for the analysis at the H1 level, examination was made of the underlying distributions of the measures in order to determine appropriate analytic methods (Hayduk, 1987).

The distributions of the 5 demographic diversity variables all fell within acceptable tolerances of skew <1.0 and kurtosis <.3.0, enabling the use of parametric methods. Three of the four dependent variables also met the criteria. Only cognitive cohesion diversity pre-discussion posed problems as it was positively skewed (S=2.56) with a Kurtosis of 10.91. Attempts to transform this (using natural log transformation, Fisher's arcsine and square root transformations) were not successful. The problem appeared to be due to two outliers, one team at each extremity of the data. By removing these two outliers, the skew was reduced to -0.15 and the Kurtosis to -1.32. Thus the analysis for this variable is on the 21 teams only.

Having established the validity of the data for analysis, the next section will report on the zero-order correlations between the study variables.

7.3.2 Zero-order Correlations (Team Level) for Hypothesis 2

In Chapter 5, H2 was stated as: team level demographic diversity will be positively associated with cognitive diversity. Initial evidence comes from the zero-order correlations between the 5 demographic and the 5 cognitive diversity variables as presented in Table 7.4. Relationships with the pre-discussion measures of cognitive diversity will be presented first, followed by cognitive change diversity, and the post-discussion measures.

In line with 'upper echelons' theory, a positive relationship between demographic diversity and cognitive diversity pre-discussion is expected. However, the findings relating to pre-discussion cognitive diversity provide very limited support for the proposition in that only tenure diversity is statistically significantly associated in the expected (positive) direction (r = .31, $p < 0.10$). This means that teams characterised by variation in tenure amongst members experience greater cognitive variation pre-discussion.

ble 7.4 Zero-order Correlation Coefficients for Demographic Diversity and Cognitive Diversity

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Age Diversity	-										
2. Functional Diversity	.11	-									
3. Educational Diversity	.15	.10	-								
4. Gender Diversity	.33†	.29†	.03	-							
5. Tenure Diversity	.11	-.10	.04	-.31†	-						
6. Cognitive Diversity pre-discussion	-.03	.15	.20	-.02	.31†	-					
7. Cognitive Cohesion Diversity pre-.	-.06	-.18	-.25	-.15	.26	.01	-				
8. Cognitive Change Diversity	-.16	-.22	-.44*	-.24	-.06	-.14	-.79**	-			
9. Cognitive Diversity post-discuss.	.19	.42*	.34†	.37*	.05	.39*	-.58**	.79**	-		
10. Cognitive Cohesion Diversity post-	.15	.04	.17	.07	.56**	.23	.16	-.04	.57**	-	
11. Team Size	.36*	.39*	.34†	.16	.44**	.55**	.65**	-.32†	.01	.52**	-

N = 21-23; † p <0.10; * p <0.05; ** p <0.01 (1 tailed).
Shaded area represents relationships of interest for H2.

As regards cognitive cohesion diversity pre-discussion (which measures the variation in the proximity of individuals' rank order to the team consensus ranking), none of the relationships was significant, thus failing to support the hypothesis.

In summary, there is little support for the notion that demographic diversity predicts pre-discussion cognitive diversity. This is a particularly important non-finding. Much of the TMT demographic research literature is predicated on the group composition literature (reviewed in Chapter 3), which asserts that a diverse group of people will have a broad, cognitively diverse base upon which to build a collective decision. The findings here suggest that only tenure diversity may be important, but none of the other types of diversity is even close to supporting such a notion.

With respect to team level cognitive change diversity (that is, the median coefficient of variation of cognitive change in the team) it was expected that demographic diversity would exhibit a negative relationship. As explained in Chapter 6, it would be unlikely for a diverse group of people to alter the spread of its rankings dramatically. The association with functional background diversity is substantial and negative as predicted ($r = -.44$, $p < 0.05$). This means that teams that are characterised by greater diversity in functional background show the least change of rankings following the team discussion

With respect to cognitive diversity post-discussion, a positive relationship was expected with demographic diversity. That is to say, greater demographic diversity should lead to greater ranking diversity. Preliminary findings offer some support for the hypothesis. Zero-order correlations with three aspects of demographic diversity are substantial, statistically significant and positive. These are: functional diversity ($r = .42$, $p<0.05$); educational diversity ($r = .34$, $p<0.10$); and gender diversity ($r = .37$, $p<0.05$). In other words, top management teams characterized by having a higher proportion of women, more diverse educational backgrounds and a greater variety of functional backgrounds, show greater cognitive diversity (i.e. less agreement) in terms of private rank-ordering of candidates after discussion.

With respect to cognitive cohesion diversity post-discussion, the expected positive relationship was observed with tenure diversity ($r = .56$, $p<.01$). This means that there is more diversity of ranking around the team consensus after a discussion in top management teams characterised by high levels of tenure diversity.

In summary, the initial findings for *post-discussion* cognitive diversity seem to provide some support for H2. Specifically, differences in gender, education and functional backgrounds and tenure are important for cognitive diversity post-discussion.

In order to control for the potentially confounding factor of team size, regression analyses were conducted. Although it is common practice to use regression analyses to explain the variance in the dependent variables using more than one predictor, and to ascertain the unique effects of predictors, it was noted earlier in this chapter, that in order not to violate cases-to-dependent-variables ratios, it is more appropriate to regress the types of demographic diversity separately onto the cognitive diversity variables. These are reported in the next section.

7.3.3 Multiple Regression Analyses (Team Level) for Hypothesis 2

The zero-order correlations, as shown in Table 7.4 and reported in the previous section, show only one statistically significant relationship for the cognitive diversity pre-discussion, that with tenure diversity which is related to team size. In the regression analysis reported in Table 7.5, team size is entered in the first step, followed by tenure diversity in the second. The effect of tenure diversity is no longer statistically significant.

With consideration to cognitive change diversity, functional background diversity is not a function of team size and singly explains 13% of the variance, ($\beta = -0.37$, $p < 0.05$). To illustrate by way of example, this finding means that a functionally diverse team, which say, comprises a couple of engineers, an accountant, a scientist and a HR consultant, will experience less cognitive shift in rankings than functionally homogeneous teams, comprising, say all engineers. That is to say, the individuals in a functionally diverse team will tend to maintain their original rankings, whereas in a functionally homogeneous team, people will change their minds more.

The regression analyses for cognitive diversity post-discussion shows that once team size is controlled, functional diversity is the only significant predictor ($\beta = 0.33$, $p < 0.05$), predicting 9% of the variance. This means that there is greater variety of opinions after discussion in a functionally diverse team. Educational diversity was no longer significant after controlling for team size. As noted in the previous section, gender diversity is also a predictor, but as neither gender diversity nor cognitive diversity post-discussion are related to team size no regression was applied, although this result means that gender diversity explains 14% of the variance in cognitive diversity post-discussion. This finding means that mixed gender teams have more diverse opinions post-discussion.

Tenure diversity had a significant association with cognitive cohesion diversity post-discussion (β = 0.41, p<0.05) over and above team size, explaining 13% of the variance. This finding means that after a discussion, tenure diverse teams exhibit high levels of variation around the team consensus. That is to say, tenure diversity predicts cognitive diversity about the team consensus.

In summary, the conclusion based on findings reported in this section, is that demographic diversity does not predict cognitive diversity pre-discussion. However, the hypothesis that demographic diversity predicts cognitive diversity did receive some support in that functional diversity predicts cognitive diversity post-discussion, and is a negative predictor of cognitive change diversity (i.e. no variation), whilst tenure diversity is a positive predictor of cognitive cohesion diversity post-discussion.

The next section will present the interpretation regarding the findings and non-findings for H2.

Table 7.5 Multiple Regression Analyses for the Effects of Demographic Diversity on Cognitive Diversity

Outcome Variable	Predictor Variable	β	$R^2 \Delta$	F	df
Cognitive diversity *pre-*.	1. Team Size	.52*	.31	9.25**	1,21
	2. Tenure Diversity	.08	.01	4.50*	2,20
Cognitive change diversity	1. Team Size	-.19	.10	2.36	1,21
	2. Functional Diversity	-.37*	.13	2.92*	2,20
Cognitive diversity *post-*	1. Team Size	.51**	.43	15.57**	1,21
	2. Functional Diversity	.33*	.09	8.41**	2,20
	1. Team Size	.58**	.43	15.57**	1,21
	2. Educational Diversity	.19	.03	8.41**	2,20
Cognitive cohesion div. *post-*.	1. Team Size	.34†	.27	7.7**	1,21
	2. Tenure Diversity	.41*	.13	6.8**	2,20

N = 23; † = p <0.10; * = p <0.05; ** = p <0.01. One-tailed tests of significance levels for standardised β coefficients in final equation. $R^2 \Delta$ refers to discrete steps. No entry for gender diversity and cognitive diversity post-discussion as neither is related to team size.

7.3.4 Conclusions Regarding Hypothesis 2

In summary, 'upper echelons' theory in its typical form, that is, that cognitive diversity at the team level is a function of team level demographic diversity, found little support in this study.

In the case of H2, significant zero-order correlations were found between demographic diversity indices and the cognitive diversity measures in six of the 25 instances examined (five demographic diversity predictors x five cognitive diversity outcomes). Five of these were subjected to regression analyses, of which three remain statistically significant and another was not subjected to regression because there was only one predictor. The loss of statistical significance after regression may be due to a reduction in power owing to the extra degrees of freedom used, or it could be that the results were marginal in the first instance. In either case, the loss of statistical significance suggests that the results are not robust. The results for H2 are slightly more than one would expect to occur through chance alone.

Specifically with regard to cognitive diversity post-discussion, functional diversity and gender diversity are the two positive predictors, explaining 9% and 14% of the variance respectively. With regard to cognitive change diversity, functional diversity is the single negative predictor, explaining 13% of the variance. Tenure diversity is the only (positive) predictor of cognitive cohesion diversity post-discussion explaining 13% of the variance.

Such paucity of results is not uncommon in TMT demographic research. Typically, the blame is retrospectively attributed to selection biases in appointments to the TMT even where the internal homogeneity of the data has not been established (see Boone et al., 2004). In this case, the scapegoat cannot be that there was not enough demographic difference in the dataset, as sufficient variation was established using ICCs as reported earlier in this chapter.

The meagre clutch of results for H2 may simply be a problem of power. The data collection requirements to gain a sizeable sample of real top management teams are high, but an n of 23 may not be a large enough base to demonstrate the expected large effects.

On the other hand, this was a particularly rigorous study of cognitive variation. Care was taken to create meaningful indices of cognitive diversity which corresponded to measures of demographic diversity. Unlike other studies in this genre, individuals' preferences (rankings) were measured prior to a discussion. Although the link between demographic diversity and cognitive diversity prior to a team discussion is widely assumed, it has never been measured with intact TMTs before. Similarly, the link between demographic diversity and cognitive diversity post-discussion, which found little support in this study, also has not been measured before.

The next chapter will investigate the proposition across two levels.

CHAPTER 8

Findings: Hypothesis 3

8.0 Overview

This chapter presents the findings concerning hypothesis 3, which transcends the two levels of analysis already discussed in Chapter 7. Owing to the complexity of the modelling and the requisite lengthy explanation, the analysis and results for this hypothesis warrant a single chapter.

8.1 Introduction

This chapter investigates a research question arising from the previous two hypotheses, that is, do team level demographic factors influence the way individuals think? This is a concept not previously investigated in the research literature, although it is a logical extension of 'upper echelons' theory. There are two strands to the question that can be answered discretely by multi-level modelling. The first is does demographic diversity at the team level contribute to individuals' opinions being different? This is a simple question answered by a simple multi-level model. The second strand is an extension of the first; is there an interaction between two predictors, such that the effect of individual demographic dissimilarity on cognitive dissimilarity differs depending upon team diversity? In other words, is it the demographically dissimilar individual in a demographically diverse team that has the most cognitively dissimilar opinions? This is a more complex question which can be answered by a correspondingly more complex multi-level model.

8.2 Team Level Demographic Diversity & Individual Cognitive Variation (H3)

Hypothesis 3 (H3) is that team demographic diversity will relate to individual cognitive dissimilarity. That is to say, the prediction is that top managers in demographically diverse teams will show greater difference in their rank order of candidates from their team colleagues than will managers in homogeneous teams. The logical extension of this hypothesis is to ascertain whether the diversity of the team moderates the effect of individual demographic dissimilarity on cognitive dissimilarity.

8.2.1 Overview of Multi-Level Modelling Procedure

This section is designed to give an overview of the multi-level modelling procedure used in subsequent pages of this chapter and in Chapter 9. Following Hox (1995), the approach taken here follows a 5 step sequence to build the complex multi-level model. As will be explained later, each of the steps represents a discrete model in its own right, with step 3 being the simple model referred to above that answers the fundamental nub of the hypothesis. As step 3 is also a part of the sequence of building the complex model (concerning dissimilar individuals in diverse teams), it is embedded within the whole model presented in the tables below. Each step will be explained in turn, illustrating how each relates to testing the hypothesis.

8.2.1.1 Step 1

The first step is to construct a variance components model, so called because the variance in the individual level dependent variable is partitioned into components corresponding to levels in the hierarchy being studied. In this study, level 1 comprises the individual cases (n=130) nested within level 2, which covers the teams (n = 23). In this first step there are no explanatory variables, only the intercept of the independent variable is modelled.

The variance partition coefficient (Goldstein, 1995), or intraclass correlation (Hox, 1995), measures the extent to which the values on the dependent variable of individuals in the same team resemble each other as compared to those from individuals in different teams. The former is referred to as between teams or level 2 variance, and the latter as between individuals or level 1 variance.

The variance components model also provides a value of the deviance (a statistic known as the -2*loglikelihood), which is a degree of misfit of the model (McCullagh & Nelder, 1989; Hox 1995). In subsequent steps in the modelling process, it is the reduction in the deviance established in the variance components model that, when statistically significant, is used as a criterion of model fit.

With respect to the current hypothesis, a variance components model will be established (below) for each of the cognitive dissimilarity variables. Only one individual level response (dependent) variable can be modelled at a time.

8.2.1.2 Step 2

Lower level explanatory variables are added to the model in the second step. With respect to the current hypothesis, step 2 requires that the five demographic dissimilarity (individual level) variables are added in turn to the variance components model established in step 1 for the cognitive dissimilarity variables. The contribution of the individual explanatory variables can be assessed by referring to the change in the deviance. The difference is compared to a chi-squared distribution with 1 degree of freedom. This will be statistically significant if the independent predictor significantly affects the response variable. It is also expected that changes will result in a reduction of the variance in the random effects part of the model, as the addition of the explanatory variables should result in a decrease in the level 1 variance. Because team composition is not identical for all teams, the individual dissimilarity explanatory variables should explain part of the individual and part of the team level variance in the cognitive dissimilarity variables (see Hox, 1995). This means that there also should be a decrease in level 2 variance, indicating that these explanatory variables explain some of the variance between teams.

Initially, the corresponding variance components of the slopes are fixed at 0. This is sometimes referred to as simple variation at level 1, which means that the only variation between individuals is in their intercepts (Rasbash et al., 2004). The next stage is to test the random slope variation between the teams. This means that the 23 team lines are allowed to have different slopes, that is, the coefficients of the explanatory variables can vary from team to team. Both the intercept and the slope are allowed to vary randomly across teams. The intercept coefficient is modified by the addition of the explanatory variable slopes. The reduction in the deviance is tested for significance on an additional degree of freedom to allow for the covariance between the parameters.

The purpose of testing both the fixed and varying by team intercepts is to assess which achieves the best model fit as indicated by the reduction in deviance. Indeed, a coefficient which is not significant when kept fixed may achieve significance when allowed to vary by team (Goldstein, 1995). The practice is to take forward the best one to the next step in the analysis.

In modest datasets such as the one in this study, it is not uncommon for the fixed and varying slopes to result in the same coefficient (Goldstein, 1995). All intercepts in this study were tested both being kept fixed and being allowed to vary by team, and gave rise to exactly the same pattern of results. To avoid redundancy, only the fixed coefficients are reported in the tables below.

In interpreting the tables which follow, one is looking for a statistically significant dissimilarity slope coefficient accompanied by a statistically significant reduction in the deviance. Together, these two test the effect of the predictor on the response variable. The slope coefficient indicates how the predictor affects the intercept. A positive slope coefficient means a positive relationship between cognitive and demographic dissimilarity and a negative slope coefficient means a negative relationship between cognitive and demographic dissimilarity.

In some cases, as will be explained in the appropriate sections, it is possible to have a significant reduction in the deviance which appears to have a 0.00 slope coefficient. If this occurs in steps 2, 3 and 4 of the model, it is usual to keep testing the predictor by carrying it forward to the next step (Hox, 1995). It is also possible for predictors significant at $p<0.10$ during these steps to increase in significance when married up with other predictors in subsequent steps. In such cases the practice is to retain such significant variations for further analysis (Hox, 1995).

One could assume that the results for step 2 of the multi-level model should mirror those arising from H1, which also tested relationships between individual demographic dissimilarity and cognitive dissimilarity. However, for H1 the sample was 130 executives and relationships observed were for the whole sample irrespective of team considerations (for example, there was no control for teamsize). Hence, it is possible that team level factors are confounded within the results for H1. By way of contrast, the strength of the multi-level model is that it is testing for the relationships within the context of each team. Thus, the multi-level model is much more stringent than the tests applied for H1 and results may differ for H3, which appears, ostensibly, to test the same relationships.

8.2.1.3 Step 3

Higher order variables, that is, the team level demographic diversity predictors, are tested in step 3. They are treated in exactly the same way as the individual level predictors. The purpose of step 3 is twofold. As a stand alone model, it tests H3 - whether diversity at the team level predicts individual cognitive dissimilarity. In the sequence of the 5 steps which build the complex model, it serves to determine which, if any, of the team level (diversity) variables need to be considered as possible predictors of individual level outcomes, before testing more stringently whether they add to the variance explained by the individual level variables and examining interactions.

In order to test the hypothesis, one is looking for a reduction in the deviance accompanied by a statistically significant slope coefficient (this is the principle established in step 2 and is the same for all steps). The team level predictors can only explain the team level variance component. In other words, the team level demographic diversity predictors, explain differences between teams, but not differences between individuals.

In order to test the first part of the hypothesis that demographic diversity (team level) will predict individual cognitive dissimilarity (individual), each of the team level diversity variables is kept fixed in exactly the same way. As with step 2, one is looking for a statistically significant slope coefficient accompanied by a statistically significant reduction in the deviance in order to interpret how the predictors affect the response variable.

8.2.1.4 Step 4

The previous steps have regressed each predictor separately on to the response variable, as if each were in fact, a separate model. The next step is to establish the main effects of testing the predictors at the same time. In this step, the relevant higher level (team diversity) explanatory variables are entered to the model together with the lower level (dissimilarity) variables, to see if they explain additional variance. It is those predictors that were observed to be statistically significant or approaching statistical significance that are taken forward to step 4 (Hox, 1995). It can be argued of course, that seemingly insignificant predictors when combined with other seemingly insignificant predictors will result in a statistically significant main effect and interaction. In the present study, this would mean testing all 25 (five dissimilarity x five demographic diversity) predictors in combination. This was done, but effects were only noted amongst predictors that had already been statistically significant or marginally significant at step 2 and 3. Thus this chapter follows Hox (1995) in only taking forward those predictors that achieved a better fit at steps 2 and 3.

In order to test the hypothesis that diversity is important over and above dissimilarity in terms of how individuals' rankings differ, the tables below must show a statistically significant coefficient for both slopes (dissimilarity and diversity) accompanied by a statistically significant reduction in the deviance at step 4 relative to step 2. Changes in the variance components indicate which part of the variance (i.e. between teams or between individuals) that is being explained.

8.2.1.5 Step 5

In the final step, cross-level interaction terms (dissimilarity * diversity) are computed for the variables used to test the main effects in step 4. Here, the reduction in deviance is compared against that derived for the two terms in the main effects model at step 4, in order to test the hypothesis. For an interaction to be significant, the interaction slope must be statistically significant and be accompanied by a statistically significant reduction in the deviance.

It is also appropriate to mention here that the predictor variables are centred (i.e. the mean is subtracted from each raw value) before beginning the multi-level modelling process (Goldstein, 1995). This is done in order to provide a mean of 0 and to avoid any convergence problems resulting from multi-collinearity in the modelling process.

The following sections of this chapter will apply the five step procedure outlined above to the cognitive dissimilarity variables.

8.2.2 Null Model: Cognitive Dissimilarity Pre-Discussion

The first cognitive individual response variable to be investigated using the multi-level model is cognitive dissimilarity pre-discussion. The multi-level analysis of this variable is explained in detail to make the procedure and its application explicit. This means that the four subsequent variables, which are treated in exactly the same way, can be described more succinctly.

The first step is to establish the variance components or null model. By referring to the variance components section of Table 8.1, it can be seen that the overall mean of cognitive dissimilarity pre-discussion is estimated as β_0 = 4.59. The means for the different teams are distributed around the overall mean with an estimated variance of σ^2_u 2.93. The variance between individuals within teams is σ^2_e 3.39. From a normal test of H_0: $\sigma^2_{u0} = 0$ ($Z = 2.93/1.05 = 2.79$, $p<0.05$), it can be seen that the variance between teams is significantly different from 0. The -*2Log-Likelihood or deviance statistic is 567.82. The next step will be compared for significant reduction against this figure.

In order to establish the percentage of variance attributable to between team differences, the variance partition coefficient (VPC) is established as follows:

$$\frac{2.93}{2.93 + 3.39} = 0.46$$

This means that approximately 46% of the total variance in cognitive dissimilarity pre-discussion may be attributed to differences between teams. That is to say, there is variation in the individual response variable that is directly attributable to team level factors. The reduction in the level 2 variance will be referred to repeatedly in the analysis which follows as it is the reduction in this variance across the function of the response variable that shows the relative contribution of each of the explanatory variables in testing H3.

8.2.2.1 Dissimilarity Effects: Cognitive Dissimilarity Pre-Discussion

In step 2, the individual demographic dissimilarity predictors are added, and are examined first by holding constant the individual variation in the response variable and then allowing it to vary by team. The five explanatory variables are investigated

Table 8.1 Cognitive Dissimilarity Pre-discussion and Demographic Variation Steps 1 - 5

	Intercept Coefficient		Slope #1 Dissimilarity Coefficient		Slope #2 Diversity Coefficient		Slope #3 Interaction Coefficient		Level 2 Variance (Between Teams)		Level 1 Variance (Between Individuals)		Deviance	Change in Deviance from Previous Step
Model Steps & Predictors	β	(SEβ)	β	(SEβ)		(SEβ)	β	(SEβ)	σ^2_u	SE	σ^2_e	SE		
Step 1: Variance Components:														
(null model)	4.59	(0.40)	n/a	n/a	n/a	n/a	n/a	n/a	2.93	(1.05)	3.39	(0.47)	567.82	n/a
Step 2: Individual Predictors:														
Age Dissimilarity	4.63	(0.40)	0.02	(0.02)	n/a	n/a	n/a	n/a	2.88	(1.04)	3.36	(0.46)	566.47	1.37
Functional Dissimilarity	4.66	(0.38)	0.24	(0.15)	n/a	n/a	n/a	n/a	2.61	(0.96)	3.38	(0.46)	565.37	2.45
Educational Dissimilarity	4.68	(0.39)	0.09	(0.07)	n/a	n/a	n/a	n/a	2.72	(0.99)	3.38	(0.46)	566.14	1.35
Gender Dissimilarity	4.59	(0.40)	-0.04	(1.01)	n/a	n/a	n/a	n/a	2.94	(1.06)	3.39	(0.46)	567.82	0.00
Tenure Dissimilarity	4.59	(0.40)	0.00	(0.00)	n/a	n/a	n/a	n/a	2.95	(1.06)	3.38	(0.46)	567.81	0.01
Step 3: Team Level Predictors:														
Age Diversity	4.59	(0.40)	n/a	n/a	-0.30	(4.14)	n/a	n/a	2.94	(1.06)	3.39	(0.46)	567.82	0.00
Functional Diversity	4.61	(0.39)	n/a	n/a	0.84	(1.87)	n/a	n/a	2.91	(1.05)	3.39	(0.47)	567.82	0.00
Educational Diversity	4.61	(0.39)	n/a	n/a	2.46	(3.74)	n/a	n/a	2.86	(1.04)	3.39	(0.46)	567.40	0.42
Gender Diversity	4.59	(0.40)	n/a	n/a	-0.16	(2.31)	n/a	n/a	2.94	(1.06)	3.39	(0.47)	567.82	0.00
Tenure Diversity	4.62	(0.38)	n/a	n/a	3.39	(2.63)	n/a	n/a	2.18	(0.99)	3.39	(0.46)	566.21	1.71
Step 4: Main Effects (Individual and Team Predictors Together):														
Not Applicable	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Step 5: Cross-level Interactions:														
Not applicable	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

* significant at p <0.05; and ** significant at p <0.01 Steps 2 – 5: d/f 2,117 to 2,127 depending on missing values

separately and are in effect, discrete models. In this step the intercept coefficient ranges from 4.59 (gender and tenure dissimilarity) to 4.68 (educational dissimilarity). The change in the deviance ranges from 0.00 (gender dissimilarity) to 2.45 (educational dissimilarity) on 2 degrees of freedom. None of the reductions in the deviance is significant.

These results at the end of step 2 mean that, contrary to expectation, demographically dissimilar individuals in TMTs do not exhibit more cognitive dissimilarity than their peers. This is consistent with the findings concerning H1 which are that demographic dissimilarity does not predict cognitive dissimilarity pre-discussion (see p 144).

8.2.2.2 Diversity Effects: Cognitive Dissimilarity Pre-Discussion

In step 3, the team diversity effects are examined. The hypothesis, which posits that demographic diversity will predict cognitive dissimilarity pre-discussion, receives no support. This is evidenced in the section of Table 8.1 which reports the coefficients for the team level diversity variables, where no reduction in deviance is noted for any of the predictors. This means that demographic diversity, contrary to the hypothesis, is not responsible for influencing individuals' private judgements prior to a team discussion. In terms of the second part of the hypothesis being considered, diversity is unable to explain any additional variance to that explained by individual dissimilarity as neither is statistically significant.

The next steps (4 and 5) of the model, main effects and interactions, are established by taking forward the significant results from steps 2 and 3. In this case, there are no statistically significant results to take forward, negating this part of the model. It could be argued of course, that a relationship between two seemingly insignificant variables may emerge when both are tested together. As explained earlier, all 25 relationships were tested as both main effects and interactions, but none was significant and hence are not reported.

8.2.2.3 Summary: Cognitive Dissimilarity Pre-Discussion

In summary, the analysis using multi-level modelling presented in this section, rejects both of the arguments posited in the hypothesis. It was found at step 3 that diverse teams are not characterised by more individual cognitive dissimilarity pre-discussion. As for the second argument, it was discovered at step 2, that there is no statistically significant difference between the way dissimilar individuals within teams rank the candidates pre-discussion and as opposed to more similar individuals within teams.

The next section applies the multi-level modelling process to cognitive cohesion dissimilarity pre-discussion.

8.2.3 Null Model: Cognitive Cohesion Dissimilarity Pre-Discussion

The second cognitive individual response variable to be investigated using the multi-level model is cognitive cohesion dissimilarity pre-discussion. As detailed in Chapter 6 (pp 124 – 125), cognitive cohesion dissimilarity measures the proximity of the individuals' pre-discussion ranking to their teams' subsequent consensus ranking on a scale of -1.0 to 1.0, where 1.0 corresponds to a complete match with the team. As explained on page 114, it is expected that the association between the predictors and this response variable will be negative. That is to say, dissimilar persons will tend towards the least conformist (-1.0) rankings.

By referring to the variance components section of Table 8.2, it can be seen that the overall mean of cognitive cohesion dissimilarity pre-discussion is estimated as β_0 = 0.36. The means for the different teams are distributed around the overall mean with an estimated variance of σ^2_u 0.01, whilst the between individuals variance is σ^2_e 0.32. This means that team level differences explain 3% of the variance. The deviance is 225.64. Although the variance is small, it is still worth seeking the team level factors which explain it.

8.2.3.1 Dissimilarity Effects: Cognitive Cohesion Dissimilarity Pre-Discussion

In step 2, the five individual demographic dissimilarity predictors are added in turn. One is statistically significant and one is approaching significance. Tenure dissimilarity has a negative slope coefficient of (β -0.00, p <0.05), the reduction in the deviance is statistically significant (Δ2*Log-Li = 3.51, p <0.05). The level 2 variance is also reduced, meaning that tenure dissimilarity does explain variation in cognitive cohesion dissimilarity at the team level. The β = -0.00 slope coefficient for tenure dissimilarity (albeit in the negative, expected direction) is somewhat unusual, but should not be interpreted to mean that there is zero change in the slope. It should be pointed out that as the multi-level model computes the regression coefficient to several decimal places, a very tiny effect can be obscured by the convention of reporting only to two decimal places. (If one were to report this particular result to three decimal places for example, the slope coefficient would be β = -0.001). This result is in line with prediction, as one would expect dissimilar individuals to choose a different ranking to that of their team consensus ranking.

These results at the end of step 2 mean that tenure dissimilar individuals in teams exhibit the least proximity to the team consensus pre-discussion, that is, they are the least likely to make the same preferential ranking as the team consensus. This is an interesting result when one compares it with the statistically significant finding for H1, which was, that age dissimilar individuals (not tenure dissimilar ones) that have the most cognitive cohesion dissimilarity pre-discussion. This apparent discrepancy draws attention to the difference between the two analytical methods. The original correlation analysis (no regression was conducted as age dissimilarity was the single predictor in H1) ignores team effects. Consequently, team effects are confounded with individual level effects. The zero-order correlations assume a single best-fit linear association through all data-points, including any outliers. In contrast, a multi-level model, assumes multiple parallel lines representing the best-fit through the data-points for each team (23 lines through the 130 data-points). Hence, an outlier will have a much reduced effect when team effects are controlled for, as is the case in multi-level modelling.

The present analysis means that the earlier finding that age dissimilarity predicted cognitive cohesion dissimilarity is not contested; but at the same time it reveals that, once team effects are taken into consideration, tenure dissimilarity is the only predictor of cognitive cohesion dissimilarity. The analysis for H1 refers to top managers more generally, whereas the multi-level model refers to managers within the context of their particular TMT.

8.2.3.2 Diversity Effects: Cognitive Cohesion Dissimilarity Pre-Discussion

When the demographic diversity predictors are tested in step 3, only one achieves significance, that is, tenure diversity (Δ*2L-Li = 2.77, p <0.05) accompanied by a reduction in the team level variance (σ^2_u 0.00). The negative slope coefficient (β -

0.65, p <0.05) indicates a negative relationship as expected. This result means that individuals in tenure diverse teams experience less proximity to the team consensus pre-discussion than those in teams whose members have similar tenure.

As tenure diversity is the only form of demographic diversity that is responsible for influencing an individual's proximity to the team consensus pre-discussion, the hypothesis receives little support.

The next step is to test the effects of entering tenure dissimilarity and tenure diversity, together.

8.2.3.3 Additive Effects of Diversity: Cognitive Cohesion Dissimilarity Pre-Discussion

Table 8.2 reports the main effects of modelling tenure diversity and dissimilarity together on cognitive cohesion dissimilarity pre-discussion. They are not statistically significant, which means that tenure diversity does not account for a statistically significant amount of variance once tenure dissimilarity has been controlled.

8.2.3.4 Interaction Effects: Cognitive Cohesion Dissimilarity Pre-Discussion

The lower portion of Table 8.2 reports the intercept, slope and interaction coefficients for the possible interaction term. It is not statistically significant.

8.2.3.5 Summary: Cognitive Cohesion Dissimilarity Pre-Discussion

In summary, the results reported in this section provide a modicum of support for one of the strands of the hypothesis whilst rejecting the other. The key finding to emerge is that tenure dissimilarity and tenure diversity, separately predict cognitive cohesion dissimilarity. That is to say, tenure dissimilar individuals, and those persons in a tenure diverse team, are more likely to select a ranking that does not agree with the team consensus. Even so, the finding is very weak. Moreover, because tenure diversity does not explain any additional, statistically significant variance, demographic diversity cannot be held to influence dissimilar individuals in relation to their proximity to the team consensus pre-discussion.

The next section applies the multi-level modelling procedure to cognitive change dissimilarity.

Table 8.2 Cognitive Cohesion Dissimilarity Pre-discussion and Demographic Variation Steps 1 - 5

	Intercept Coefficient		Slope #1 Dissimilarity Coefficient		Slope #2 Diversity Coefficient		Slope #3 Interaction Coefficient		Level 2 Variance (Between Teams)		Level 1 Variance (Between Individuals)		Deviance	Change in Deviance from Previous Step
Model Steps & Predictors	β	(SEβ)	β	(SEβ)		(SEβ)	β	(SEβ)	σ^2_u	(SE)	σ^2_e	(SE)		
Step 1: Variance Components:														
(null model)	0.36	(0.05)	n/a	n/a	n/a	n/a	n/a	n/a	0.01	(0.00)	0.32	(0.04)	225.64	-
Step 2: Individual Predictors:														
Age Dissimilarity	0.35	(0.05)	-0.01	(0.00)	n/a	n/a	n/a	n/a	0.01	(0.00)	0.32	(0.04)	223.69	1.95
Functional Dissimilarity	0.36	(0.05)	-0.02	(0.03)	n/a	n/a	n/a	n/a	0.01	(0.02)	0.32	(0.04)	222.34	0.30
Educational Dissimilarity	0.35	(0.05)	-0.01	(0.02)	n/a	n/a	n/a	n/a	0.01	(0.02)	0.32	(0.04)	225.53	0.11
Gender Dissimilarity	0.36	(0.05)	0.34	(0.24)	n/a	n/a	n/a	n/a	0.01	(0.00)	0.32	(0.04)	223.70	1.94
Tenure Dissimilarity	0.36	(0.05)	-0.00*	(0.00)	n/a	n/a	n/a	n/a	0.00	(0.00)	0.32	(0.04)	222.13	3.51*
Step 3: Team Level Predictors:														
Age Diversity	0.36	(0.05)	n/a	n/a	-0.63	(0.58)	n/a	n/a	0.01	(0.02)	0.32	(0.04)	225.45	0.19
Functional Diversity	0.36	(0.05)	n/a	n/a	-0.24	(0.28)	n/a	n/a	0.01	(0.00)	0.32	(0.04)	224.94	0.70
Educational Diversity	0.36	(0.05)	n/a	n/a	0.57	(0.53)	n/a	n/a	0.01	(0.00)	0.32	(0.04)	224.51	0.14
Gender Diversity	0.35	(0.08)	n/a	n/a	-0.00	(0.32)	n/a	n/a	0.01	(0.00)	0.32	(0.04)	224.64	0.00
Tenure Diversity	0.36	(0.05)	n/a	n/a	-0.65*	(0.38)	n/a	n/a	0.01	(0.00)	0.32	(0.04)	222.87	2.77*
Step 4: Main Effects:														
Tenure Dissimilarity + Tenure Diversity	0.36	(0.05)	-0.00*	(0.00)	-0.45	(0.40)	n/a	n/a	0.00	(0.00)	0.32	(0.04)	220.86	1.27
Step 5: Interactions:														
Tenure Dissimilarity * Tenure Diversity	0.37	(0.05)	-0.00	(0.00)	-0.40	(0.40)	0.00	(0.00)	0.00	(0.02)	0.32	(0.04)	220.86	0.00

* significant at p <0.05; and ** significant at p <0.01, one-tailed. Steps 2: d/f 2,117 to 2,127; Step 3: d/f, 1,118 to 1,128; Steps 4 & 5: d/f 2,117 to 2,127 and 3,116 to 3,126 depending on missing values. Δ*2L_li for 2 way interactions at Step 5 are compared against main effects at Step 4 as a proxy for full formula of 2 main effects and 2 two-way interactions.

8.2.4 Null Model: Cognitive Change Dissimilarity

The third cognitive individual response variable to be investigated using the multi-level model is cognitive change dissimilarity. As detailed in Chapter 6 (p 126), cognitive change dissimilarity measures the proximity of the individuals' pre-discussion ranking to their post-discussion ranking on a scale of -1.0 to 1.0, where 1.0 = no change (complete agreement). As explained on page 115, it is expected that the association between the predictors and the response variable will be positive. That is to say, dissimilar persons will tend to hold to their opinion (1.0).

By referring to the variance components section of Table 8.3, it can be seen that the overall mean of cognitive change dissimilarity is estimated as $\beta_0 = 0.75$. The means for the different teams are distributed around the overall mean with an estimated variance of σ^2_u 0.02, whilst the between individuals variance is σ^2_e 0.18. This means that team level differences explain 10% of the variance. The deviance is 152.99. Although the variance at level 2 is relatively small, it is statistically significant at $p<0.05$ and is worth exploring to see which team level factors explain it.

8.2.4.1 Dissimilarity Effects: Cognitive Change Dissimilarity

In step 2, the five individual demographic dissimilarity predictors are added in turn. Only gender dissimilarity achieves a significant reduction in deviance of 4.98 (p <0.05). The variance between teams is reduced as is that between individuals. The positive slope coefficient (β 0.42, p <0.05) suggests a positive relationship between gender dissimilarity and no change in ranking, consistent with the prediction. This is a particularly striking effect given that there were so few women in the sample (12 women in 9 of the 23 teams). Moreover, as explained on pages 104 - 105, no other study of TMTs has included gender as a study variable.

8.2.4.2 Diversity Effects: Cognitive Change Dissimilarity

Next, in step 3 the five team diversity predictors are added in turn. Three of them are statistically significant, and in order of significance these are: functional background diversity (Δ*2L-Li = 6.63, p <0.05); gender diversity (Δ*2L-Li = 6.23, p <0.05); and educational diversity (Δ*2L-Li = 4.47, p <0.05). All have statistically significant positive slope coefficients indicative of relationships as predicted, whilst team level variance is reduced to σ^2_u 0.01.

These three results support the first part of the hypothesis that team level demographic factors influence individual cognitive variation. Specifically, these findings are that individuals' privately held opinions tend to stay the same (i.e. change least) in gender diverse teams, in functionally diverse teams and in educationally diverse teams. In other words, there is very little cognitive shift between pre- and post-discussion rankings in functionally diverse teams, or in teams in whose members have been educated to different levels. Similarly, in mixed gender teams, neither men nor women exhibit a great deal of cognitive shift between their private pre- and post-discussion rankings.

8.2.4.3 Additive Effects of Diversity: Cognitive Change Dissimilarity

Table 8.3 reports the modelling (step 4) of gender dissimilarity and diversity together on cognitive change dissimilarity as main effects. This shows whether or not the diversity variable explains variance over and above the dissimilarity one.

When entered into the model together, gender dissimilarity and functional diversity reduce the deviance by 4.41 ($p<0.05$), achieving corresponding positive slope coefficients and reduction in level 2 and level 1 variance. This means that functional diversity explains variance in cognitive change over and above that explained by gender dissimilarity. It is women, and people in functionally diverse teams, who change their pre-discussion opinions least. Following Hox (1995), this result will be taken forward to Step 5. The second result, regarding gender dissimilarity and educational diversity is similar ($\Delta 2$*Log-Li = 3.54 $p<0.05$), meaning that it is women and people in educationally diverse teams that do not shift from their pre-discussion rankings. The final test for gender dissimilar individuals and gender diverse teams is not significant. In other words, the gender diversity effect observed in Step 3 is part and parcel of the gender dissimilarity one (step 2).

The next step will ascertain whether it is gender dissimilar individuals (i.e. particularly women as they took highest values), embedded in these diverse teams that are not changing their minds.

Table 8.3 Cognitive Change Dissimilarity and Demographic Variation Steps 1 - 5

Model Steps & Predictors	Intercept Coefficient β	(SEβ)	Slope #1 Dissimilarity Coefficient β	(SEβ)	Slope #2 Diversity Coefficient	(SEβ)	Slope #3 Interaction Coefficient β	(SEβ)	Level 2 Variance (Between Teams) σ^2_u	(SE)	Level 1 Variance (Between Individuals) σ^2_e	(SE)	Deviance	Change in Deviance from Previous Step
Step 1: Variance Components:														
(null model)	0.75	(0.05)	n/a	n/a	n/a	n/a	n/a	n/a	0.02	(0.02)	0.18	(0.04)	152.99	n/a
Step 2: Individual Predictors:														
Age Dissimilarity	0.75	(0.79)	0.00	(0.00)	n/a	n/a	n/a	n/a	0.02	(0.00)	0.18	(0.02)	152.93	0.06
Functional Dissimilarity	0.76	(0.11)	0.04	(0.03)	n/a	n/a	n/a	n/a	0.01	(0.02)	0.18	(0.03)	151.32	1.67
Educational Dissimilarity	0.76	(0.05)	0.00	(0.01)	n/a	n/a	n/a	n/a	0.02	(0.02)	0.18	(0.03)	152.93	0.06
Gender Dissimilarity	0.76	(0.04)	0.42*	(0.17)	n/a	n/a	n/a	n/a	0.01	(0.00)	0.18	(0.02)	148.01	4.98*
Tenure Dissimilarity	0.75	(0.05)	0.00	(0.00)	n/a	n/a	n/a	n/a	0.00	(0.00)	0.18	(0.04)	152.80	0.19
Step 3: Team Level Predictors:														
Age Diversity	0.76	(0.40)	n/a	n/a	0.18	(0.49)	n/a	n/a	0.02	(0.02)	0.18	(0.02)	152.86	0.13
Functional Diversity	0.76	(0.04)	n/a	n/a	0.50*	(0.21)	n/a	n/a	0.01	(0.00)	0.18	(0.02)	147.93	6.63*
Educational Diversity	0.76	(0.04)	n/a	n/a	0.91*	(0.43)	n/a	n/a	0.01	(0.00)	0.18	(0.04)	148.52	4.47*
Gender Diversity	0.76	(0.05)	n/a	n/a	0.62*	(0.23)	n/a	n/a	0.01	(0.00)	0.18	(0.02)	146.76	6.23*
Tenure Diversity	0.77	(0.05)	n/a	n/a	-0.12	(0.31)	n/a	n/a	0.01	(0.02)	0.18	(0.02)	150.47	2.52
Step 4: Main Effects:														
Gender Diss. + Functional Diversity	0.76	(0.04)	0.37*	(0.17)	0.44*	(0.20)	n/a	n/a	0.00	(0.00)	0.18	(0.02)	143.60	4.41*
Gender Diss. + Educational Diversity	0.76	(0.04)	0.36*	(0.17)	0.75*	(0.38)	n/a	n/a	0.00	(0.00)	0.18	(0.02)	144.47	3.54*
Gender Diss. + Gender Diversity	0.76	(0.04)	0.17	(0.26)	0.45	(0.34)	n/a	n/a	0.01	(0.01)	0.18	(0.02)	146.31	1.70
Step 5: Interactions:														
Gender Diss. * Functional Diversity	0.76	(0.05)	0.40*	(0.19)	0.37	(0.26)	-0.59	(1.58)	0.00	(0.00)	0.18	(0.02)	143.46	0.14
Gender Diss. * Educational Diversity	0.76	(0.05)	0.40*	(0.18)	0.66	(0.41)	-1.38	(2.09)	0.01	(0.01)	0.18	(0.02)	144.03	0.44
Gender Diss. * Gender Diversity	0.79	(0.06)	0.35*	(0.36)	0.41	(0.35)	-1.21	(1.73)	0.01	(0.01)	0.18	(0.02)	145.82	0.49

* significant at p <0.05; and ** significant at p <0.01, one-tailed. Steps 2: d/f 2,117 to 2,127; Step 3: d/f, 1,118 to 1,128; Steps 4 & 5: d/f 2,117 to 2,127 and 3,116 to 3,126 depending on missing values. Δ*2L_li for 2 way interactions at Step 5 are compared against main effects at Step 4 as a proxy for full formula of 2 main effects and 2 two-way interactions.

8.2.4.4 Interaction Effects: Cognitive Change Dissimilarity

The lower portion of Table 8.3 reports the intercept, slope and interaction coefficients for the three possible interaction terms. By referring to the reduction in the deviance column, it is evident that none achieves a significant reduction. The positive slope coefficients are all due to gender dissimilarity, the first effect noted at step 2.

8.2.4.5 Summary: Cognitive Change Dissimilarity

In summary, the key findings to emerge from this section are that women change their minds least, and that people in teams which are functionally, educationally and gender diverse will tend to exhibit least cognitive shift in their rankings.

8.2.5 Null Model: Cognitive Dissimilarity Post-Discussion

The fourth cognitive individual response variable to be investigated using the multi-level model is cognitive dissimilarity post-discussion. As detailed in Chapter 6, this is the Euclidean distance of individuals' rankings within the team ranging from 0 – 11.20. It is expected that the association between the predictors and the response variable will be positive. That is to say, dissimilar persons in diverse teams will have a different opinion to their peers.

By referring to the variance components section of Table 8.4, it can be seen that the overall mean of cognitive change dissimilarity is estimated as $\beta_0 = 4.00$. The means for the different teams are distributed around the overall mean with an estimated variance of σ^2_u 5.55, whilst the between individuals variance is σ^2_e 3.47 (0.47). This means that team level differences explain a substantial 42% of the variance. The deviance is 582.77.

8.2.5.1 Dissimilarity Effects: Cognitive Dissimilarity Post-Discussion

In step 2, the five individual demographic dissimilarity predictors are added in turn. The only statistically significant result is for tenure dissimilarity which achieves a reduction in the deviance of 5.44, (p <0.05), the variance between teams decreases to β 5.35, and the negative slope coefficient (β -0.01, p <0.05) indicates a negative relationship between tenure dissimilarity in teams and post-discussion cognitive dissimilarity, contrary to prediction. This result means that tenure dissimilar individuals are more likely to have a similar opinion to their peers post-discussion.

Based upon author's experience at this level in organisations, there is a reasonable explanation for this finding. It is not uncommon for persons who have just joined a TMT to assume that as they are learning the ropes as it were, they may tend to think that their peers who have been around the company and the team for much longer periods are correct in their collective judgement. Although it is true to say that tenure dissimilar individuals can be those who have been in the TMT the longest, or indeed the shortest length of time, in this study the persons who took the highest values were typically those with the shortest tenure. Hence the fact that this study finds that tenure dissimilar individuals agree with their peers subsequent to a discussion is consistent with reality.

Moreover, this relationship does not mirror the findings at H1, for which it was found that amongst top managers generally, educationally dissimilar people held a more disparate viewpoint, consistent with expectation. As discussed earlier, it is legitimate to have different results when taking into account team factors. It is of interest to note that perhaps the contrary relationship with tenure dissimilarity was latent in the zero-order correlations for H1, as it was noted then that tenure dissimilarity was the only predictor variable that was negative, although it was not significant. What the multi-level model analysis reveals is that when team context is considered, tenure dissimilarity is the only, negative, predictor of cognitive dissimilarity post-discussion.

8.2.5.2 Diversity Effects: Cognitive Dissimilarity Post-Discussion

Next, in step 3, the five team diversity predictors are added in turn. Only educational diversity is significant at (Δ*2L-Li = 3.30, p <0.05), and achieves a reduction in the team level variance (σ^2_u 4.82). The positive slope coefficient reveals that educationally diverse teams take ranking distance values β = 8.05, (p <0.05) points higher than the mean.

This result supports the first strand of H3 that demographic diversity at the team level explains individual cognitive differences. Following Hox (1995), this will be taken forward to step 4, which will determine whether educational diversity explains a proportion of the variance over and above tenure dissimilarity.

Table 8.4 Cognitive Dissimilarity Post-discussion and Demographic Variation Steps 1 - 5

Model Steps & Predictors	Intercept Coefficient β	(SEβ)	Slope #1 Dissimilarity Coefficient β	(SEβ)	Slope #2 Diversity Coefficient	(SEβ)	Slope #3 Interaction Coefficient β	(SEβ)	Level 2 Variance (Between Teams) σ^2_u	(SE)	Level 1 Variance (Between Individuals) σ^2_e	(SE)	Deviance	Change in Deviance From Previous Step
Step 1: Variance Components;														
(null model)	4.00	(0.52)	n/a	n/a	n/a	n/a	n/a	n/a	5.55	(1.84)	3.47	(0.47)	582.77	n/a
Step 2: Individual Predictors:														
Age Dissimilarity	4.01	(0.52)	0.01	(0.02)	n/a	n/a	n/a	n/a	5.49	(1.81)	3.47	(0.47)	582.72	0.05
Functional Dissimilarity	4.08	(0.50)	0.27	(0.16)	n/a	n/a	n/a	n/a	5.00	(1.68)	3.45	(0.47)	580.20	2.57
Educational Dissimilarity	4.04	(0.52)	0.04	(0.07)	n/a	n/a	n/a	n/a	5.37	(1.78)	3.48	(0.47)	582.45	0.32
Gender Dissimilarity	4.00	(0.52)	-0.58	(1.06)	n/a	n/a	n/a	n/a	5.59	(1.84)	3.45	(0.50)	582.46	0.31
Tenure Dissimilarity	4.01	(0.53)	-0.01*	(0.00)	n/a	n/a	n/a	n/a	5.35	(1.85)	3.27	(0.45)	577.33	5.44*
Step 3: Team Level Predictors;														
Age Diversity	4.03	(0.52)	n/a	n/a	1.72	(1.83)	n/a	n/a	5.52	(1.83)	3.47	(0.47)	582.66	0.11
Functional Diversity	4.01	(0.52)	n/a	n/a	0.18	(2.44)	n/a	n/a	5.55	(1.84)	3.47	(0.47)	582.76	0.01
Educational Diversity	4.08	(0.49)	n/a	n/a	8.05*	(4.65)	n/a	n/a	4.82	(1.59)	3.47	(0.47)	579.47	3.30+
Gender Diversity	4.01	(0.52)	n/a	n/a	0.51	(3.03)	n/a	n/a	5.54	(1.83)	3.47	(0.47)	582.74	0.03
Tenure Diversity	4.03	(0.52)	n/a	n/a	2.06	(3.59)	n/a	n/a	5.47	(1.81)	3.47	(0.47)	582.42	0.35
Step 4: Main Effects:														
Tenure Dissim. + Educational Diversity	4.07	(0.50)	-0.01*	(0.00)	7.25	(4.79)	n/a	n/a	4.17	(1.62)	3.47	(0.47)	575.17	2.16
Step 5: Interactions:														
Tenure Diss. * Educational Diversity	4.07	(0.50)	-0.01	(0.00)	7.25	(4.79)	0.00	(0.00)	5.16	(1.70)	3.28	(0.45)	575.17	0.00

* significant at p <0.05; and ** significant at p <0.01, one-tailed. Steps 2: d/f 2,117 to 2,127; Step 3: d/f, 1,118 to 1,128; Steps 4 & 5: d/f 2,117 to 2,127 and 3,116 to 3,126 depending on missing values. Δ*2L_li for 2 way interactions at Step 5 are compared against main effects at Step 4 as a proxy for full formula of 2 main effects and 2 two-way interactions.

8.2.5.3 Additive Effects of Diversity: Cognitive Dissimilarity Post-Discussion

Table 8.4 reports the main effects modelling (step 4) of educational diversity and tenure dissimilarity together on cognitive dissimilarity post-discussion. The relationship is not statistically significant.

8.2.5.4 Interaction Effects: Cognitive Dissimilarity Post-Discussion

Table 8.4 reports that the interaction term is clearly not significant. Indeed, the coefficients and the reduction in the deviance are virtually identical to the main effects model.

8.2.5.5 Summary: Cognitive Dissimilarity Post-Discussion

In summary, the findings for this section support the first argument posited in the hypothesis that demographic diversity predicts cognitive dissimilarity post-discussion, but only with respect to educational diversity. Moreover, the second argument, that demographically dissimilar individuals in diverse teams will select dissimilar rankings to their peers after a discussion, is not supported.

The next section investigates the final post-discussion measure, cognitive cohesion dissimilarity.

8.2.6 Null Model: Cognitive Cohesion Dissimilarity Post-Discussion

The fifth cognitive individual response variable to be investigated using the multi-level model is cognitive cohesion dissimilarity post-discussion. As detailed in previous chapters, cognitive cohesion dissimilarity measures the proximity of the individuals' post-discussion ranking to their team's consensus ranking. It is of particular interest to measure cohesion post-discussion, as by this stage all individuals know what the team consensus ranking is, and the measure computes the extent to which they concur.

Table 8.5 Cognitive Cohesion Dissimilarity Post-discussion and Demographic Variation Variance Components Model

	Intercept Coefficient		Slope #1 Dissimilarity Coefficient		Slope #2 Diversity Coefficient		Slope #3 Interaction Coefficient		Level 2 Variance (Between Teams)		Level 1 Variance (Between Individuals)		Deviance
Model Steps & Predictors	β	(SEβ)	β	(SEβ)		(SEβ)	β	(SEβ)	σ^2_u	(SE)	σ^2_e	(SE)	
Step 1: Variance Components:													
(null model)	0.64	(0.04)	n/a	n/a	n/a	n/a	n/a	n/a	0.00	(0.00)	0.21	(0.04)	162.99

By referring to the variance components section of Table 8.5, it can be seen that the overall mean of cognitive cohesion dissimilarity post-discussion is estimated as $\beta_0 = 0.64$. The means for the different teams are distributed around the overall mean with an estimated variance of σ^2_u 0.00, whilst the between individuals variance is σ^2_e 0.21 (0.04). This means that almost all of the variance is at the individual level and there is perhaps a trace element attributable to teams. The variance at the team level is so tiny as to negate the multi-level modelling process.

8.2.6.1 Summary: Cognitive Cohesion Dissimilarity Post-Discussion

In summary, the findings for this section do not support the hypothesis. There is virtually no level 2 variance between the teams to account for, which there would be if team level factors were salient. Specifically this means that demographic diversity at the team level cannot predict individual cognitive variation in cognitive cohesion dissimilarity post-discussion. Moreover, there can be no effects of dissimilarity and diversity together in relation to this response variable.

8.3 Conclusion

Essentially there were 25 core relationships tested in H3 (five cognitive dissimilarity variables x five demographic diversity variables) from which five statistically significant results that support the hypothesis are achieved, slightly more than the number expected by chance alone.

Additional relationships were also tested (five demographic dissimilarity variables within the context of teams, and additive main effects and interactions) yet only two of these were statistically significant in support of the hypothesis.

Specifically, the findings in support of the first strand of H3 are that: tenure diverse teams will have the least proximity to the team consensus pre-discussion; functionally, educationally and gender diverse teams will experience the least change in opinions; and educationally diverse teams experience more cognitive dissimilarity post-discussion. There is no clear pattern in these results as to which type of demographic diversity consistently affects cognitive dissimilarity, and indeed, two of the five cognitive dissimilarity variables were not related to demographic diversity at all.

The findings for the second strand of H3, that demographic diversity at the team level would interact with individual demographic dissimilarity to affect cognitive dissimilarity, is not supported. Although it was found that gender dissimilar individuals tend not to change their minds, this was a discrete relationship that was not moderated by the effects of demographic diversity. Concerning this latter finding, it is interesting to note that the least studied demographic factor in TMT demographic variation research is gender (See Chapter 2 for a review). There is clearly much work still to do.

The last two chapters have considered cognitive variation through reference to rankings of candidates pre- and post-discussion. Of particular interest now is the team discussion. The question which arises is: to what extent do team processes exhibited during the discussion affect the way different people in the team rank the candidates? It is to this topic that the next chapter now turns.

CHAPTER 9

Findings: Hypotheses 4 - 6

9.0 Overview

This chapter presents the findings for hypotheses 4 to 6, dealing with the relationship between diversity (demographic and cognitive) and team processes at the individual and team levels (as shown in the guiding conceptual model on p 76).

9.1 Introduction

Four aspects of team process were put forward in Chapter 5 that are likely to influence demographic variation and cognitive variation. These were procedural rationality (analytical and systematic consideration of information); frequency of team meetings; reflexivity (objectively reviewing processes without expectation of censure); and psychological safety.

Hypothesis 4 (H4) focuses on the team level (n = 23), predicting a negative relationship between demographic diversity and all four of the team processes. This is because it is widely argued that demographic diversity increases conflict (Jackson, 1996) and disagreement amongst team members (Souder, 1987). Specifically, H4 proposes that those teams characterized by greater variation in age, functional background, educational attainment, gender and tenure, will tend to be more intuitive rather than analytical in their decision-making process, are likely to meet less often, to be the least reflexive, and to experience the least psychological safety.

Hypothesis 5 (H5) also focuses on the team level, predicting a relationship between cognitive diversity and team processes. As discussed in Chapter 4, some team processes (such as systematic decision processes and frequency of team meetings) are widely held to be beneficial in assuaging negative aspects of team functioning. Specifically, H5 argues that teams which experience less cognitive variation (i.e. more agreement in how they rank candidates), will use procedural rationality for strategic decision making, and tend to meet more frequently. Further, as discussed in Chapter 3, creative or effective decisions are usually considered to be those that are based upon conflicting viewpoints within the team (Nemeth & Owens, 1996). Processes argued to facilitate constructive yet conflicting views are reflexivity and psychological safety (West et al., 1997; Tjosvold, 1996). Hence, a positive relationship is expected between reflexivity and psychological safety on the one hand, and cognitive variation on the other. That is to say, teams that are reflexive (i.e. those that objectively identify problems without expectation of censure) and enjoy high levels of psychological safety will show the greatest diversity (least agreement) in how they rank the candidates.

Hypothesis 6 (H6) which transcends both the team and individual levels of analysis, is that team processes will predict individual cognitive dissimilarity. Procedural rationality and frequency of team meetings are expected to engender more agreement amongst team members, whilst it is expected that individuals in reflexive teams and psychologically safe teams, will show greater difference in their rank order of candidates relative to those in less reflexive or less safe teams. It is also of interest to investigate whether there are cross-level interactions between the aspects of demographic variation found to influence cognitive dissimilarity in the previous chapter, and team processes. For example, findings testing H3 found that educationally dissimilar individuals experienced more cognitive dissimilarity pre-discussion. Part of the analysis for H6 will test for whether there is an interaction between educational dissimilarity, cognitive dissimilarity pre-discussion and team processes.

As this hypothesis transcends the two levels of analysis involving individual and team independent variables with an individual dependent variable, it is addressed using multilevel modeling.

For each of the hypotheses, statistical considerations are presented first, followed by zero-order correlations and multiple regression analyses, or multi-level modeling for H6. Within each relevant section, the two pre-discussion measures of cognitive variation are reported first, followed by cognitive change, and the two post-discussion measures.

9.2 The Relationship Between Demographic Diversity and Team Processes (H4)

9.2.1 Statistical Considerations

For H4 the team is the unit of analysis and the sample comprises 23 teams. As reported in Chapter 6, the underlying distributions of the demographic diversity variables were sufficiently close to normal to allow the use of parametric statistics. An investigation was made of the team process variables, and they too fell within the tolerances for normality meaning that parametric analytic methods were appropriate. Initial testing of H4 involved determining the extent to which each of the five demographic diversity variables is related to the four team process measures. Zero-order correlations are reported in the section below (see Table 9.1, pink box).

9.2.2 Zero-order Correlations

Preliminary findings concerning the relationship of the demographic diversity variables and procedural rationality are such that four of the five possible relationships are negative as expected. That with gender diversity is substantial and statistically significant, $r = -0.35$, $p < 0.05$. This finding means that gender diverse teams are more intuitive and less analytical than all male teams. As this is a single predictor, and neither gender diversity nor procedural rationality is related to team size, no multiple regression analysis will be conducted. It can be said at this point that gender diversity explains 12% of the variance in procedural rationality.

Initial findings concerning demographic diversity and the frequency of team meetings are that all five of the possible relationships are in the expected (negative) direction, with functional and gender diversity being substantial and statistically significant ($r = -0.47$, $p < 0.05$ and $r = -0.39$, $p < 0.05$ respectively). These findings mean that functionally diverse TMTs and mixed gender teams meet less frequently than functionally homogeneous and all male teams. This is consistent with a finding by Ancona & Caldwell (1992) that functionally diverse groups tend to communicate more outside the boundaries of formal team meetings, perhaps negating the need for more, regular formal meetings. Of course, they were not investigating TMTs, but the substance of the finding is similar.

Table 9.1 Zero-order Correlation Coefficients for Hypotheses 4, 5, 7, 8 and 9

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.
1. Age Diversity	-																
2. Functional Diversity	.15	-															
3. Educational Diversity	.15	.34†	-														
4. Gender Diversity	.33†	.19	.29†	-													
5. Tenure Diversity	.26	-.02	-.10	-.31†	-												
6. Cognitive Diversity Pre-	-.03	.16	.15	-.02	.31†	-											
7. Cognitive Cohesion Diversity Pre-	-.08	-.47*	-.19	-.14	.25	.01	-										
8. Cognitive Change Diversity	-.16	-.63**	-.22	-.24	-.06	-.14	.80**	-									
9. Cognitive Diversity Post-	.19	.56**	.42*	.37*	.05	.39*	-.60**	-.79**	-								
10. Cognitive Cohesion Div. Post-	.15	.12	.04	.07	.56**	.23	.55**	-.04	.16	-							
11. Procedural Rationality	-.26	-.14	.06	-.35*	-.06	.02	-.22	-.02	.15	-.24	-						
12. Frequency of Team Meetings	-.20	-.47*	-.27	-.39*	-.07	-.44*	.31†	.43*	-.66**	-.16	-.02	-					
13. Reflexivity	-.32†	.08	.09	-.01	-.18	-.06	-.05	.01	.23	-.02	.33†	-.20	-				
14. psychological safety	-.21	.10	-.08	-.45*	.02	-.00	.20	.30†	-.28	-.06	.31†	.33†	.13	-			
15. Satisfaction	-.04	-.07	-.41*	-.23	.33†	-.07	-.10	-.28†	-.03	.24	.13	.20	-.15	-.01	-		
16. Confidence	.08	-.38*	-.39*	-.33†	.53**	.06	.30†	.15	-.46**	.26	-.10	.28†	-.40*	.02	.53**	-	
17. Perceived Effectiveness	.04	-.34†	-.39*	-.26	.24	-.05	.20	.12	-.33†	.16	.07	.34*	-.21	.18	.56**	.68**	-
18. Team Size	.36*	.44*	.39*	.16	.44*	.55**	-.02	-.32	.65**	.52**	.22	-.62**	.20	.04	-.17	-.20	-.22

N = 23; † = *p* <0.10; * = *p* <0.05; ** = *p* <0.01

H 4= ▢ H 5 = ▢ H 7 = ▢ H 8 = ▢ H 9 = ▢

The findings relating to reflexivity also offer a modicum of support for H4, with one relationship, that with age diversity achieving significance (r = -0.32, p < 0.10). This means that people in mixed age teams are less likely to reflect upon their processes or to challenge the way things are done in the team or the organisation. As age diversity is positively associated with team size, a regression analysis between age diversity and reflexivity whilst controlling for team size is appropriate. This will be reported on in the next section.

Finally, for psychological safety, the findings offer modest support for the hypothesis in that the bivariate correlation with gender diversity is substantial and negative (r = -.45, p <0.05). This finding means that people in mixed gender teams feel more uncomfortable about freely sharing ideas and information than those in all male teams. Neither psychological safety nor gender diversity is significantly associated with team size, which means that men and women feel just as safe in smaller or larger teams. Gender diversity explains 20% of the variance in psychological safety.

In summary, initial findings presented in this section support H4 in some respects. Negative relationships are observed between some aspects of demographic diversity and team processes. Specifically, gender diversity is negatively associated with procedural rationality, frequency of team meetings, and psychological safety. Functional diversity also is negatively associated with frequency of team meetings, whilst age diversity is negatively associated with reflexivity.

Multiple regression analyses were conducted to test the preliminary finding regarding diversity and team processes more stringently. These are reported in the next section.

9.2.3 Multiple Regression Analyses for H4

The results of the multiple regression analyses for H4 are shown in Table 9.2. For each of the team process dependent variables, the analyses were conducted using those demographic diversity variables found to correlate with them in the zero-order analyses. Where the zero-order correlations indicated that there could be a potentially confounding association with team size, this was entered first and the demographic explanatory variables in a second step.

Table 9.2 Multiple Regression Analyses for the Effects of Demographic Diversity on Team Processes

Outcome Variable	Predictor Variable	β	$R^2 \Delta$	F	df
Frequency of Team Meetings	1. Team Size	-.51*	.38	13.07**	1,21
	2. Functional Diversity	-.24	.05	7.58**	2,20
Frequency of Team Meetings	1. Team Size	-.57**	.38	13.07**	1,21
	2. Gender Diversity	-.30†	.09	8.98**	2,20
Reflexivity	1. Team Size	.36	.04	.86	1,21
	2. Age Diversity	-.45*	.19	2.71*	2,20

N = 23; † = p <0.10; * = p <0.05; ** = p <0.01. One-tailed tests of significance levels for standardised β coefficients in final equation. $R^2 \Delta$ refers to discrete steps.
No entry for procedural rationality nor psychological safety as there is only one predictor in each case

In the first analysis, functional diversity was regressed onto frequency of team meetings. Table 9.2 shows that once team size is controlled for, functional diversity is no longer statistically significant. The beta coefficients reveal that team size is the unique statistically significant predictor (β = -0.51, p<0.05) and explains 38% of the variance.

In the second analysis, gender diversity was regressed onto frequency of team meetings. Table 9.2 shows that once team size is controlled for, gender diversity is still a statistically significant predictor (β = -0.30, p<0.10) and explains 9% of the variance.

In summary, these two findings mean that larger teams, and gender diverse teams, meet less frequently. The former is not surprising. As reported in Chapter 6, it was very difficult to arrange a full TMT meeting for the purposes of this research, a likely indication of regular experience in these organisations. The initial correlations that suggested that functional diversity and gender diversity were also associated with less frequent meetings are probably not inconsistent either. With larger teams incorporating more diverse functional backgrounds it is not difficult to imagine that schedules for production may not coincide with marketing or dispatch for example, so that getting the directors of these departments together for a TMT meeting would not be as easy as in a smaller team where the scope of responsibility for one director may be narrower. As to why gender diverse teams meet less often, one might surmise that this too may be related to functional diversity, perhaps with a gender split across the functions. In one's own experience, despite greater equality in the workplace, women executives still typically tend to hold positions in the soft management functions (e.g. marketing, HR, finance) where as men typically hold the traditional functions (e.g. engineering, production). However, this explanation is unlikely to be valid, because gender diversity and functional diversity are not correlated in this study.

A regression analysis was conducted to ascertain whether age diversity was a unique predictor of reflexivity (and not confounded by team size). The bivariate correlation reported in Table 9.1 showed that age diversity is positively associated with team size, and associated in the expected negative direction with reflexivity. Team size was entered first, followed by age diversity. The results, reported in Table 9.2 show that once team size is controlled for, age diversity singly explains 21% of the variance in reflexivity (β = -.45, p<0.05), confirming the earlier supposition that mixed age teams are less reflexive, that is to say, they avoid discussion of problems in order to preserve morale (West, 1996).

In summary, the findings reported in this section are that demographic diversity is negatively associated with team processes. Specifically, the findings are that gender diversity is a negative predictor of procedural rationality, frequency of team meetings and psychological safety, whilst age diversity is a negative predictor of reflexivity. The next section will present the conclusions regarding the findings for H4.

9.2.4 Conclusions Regarding Hypothesis 4

H4 received some support, but only 4 findings emerged out of a possible 20. All conformed to prediction, and were statistically significant. Gender diversity is a negative predictor of frequency of team meetings, procedural rationality and psychological safety, whilst age diversity is a negative predictor of reflexivity. The fact that other demographic diversity factors did not contribute in the same way, may be due to the limited power in a relatively small dataset.

On the other hand, it may be that age and gender diversity are the real keys to understanding team process. The potential for dissonance between older, more risk averse managers and their youthful, creative risk-taking counterparts, has been alluded to in the literature (e.g. Hambrick & Mason, 1984). It has also been suggested that age diverse teams may experience more internal conflict (Williams & O'Reilly, 1998). The finding from this study bears out both of these assumptions. It suggests that age heterogeneity means that teams do not talk about their problems or seek better ways of doing things.

As regards gender diversity being a negative predictor of procedural rationality, this is consistent with studies of cognition that find that women generally speaking are more intuitive whilst men are more analytical (CfWBR, 1994). However, the same caveat as offered previously is current, that is, the finding concerns gender diverse teams, not men or women in teams. As before, this is an important finding considering that gender diversity in TMTs has not been studied before.

The finding that gender diversity is a negative predictor of psychological safety is consistent with studies of the effects of being in a gender minority (Tsui et al., 1992). However, what should be remembered about this study is that it is investigating team level gender diversity (i.e. not female minorities in male teams), and the measure of psychological safety was determined by independent observers. The finding is that reduced levels of psychological safety are experienced by mixed gender teams. This finding makes an important contribution to the literature, which at this level in organisations has not studied gender diversity before.

9.3 The Relationship Between Cognitive Diversity and Team Process (H5)

9.3.1 Statistical Considerations

For H5 the team is once again the unit of analysis and the sample comprises 23 teams. As reported in Chapter 6, the underlying distribution of the cognitive diversity variables was sufficiently close to normal to allow the use of parametric statistics with the exception of cognitive cohesion diversity pre-discussion, from which two outliers at the extremities of the data were dropped in order to meet parametric requirements. All four team process variables were normally distributed as reported above.

9.3.2 Zero-order Correlations

In Chapter 4, H5 was stated as team level procedural rationality and frequency of team meetings will be negatively associated with cognitive diversity (i.e. more agreement), whilst reflexivity and psychological safety will be positively associated with cognitive diversity (i.e. less agreement). The first two team processes are supposed to mitigate negative communication difficulties and help facilitate team consensus (Sniezek & Henry, 1988). Therefore, cognitive variation should be less. Only with respect to cognitive change diversity would one expect a positive relationship. This is because, as explained in Chapter 6, (p 126), cognitive change diversity is a measure of the extent to which people within the team change their minds. In order to achieve a consensus, individuals in the team would need to be willing to compromise on their preferences, that is to say, bring their opinions in line with their peers (i.e. more agreement). Therefore a positive relationship between procedural rationality, frequency of team meetings and cognitive change diversity (team level) is consistent with the hypothesis.

The second two team processes, reflexivity and psychological safety, are supposed to increase constructive controversy and variety of opinions (West et al., 1997; Tjosvold, 1996). Therefore, cognitive variation should increase. Only with respect to cognitive change diversity would one expect a negative association. This is because, in psychologically safe and reflexive teams, it should be acceptable to hold on to one's own opinion. This means that there should be less cognitive shift (i.e. less agreement) in teams that are characterized by reflexivity and psychological safety. The four processes will be addressed in turn, starting with the two expected negative predictors of cognitive variation, procedural rationality and frequency of team meetings.

As shown in Table 9.1 (green box), procedural rationality has no relationships that achieve statistical significance, nor are any approaching significance. This means that the type of decision process regularly used by a team is not important for reducing cognitive variation. With respect to frequency of team meetings, the preliminary findings are that three of the possible five relationships are statistically significant in support of the hypothesis, and a further one is statistically significant but in the opposite direction.

The expected negative relationship was observed between frequency of team meetings and cognitive diversity pre-discussion (r = -.44, $p<0.05$). This means that there is less diversity of rankings prior to a team discussion in teams which meet frequently. In other words, teams which meet more often are more likely to reach a similar conclusion amongst members, even when they have not yet met to discuss the matter. In this case, there was less variation around the candidate selection, pre-discussion, in teams which meet between three and six times per month. The expected positive relationship was also observed between frequency of team meetings and cognitive change diversity (r = .43, $p<0.05$). This means that people change their candidate rankings more after a discussion in teams which meet more often. The expected negative relationship was also observed between frequency of team meetings and cognitive diversity post-discussion (r = -.66, $p<0.01$). This means that the more often a team meets, the less variation there will be in candidate selection after a team discussion.

An unexpected positive relationship was observed between frequency of team meetings and cognitive cohesion diversity pre-discussion (r = .31, $p<0.10$). This means that there is more variation away from the team consensus pre-discussion in teams which meet frequently, explaining 9% of the variance.

In summary, the initial findings for this section are mixed. On the one hand, frequency of team meetings appears to be important for reducing cognitive diversity.

Indeed, in teams which meet often, it seems that individuals tend to look at the information and make similar rankings even when they have not yet held a meeting to discuss the matter. Post-discussion, there is even less cognitive variation in teams which meet often. On the other hand, and contrary to prediction, frequency of team meetings appears to increase variation away from the team consensus pre-discussion (i.e. less agreement) but does not influence variation in proximity to the team consensus post-discussion. Both the cognitive diversity measures are associated with team size, hence, further analyses will control for this potentially confounding factor.

With respect to reflexivity and psychological safety, mixed support is evidenced in the zero-order correlations. No statistically significant relationships emerge between reflexivity and the cognitive diversity variables. Only one relationship with psychological safety, that with cognitive change diversity is statistically significant ($r = 0.30, p < 0.10$). This means that psychological safety is a facilitator of cognitive shift around individual rankings in teams as predicted. As neither variable is associated with team size, it is not necessary to conduct a regression analysis in this case. Psychological safety explains 9% of the variance in cognitive change diversity.

In summary, the findings for this section are that reflexivity is not important for increasing cognitive variation and that psychological safety only increases variation with respect to people changing their minds. That is to say, in teams characterised by psychological safety, individuals' personal opinions are more likely to undergo change.

In order to control for third factors, explain the variance in the dependent variables using more than one predictor, and to ascertain the unique effects of predictors, regression analyses for frequency of team meetings and cognitive diversity were conducted. These are reported in the next section.

9.3.3 Multiple Regression Analyses for H5

Table 9.3 shows that cognitive diversity pre- and post-discussion are positively related to team size. Therefore, for findings with this dependent variable team size was controlled for by entering it first in the regression equation. This was followed by entering frequency of team meetings. The same procedure was applied to cognitive change diversity. The results for all three regression analyses are reported in Table 9.3.

The regression analysis for cognitive diversity pre-discussion shows that once team size is controlled for, frequency of team meetings is no longer statistically significant, with team size itself being statistically significant ($\beta = 0.46, p < 0.10$).

The regression analysis for cognitive change diversity shows that once team size is controlled for, frequency of team meetings is no longer statistically significant. This means that frequency of team meetings does not predict a propensity towards compromise, that is, it does not predict cognitive shift with regard to change in individuals personal rankings in the team.

By way of contrast, for cognitive diversity post-discussion, the effect of frequency of team meetings is not a function of team size and singly explains 10% of the variance, (β -0.41, p<0.05). This means that teams which meet often will experience less cognitive diversity after a team discussion irrespective of size.

In summary, the findings for this section partially support the hypothesis, but only in respect of cognitive diversity post-discussion and frequency of team meetings. Teams which meet between three and six times per month experience less cognitive variation after a team discussion.

9.3.4 Conclusions Regarding Hypothesis 5

H5 put forward two strands, the first that certain team processes (procedural rationality and frequency of team meetings) would facilitate less cognitive variation, and the second that other processes (reflexivity and psychological safety) would increase cognitive variation. Four out of 20 relationships emerged in the zero-order correlations to support the hypothesis, whilst another was in the opposite direction. After being subjected to stringent regression analyses, only two relationships are statistically significant and in the right direction, whilst one opposes the hypothesis.

Specifically, the findings are that teams which meet more often experience less cognitive diversity after a discussion, and teams characterized by psychological safety experience more cognitive change. Contrary to expectation, there is less proximity to the team consensus after a discussion in teams which meet more often. The real conclusion is that there are occasional isolated effects between team processes and cognitive diversity, but many more predictions are unsubstantiated.

The sections of this chapter so far have considered the team level proposition for the effect of team processes on demographic and cognitive diversity. The next section will present analyses which test the team level process effects on individual level cognitive dissimilarity.

Table 9.3 Multiple Regression Analyses for the Effects of Team Processes on Cognitive Diversity

Outcome Variable	Predictor Variable	β	$R^2 \Delta$	F	Df
Cognitive diversity pre-discussion	1. Team Size	.46†	.31	9.25**	1,21
	2. Frequency of Team Meetings	-.16	.01	4.73*	2,20
Cognitive change diversity	1. Team Size	-.09	.10	2.36	1,21
	2. Frequency of Team Meetings	.37	.09	2.32	2,20
Cognitive diversity post-discussion	1. Team Size	.40*	.43	15.57**	1,21
	2. Frequency of Team Meetings	-.41*	.10	11.18**	2,20

N = 23; † = p <0.10; * = p <0.05; ** = p <0.01. One-tailed tests of significance levels for standardised β coefficients in final equation. R^2 Δ refers to discrete steps. No entry for cognitive cohesion diversity pre-discussion as only one predictor.

9.4 The Relationship Between Team Processes & Individual Cognitive Variation (H6)

H6 is that team processes will predict individual cognitive dissimilarity. Two processes were predicted to *reduce* cognitive dissimilarity (i.e. foster greater agreement), procedural rationality and frequency of meetings. Two processes were predicted to *increase* cognitive dissimilarity (i.e. more disagreement), reflexivity and psychological safety.

The following sections of this chapter will apply the multi-level modelling procedure described in Chapter 9 to the cognitive dissimilarity variables along with the team process variables. The two pre-discussion measures will be addressed first, then cognitive change dissimilarity, followed by the two post-discussion measures.

9.4.1 Statistical Considerations

The variables used to test H6 are the individual cognitive dissimilarity dependent variables used for H1 and H3 as described in Chapter 7, and the team independent variables of procedural rationality, frequency of meetings, reflexivity and psychological safety. The pattern of the multi-level modelling follows that for H3 in Chapter 8 (pp 157 - 161). The variance components models (step 1) for the individual cognitive dissimilarity variables and the deviance statistics have already been established in Tables 8.1, 8.2, 8.3, 8.4, and 8.5, however, they are repeated again in the corresponding tables[1] (9.4, 9.5, 9.6 and 9.7) for the sake of convenience. Furthermore, demographic predictors from steps 2, 3 and 4 from Chapter 8 (dissimilarity and diversity effects entered separately and then together), are included in the tables here, because these provide the baseline against which the team process factors are being tested.

The tables in this chapter follow the same overall format as those in Chapter 8. However, in several cases below there are multiple predictors (demographic dissimilarity, demographic diversity and team processes), so the slope coefficients are labelled #1 - #4 at the top of each column. Slope coefficients refer to the predictors in the order they appear in the Table and are entered into the analysis.

As discussed in Chapter 8, one is looking for a statistically significant reduction in the deviance (which is an indicator of model fit), a reduction in the between teams variance (which means that the predictor explains variance between teams) and a statistically significant slope coefficient (which shows the direction of the effect of the predictor on the response variable).

9.4.2 Team Processes and Cognitive Dissimilarity Pre-Discussion

The first cognitive individual response variable to be investigated is cognitive dissimilarity pre-discussion. As established in Chapter 8 (section 8.3.2), the deviance is 567.82 and the variance partition coefficient (VPC) is 46%, which is attributable to between team differences. Also established in Chapter 8 (and repeated in section 2 of Table 9.4), none of the individual dissimilarity variables, nor the diversity variables was statistically significantly related to cognitive dissimilarity pre-discussion. Hence, these are not included in the model, which means that there is only one slope coefficient column in the analysis of team processes.

---

[1] There is no corresponding table to 8.5, as it was established earlier that the amount of team level [illegible] explained for cognitive cohesion dissimilarity is too tiny to model.

Table 9.4 Cognitive Dissimilarity Pre-discussion and Team Processes Steps 1 - 5

Model Steps & Predictors	Intercept Coefficient β	(SEβ)	Slope #1 Coefficient β	(SEβ)	Level 2 Variance (Between Teams) σ^2_u	(SE)	Level 1 Variance (Between Individuals) σ^2_e	(SE)	Deviance	Change in Deviance From Previous Step
Step 1: Variance Components:										
(null model)	4.59	(0.40)	n/a	n/a	2.93	(1.05)	3.39	(0.47)	567.82	n/a
Step 2: Demographic Predictors (Established in Chapter 8):										
Nil	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Step 3: Team Level Predictors:										
Procedural Rationality	4.56	(0.37)	0.26	(0.36)	2.87	(1.04)	3.39	(0.46)	567.30	0.52
Frequency of Meetings	4.65	(0.39)	-0.82*	(0.33)	2.56	(0.94)	3.38	(0.46)	562.40	5.42**
Reflexivity	4.59	(0.39)	0.72	(0.92)	2.83	(1.03)	3.40	(0.46)	567.21	0.61
Psychological Safety	4.59	(0.40)	-0.02	(0.61)	2.94	(1.05)	3.39	(0.47)	567.82	0.00
Step 4: Main Effects (Individual Dissimilarity and Team Process Predictors Together):										
Not Applicable	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Step 5: Cross-level Interactions:										
Not Applicable	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

* significant at p <0.05; and ** significant at p <0.01, one-tailed. Step 3: d/f 1,118 to 1,128 depending on missing values.

9.4.2.1 Team Process Effects: Cognitive Dissimilarity Pre-Discussion

Table 9.4 reports the addition of team processes as fixed effects[2]. The only statistically significant result, for frequency of team meetings, reduces the deviance by 5.42 (p <0.01) to 562.40. The between team variance is reduced to σ^2_u 2.56 (0.94), which means that frequency of team meetings explains some of the variance between teams. The slope coefficient is negative and statistically significant (β = -0.82, p <0.05), which conforms to prediction. It means that in teams which meet more often, individuals will take distance values on ranking pre-discussion -0.82 below the mean. In other words, the level of disagreement in teams that meet more frequently is less acute.

9.4.2.2 Summary: Team Processes and Cognitive Dissimilarity Pre-Discussion

In summary, the results for this section provide very limited and partial support for H6. Specifically, frequency of meetings, reduces the levels of disagreement amongst team members' rankings pre-discussion. However, as this was the only team process that had a statistically significant effect, it provides far from convincing support for the hypothesis overall. The next section will investigate the second pre-discussion measure, cognitive cohesion dissimilarity.

9.4.3 Team Processes and Cognitive Cohesion Dissimilarity Pre-Discussion

As detailed in Chapter 6 (pp 123 – 125), cognitive cohesion dissimilarity measures the proximity of the individuals' pre-discussion ranking to their teams' ultimate consensus ranking on a scale of -1.0 to 1.0, where 1.0 corresponds to a complete match with the team. As explained on page 113, it is expected that the association of cognitive cohesion dissimilarity pre-discussion with procedural rationality and frequency of meetings will be positive, whilst the association with reflexivity and psychological safety and this response variable will be negative. That is to say, the former processes will tend towards more agreement (1.0) whilst the latter will tend towards more disagreement (-1.0).

---

2 As discussed in Chapter 8 (p 148), analysis allowing varying effects gives the equivalent pattern of findings.

Table 9.5 Cognitive Cohesion Dissimilarity Pre-discussion and Team Processes Steps 1 - 5

Model Steps & Predictors	Intercept Coefficient		Slope #1 Coefficient		Slope #2 Coefficient		Slope #3 Coefficient		Slope #4 Coefficient		Level 2 Variance (Between Teams)		Level 1 Variance (Between Individuals)		Dev.	Δ*2L_li
	β	(SEβ)	β	(SEβ)	β	(SEβ)	β	(SEβ)	β	(SEβ)	σ^2_u	(SE)	σ^2_e	(SE)		
Step 1: Variance Components: (Null)	0.36	(0.05)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.01	(0.00)	0.32	(0.04)	225.64	-
Step 2: **Demographic Predictors (Established in Chapter 8):**																
Tenure Dissimilarity	0.36	(0.05)	-0.00*	(0.00)	n/a	n/a	n/a	n/a	n/a	n/a	0.00	(0.00)	0.31	(0.04)	222.13	3.51*
Tenure Diversity	0.36	(0.05)	-0.65*	(0.38)	n/a	n/a	n/a	n/a	n/a	n/a	0.01	(0.00)	0.32	(0.04)	222.87	2.77*
Tenure Dissim. + Tenure Diversity	0.36	(0.05)	-0.00*	(0.00)	-0.45	(0.40)	n/a	n/a	n/a	n/a	0.00	(0.00)	0.32	(0.04)	220.86	1.27
Tenure Dissim. * Tenure Diversity	0.37	(0.05)	-0.00	(0.00)	-0.40	(0.40)	0.00	(0.00)	n/a	n/a	0.00	(0.02)	0.32	(0.04)	220.86	0.00
Step 3: Team Process Predictors:																
Procedural Rationality	0.33	(0.05)	0.10*	(0.05)	n/a	n/a	n/a	n/a	n/a	n/a	0.00	(0.02)	0.32	(0.04)	221.93	3.71*
Frequency of Meetings	0.36	(0.06)	-0.01	(0.05)	n/a	n/a	n/a	n/a	n/a	n/a	0.01	(0.02)	0.32	(0.04)	225.63	0.01
Reflexivity	0.36	(0.06)	-0.04	(0.13)	n/a	n/a	n/a	n/a	n/a	n/a	0.01	(0.02)	0.32	(0.04)	225.54	0.10
Psychological Safety	0.36	(0.06)	0.04	(0.09)	n/a	n/a	n/a	n/a	n/a	n/a	0.01	(0.02)	0.32	(0.04)	225.42	0.22
Step 4: Main Effects:																
Tenure Dissim. + Procedural Rationality	0.35	(0.05)	-0.00*	(0.00)	0.09*	(0.05)	n/a	n/a	n/a	n/a	0.00	(0.00)	0.31	(0.04)	218.67	3.46*
Tenure Diversity + Proced. Rationality	0.34	(0.05)	-0.61*	(0.36)	0.09*	(0.05)	n/a	n/a	n/a	n/a	0.00	(0.00)	0.32	(0.04)	219.15	3.72*
Tenure Dissim. + Ten. Div. + Proced. Rat.	0.36	(0.05)	-0.00	(0.00)	-0.40	(0.39)	0.07	(0.05)	n/a	n/a	0.00	(0.00)	0.32	(0.04)	218.75	2.11
Step 5: Interactions:																
Tenure Dissim. * Procedural Rationality	0.35	(0.05)	-0.00*	(0.00)	0.09*	(0.05)	0.00	(0.00)	n/a	n/a	0.00	(0.00)	0.31	(0.04)	216.66	2.01
Tenure Diversity * Procedural Rationality	0.34	(0.05)	-0.61*	(0.37)	0.09*	(0.05)	-0.11	(0.39)	n/a	n/a	0.00	(0.00)	0.32	(0.04)	219.07	0.08
Tenure Dissim. * Diversity * Proced. Rat.	0.36	(0.05)	-0.00	(0.00)	-0.35	(0.39)	0.04	(0.05)	0.01*	(0.00)	0.00	(0.00)	0.32	(0.04)	214.07	2.59

* significant at p <0.05; and ** significant at p <0.01, one-tailed. Steps 2: d/f 2,117 to 2,127; Step 3: d/f, 1,118 to 1,128; Steps 4 & 5: d/f 2,117 to 2,127 and 3,116 to 3,126 depending on missing values. Δ*2L_li for 3 way interactions at Step 5 are compared against main effects at Step 4 as a proxy for full formula of 3 main effects and 3 two-way interactions.

By referring to the variance components section of Table 9.5, it can be seen that the overall mean of cognitive cohesion dissimilarity pre-discussion is estimated as $\beta_0 = 0.36$. The means for the different teams are distributed around the overall mean with an estimated variance of σ^2_u 0.01, whilst the between individuals variance is σ^2_e 0.32. This means that team level differences explain 3% of the variance. The deviance is 225.64. As described in Chapter 8, this is a very small portion of variance to explain, but as demographic factors did not explain all of this variance (tenure dissimilarity and tenure diversity were the only predictors), there is still variance to be explained that may be attributable to team processes.

In Step 2, the slope#1 coefficient refers to the first predictor named (demographic dissimilarity) and the slope#2 coefficient refers to the second predictor, at the team level (demographic diversity).

9.4.3.1 Team Process Effects: Cognitive Cohesion Dissimilarity Pre-Discussion

In step 3, each of the 4 team process predictors are added, and the slope coefficients are shown in slope #2 column. The finding for procedural rationality is statistically significant, reducing the deviance by 3.71, $p < 0.05$. The between team variance is reduced to σ^2_u 0.00. The slope coefficient is positive and statistically significant (β = 0.10, $p < 0.05$), which conforms to prediction. This result suggests that in teams where it is common practice to analytically and systematically review information for decision making, individuals' rankings are more likely to agree with the team consensus prior to discussion. In other words, agreement with the team consensus pre-discussion is more likely in teams that are analytical rather than intuitive. It is, however, a very small effect.

9.4.3.2 Additive Team Process Effects: Cognitive Cohesion Dissimilarity Pre-Discussion

The next step takes forward the statistically significant demographic predictors of cognitive cohesion dissimilarity pre-discussion established in Chapter 8, tenure dissimilarity and tenure diversity, to test for additive main effects of, and interactions with, team process variables. The slope coefficients are entered in columns #1, #2, or #3 according to the order in which the predictors are listed under step 4.

For two of the demographic predictors, tenure dissimilarity and tenure diversity, procedural rationality explains additional variance, in both cases there is a statistically significant reduction in the deviance (3.46 and 3.72 respectively, both p <0.05). In each case, the slopes (demographic variation and procedural rationality) are statistically significant and there is a reduction in the level 2 variance.

This means that procedural rationality explains variance in cognitive cohesion dissimilarity pre-discussion over and above that explained by tenure dissimilarity, and that explained by tenure diversity. It is tenure dissimilar individuals and those individuals in tenure diverse teams that will typically tend to disagree with the team consensus prior to discussion (as indicated by the negative slopes). But, in teams which use procedural rationality, individuals will be closer to the team consensus (as indicated by the positive slope). When all three are added into the model together, there is no significant reduction in the deviance, meaning that procedural rationality does not explain a statistically significant effect amount of variance over and above that explained by tenure dissimilarity and tenure diversity together.

9.4.3.3 Interaction Effects: Cognitive Cohesion Dissimilarity Pre-Discussion

The next step is to compute the interaction terms, which are shown in the lower section of Table 9.5. As previously, the slope coefficients are shown in columns #1 – #3 in the order listed, with the multiplicative term shown in column #4. None of the interaction terms is statistically significant.

9.4.3.4 Summary: Cognitive Cohesion Dissimilarity Pre-Discussion

In summary, the results for reported in this section provide a modicum of support for the hypothesis that team processes predict individual cognitive variation. Specifically, the findings are that procedural rationality alone is a statistically significant, positive predictor of cognitive cohesion dissimilarity pre-discussion. That is, procedural rationality means more agreement with the team consensus prior to discussion. Moreover, that effect is additional to either tenure dissimilarity or tenure diversity (but not both). However, one relationship out of a possible four is not totally convincing. Further, contrary to the second strand of the hypothesis, none of the team processes is statistically significant in increasing agreement with the team consensus amongst dissimilar individuals within diverse teams (the interaction effects).

9.4.4 Team Processes and Cognitive Change Dissimilarity

The third cognitive individual response variable to be investigated is cognitive change dissimilarity. As explained on page 115 cognitive change dissimilarity is measured on a scale of -1.0 to 1.0, where a value of 1.0 indicates zero change and -1.0 indicates the greatest variation. A negative relationship between team processes and the cognitive change dissimilarity variable is expected. It stands to reason that in order for a consensus decision to be reached, some individuals in the team will need to move away from their original opinions in order to have more agreement. Procedural rationality should provide a justification for individuals to change their opinions as the team engages in a systematic analysis of the information available. One would also expect that in teams which meet frequently, enhanced communication amongst members will result in more scope for people to change their minds (tend towards -1.0) based on new information or influence from colleagues. In order to support the hypothesis that procedural rationality and frequency of team meetings results in more agreement, a negative relationship is expected between these two team process variables and cognitive change dissimilarity.

In teams characterised by reflexivity and psychological safety it should be socially acceptable for individuals in teams maintain their own opinions. That is to say, one would expect that team processes would have a positive effect on zero change or tendency towards 1.0 (a complete match between individuals' pre- and post-discussion rankings), which suggests intransigence to change. In order to support the hypothesis, a positive relationship is expected between these two team process variables and cognitive change dissimilarity.

As established in Chapter 8 (section 8.3.4), and repeated in Table 9.6, the variance partition coefficient is 10%, which is attributable to between team differences. As in previous models in this chapter, slope coefficients are shown in columns #1 – #4 consistent with the order in which the predictors are listed.

9.4.4.1 Team Process Effects: Cognitive Change Dissimilarity

As shown in Table 9.6, only one of the four team processes is statistically significant, that is with frequency of meetings, which reduces the deviance by 8.58 (p = <0.01) to 144.41 and achieves a negative (expected) slope coefficient (β = -0.12, p <0.01). The between team variance is reduced to σ^2_u 0.00. This result, which conforms to prediction, suggests that in teams that meet frequently, people change their minds more, which by logical extension, means that there should be a tendency towards more agreement amongst a team. Conversely, for teams which meet less often, individuals are more likely to stick to their own opinions, not changing their original ranking and thereby tending towards divergence.

9.4.4.2 Additive Team Process Effects: Cognitive Change Dissimilarity

The next step takes forward the statistically significant predictors of cognitive change dissimilarity established in Chapter 8 (section 8.3.4), namely gender dissimilarity together with functional, educational and gender diversity, to test for additional main effects of, and interactions with, team processes. Step 4 in Table 9.6 below, shows the effects on cognitive change dissimilarity of adding frequency of team meetings, in turn, to gender dissimilarity and each of the demographic diversity predictors.

When the variable frequency of team meetings is added to the dissimilarity and diversity predictors separately, in all four cases (the first four entries for Step 4), there is a statistically significant reduction in the deviance (final column), together with a statistically significant slope coefficient (slope #2). Indeed, even when frequency of team meetings is added to the dissimilarity variable together with both of the diversity ones, it retains its statistically significant effect (slope #3). If one were to add frequency of team meetings to the three diversity predictors and the dissimilarity predictor in one equation (not shown in the Table), frequency of team meetings still maintains its statistically significant effect. This means that frequency of team meetings is a robust and reliable predictor of cognitive change dissimilarity.

The next step is to investigate whether the main effects of team processes on gender dissimilar individuals in diverse teams are due to cross-level interactions.

Table 9.6 Cognitive Change Dissimilarity and Team Processes Steps 1 – 5

Model Steps & Predictors	Intercept Coefficient β	(SEβ)	Slope #1 Coefficient β	(SEβ)	Slope #2 Coefficient β	(SEβ)	Slope #3 Coefficient β	(SEβ)	Slope #4 Coefficient β	(SEβ)	Level 2 Variance (Between Teams) σ^2_u	(SE)	Level 1 Variance (Between Individuals) σ^2_e	(SE)	Dev.	Δ*2L_li
Step 1: Variance Components: (Null)	0.75	(0.05)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.02	(0.02)	0.18	(0.04)	152.99	-
Step 2: Demographic Predictors (Established in Chapter 8):																
Gender Dissimilarity	0.76	(0.04)	0.42*	(0.17)	n/a	n/a	n/a	n/a	n/a	n/a	0.01	(0.00)	0.18	(0.02)	148.01	4.98*
Functional Diversity	0.76	(0.04)	0.50*	(0.21)	n/a	n/a	n/a	n/a	n/a	n/a	0.01	(0.00)	0.18	(0.02)	147.93	6.63*
Educational Diversity	0.76	(0.04)	0.91*	(0.43)	n/a	n/a	n/a	n/a	n/a	n/a	0.01	(0.00)	0.18	(0.04)	148.52	4.47*
Gender Diversity	0.76	(0.04)	0.62*	(0.23)	n/a	n/a	n/a	n/a	n/a	n/a	0.01	(0.00)	0.18	(0.02)	146.76	6.23*
Gender Dissim. + Functional Diversity	0.76	(0.04)	0.37*	(0.17)	0.44*	(0.20)	n/a	n/a	n/a	n/a	0.00	(0.00)	0.18	(0.02)	143.60	4.41*
Gender Dissim. + Educational Diversity	0.76	(0.04)	0.36*	(0.17)	0.75*	(0.38)	n/a	n/a	n/a	n/a	0.00	(0.01)	0.18	(0.02)	144.47	3.54*
Gender Dissim. + Gender Diversity	0.76	(0.04)	0.17	(0.26)	0.45	(0.34)	n/a	n/a	n/a	n/a	0.01	(0.01)	0.18	(0.02)	146.31	1.70
Gender Dissim. * Functional Diversity	0.76	(0.05)	0.40*	(0.19)	0.37	(0.26)	-0.59	(1.58)	n/a	n/a	0.00	(0.00)	0.18	(0.02)	143.46	0.14
Gender Dissim. * Educational Diversity	0.76	(0.05)	0.40*	(0.18)	0.66	(0.41)	-1.38	(2.09)	n/a	n/a	0.01	(0.01)	0.18	(0.02)	144.03	0.44
Gender Dissim. * Gender Diversity	0.79	(0.06)	0.35*	(0.36)	0.41	(0.35)	-1.21	(1.73)	n/a	n/a	0.01	(0.01)	0.18	(0.02)	145.82	0.49
Step 3: Team Process Predictors:																
Procedural Rationality	0.75	(0.05)	-0.01	(0.04)	n/a	n/a	n/a	n/a	n/a	n/a	0.02	(0.02)	0.18	(0.02)	152.95	0.04
Frequency of Meetings	0.75	(0.04)	-0.12**	(0.04)	n/a	n/a	n/a	n/a	n/a	n/a	0.00	(0.00)	0.18	(0.02)	144.41	8.58**
Reflexivity	0.75	(0.04)	0.05	(0.11)	n/a	n/a	n/a	n/a	n/a	n/a	0.02	(0.01)	0.18	(0.02)	152.80	0.19
Psychological Safety	0.75	(0.05)	-0.02	(0.07)	n/a	n/a	n/a	n/a	n/a	n/a	0.02	(0.01)	0.18	(0.02)	152.94	0.05
Step 4: Main Effects:																
Gender Dissim. + Frequency of Mtgs.	0.76	(0.04)	0.33*	(0.17)	-0.10**	(0.04)	n/a	n/a	n/a	n/a	0.00	(0.00)	0.17	(0.02)	140.53	7.48**
Funct. Diversity + Frequency of Mtgs.	0.76	(0.04)	0.40*	(0.20)	-0.10**	(0.04)	n/a	n/a	n/a	n/a	0.00	(0.00)	0.18	(0.02)	142.98	4.95*
Edu. Diversity + Frequency of Mtgs.	0.76	(0.04)	0.91*	(0.41)	-0.11**	(0.04)	n/a	n/a	n/a	n/a	0.01	(0.01)	0.18	(0.02)	140.90	7.62**
Gender Diversity + Frequency of Mtgs.	0.76	(0.04)	0.46*	(0.22)	-0.10**	(0.04)	n/a	n/a	n/a	n/a	0.00	(0.00)	0.17	(0.02)	140.22	6.54*
Gender Diss. + Funct. Div. + Mtgs.	0.75	(0.04)	0.30*	(0.16)	0.30	(0.20)	-0.09*	(0.04)	n/a	n/a	0.00	(0.00)	0.17	(0.02)	138.31	5.29*
Gender Diss. + Edu. Div. + Mtgs.	0.75	(0.04)	0.29*	(0.17)	0.59	(0.37)	-0.09*	(0.04)	n/a	n/a	0.00	(0.00)	0.17	(0.02)	138.04	6.43*
Gender Diss. + Gender Div. + Mtgs.	0.76	(0.04)	0.17	(0.25)	0.29	(0.33)	-0.10**	(0.04)	n/a	n/a	0.00	(0.00)	0.17	(0.02)	139.78	6.04*
Step 5: Interactions:																
Gender Dissim. * Frequency of Mtgs.	0.77	(0.04)	0.41*	(0.18)	-0.09*	(0.04)	0.24	(0.20)	n/a	n/a	0.00	(0.00)	0.17	(0.02)	139.03	1.50
Funct. Diversity * Frequency of Mtgs.	0.76	(0.04)	0.40*	(0.20)	-0.10*	(0.05)	0.34*	(0.15)	n/a	n/a	0.00	(0.00)	0.18	(0.02)	142.98	0.00
Edu. Diversity * Frequency of Mtgs.	0.76	(0.04)	0.78*	(0.38)	-0.09*	(0.04)	0.43	(0.35)	n/a	n/a	0.00	(0.01)	0.17	(0.02)	139.37	1.53
Gender Diversity * Frequency of Mtgs.	0.77	(0.04)	0.54*	(0.23)	-0.08*	(0.04)	0.33	(0.25)	n/a	n/a	0.00	(0.00)	0.17	(0.02)	138.48	1.74
Gender Diss. * Funct. Div. * Mtgs.	0.75	(0.04	0.33*	(0.17)	0.46*	(0.24)	0.09*	(0.04)	1.33	(1.08)	0.00	(0.00)	0.17	(0.02)	136.82	1.49
Gender Diss. * Edu. Div. * Mtgs.	0.75	(0.04)	0.28	(0.17)	0.58	(0.38)	-0.10**	(0.04)	-0.51	(2.07)	0.00	(0.00)	0.17	(0.02)	137.98	0.06
Gender Diss. * Gender Div. * Mtgs.	0.75	(0.04)	0.22	(0.26)	0.27	(0.33)	-0.12**	(0.04)	0.97	(1.23)	0.00	(0.00)	0.17	(0.02)	139.16	0.62

* significant at p <0.05; and ** significant at p <0.01, one-tailed. Steps 2: d/f 2,117 to 2,127; Step 3: d/f, 1,118 to 1,128; Steps 4 & 5: d/f 2,117 to 2,127 and 3,116 to 3,126 depending on missing values. Δ*2L_li for 3 way interactions at Step 5 are compared against main effects at Step 4 as a proxy for full formula of 3 main effects and 3 two-way interactions.

9.4.4.3 Interaction Effects: Cognitive Change Dissimilarity

The lower section of Table 9.6 shows that none of the eight possible interaction terms is statistically significant.

9.4.4.4 Summary: Cognitive Change Dissimilarity

In summary, the results reported in this section provide a little support for the hypothesis that team processes predict individual cognitive variation. Specifically, frequency of team meetings reduces individual variation in personal opinions However, contrary to expectation, frequency of team meetings does not exert influence on dissimilar individuals in diverse teams (i.e. no interaction effects).

9.4.5 Team Processes and Cognitive Dissimilarity Post-Discussion

The fourth cognitive individual response variable to be investigated is cognitive dissimilarity post-discussion. As established in Chapter 8 (section 8.3.5), the deviance is 582.77 and the variance partition coefficient (VPC) is a substantial 42%, which is attributable to between team differences.

9.4.5.1 Team Process Effects: Cognitive Dissimilarity Post-Discussion

Table 9.7 reports the addition of team processes as fixed effects at level 2. Two statistically significant results emerge. The first, which supports the hypothesis, is for frequency of team meetings, with the reduction in the deviance being 3.83 ($p<0.05$) and the between team variance being reduced to σ^2_u 4.60. The slope coefficient is negative and statistically significant (β = -0.92, p <0.05), which conforms to prediction. The conclusion with regard to cognitive dissimilarity post-discussion, is that frequency of team meetings reduces the level of disagreement.

The second statistically significant result, however, is in the opposite direction to that expected, and is for procedural rationality which reduces the deviance by 6.35 (p <0.01). The between team variance is reduced to σ^2_u 4.07, but the slope coefficient is positive (β = 1.13, p <0.05), which is contrary to prediction. It means that in teams that tend to systematically review information for decision-making, individuals will take distance values on ranking post-discussion 1.13 above the mean. In other words, the level of post-discussion disagreement is increased in teams which use analytical decision making strategies. The coefficients for the other two processes under consideration are not statistically significant.

The next step is to investigate whether frequency of team meetings and procedural rationality explain variance in cognitive dissimilarity post-discussion additional to that accounted for by the established demographic predictors, namely tenure dissimilarity & educational diversity as established in Chapter 8.

9.4.5.2 Additive Team Process Effects: Cognitive Dissimilarity Post-Discussion

The entry for Step 4 in Table 9.7 follows the convention established in previous tables of listing procedural rationality first, but as the effect at Step 3 was in opposition to the hypothesis, this discussion will address frequency of team meetings first.

Examination of the findings at Step 4 (Table 9.7) shows that frequency of team meetings explains additional variance over and above tenure dissimilarity. The deviance is reduced by 4.42 (p <0.05), the level 2 variance is reduced to σ^2_u 4.99, and the frequency of meetings slope coefficient is β -0.81 (p <0.05). With respect to educational diversity, frequency of team meetings does not explain additional variance as the deviance is reduced by only 2.18, which is not statistically significant. In the absence of a statistically significant improvement in model fit the effect for the frequency of meetings slope coefficient (β -1.05, p <0.05) is disregarded.

Contrary to the hypothesis, procedural rationality also explains variance over and above that explained by tenure dissimilarity, as shown by the reduction in the deviance in the final column (4.05, p <0.05), and by the reduction in the level 2 variance (σ^2_u 4.72), and by the relevant slope coefficient (slope #2) β 0.97 (p <0.05). A similar result emerges for educationally diverse teams, which reduces the deviance by 3.25 (p <0.05) and reduces the level 2 variance. The addition of procedural rationality achieves a positive slope coefficient of β = 1.10 (p <0.01). This means that procedural rationality explains variance in cognitive dissimilarity post-discussion in addition to that explained by educational diversity.

Table 9.7 Cognitive Dissimilarity Post-discussion and Team Processes Steps 1 – 5

Model Steps & Predictors	Intercept Coefficient β	(SEβ)	Slope #1 Coefficient β	(SEβ)	Slope #2 Coefficient β	(SEβ)	Slope #3 Coefficient β	(SEβ)	Slope #4 Coefficient β	(SEβ)	Level 2 Variance (Between Teams) σ^2_u	(SE)	Level 1 Variance (Between Individuals) σ^2_e	(SE)	Dev.	Δ*2L_li
Step 1: Variance Components: (Null)	4.00	(0.52)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	5.55	(1.84)	3.47	(0.47)	582.77	n/a
Step 2: Demographic Predictors (Established in Chapter 8):																
Tenure Dissimilarity	4.01	(0.53)	-0.01*	(0.00)	n/a	n/a	n/a	n/a	n/a	n/a	5.35	(1.85)	3.27	(0.45)	577.33	5.44*
Educational Diversity	4.08	(0.49)	8.05*	(4.65)	n/a	n/a	n/a	n/a	n/a	n/a	4.82	(1.59)	3.47	(0.47)	579.47	3.30*
Tenure Dissim. + Educational Diversity	4.07	(0.50)	-0.01*	(0.00)	7.25	(4.79)	n/a	n/a	n/a	n/a	4.17	(1.62)	3.47	(0.47)	575.17	2.16
Tenure Dissim. * Educational Diversity	4.07	(0.50)	-0.01	(0.00)	7.25	(4.79)	0.00	(0.00)	n/a	n/a	5.16	(1.70)	3.28	(0.45)	575.17	0.00
Step 3: Team Process Predictors:																
Procedural Rationality	3.89	(0.46)	1.13*	(0.42)	n/a	n/a	n/a	n/a	n/a	n/a	4.07	(1.40)	3.46	(0.47)	576.42	6.35**
Frequency of Meetings	4.10	(0.48)	-0.92*	(0.45)	n/a	n/a	n/a	n/a	n/a	n/a	4.60	(1.56)	3.46	(0.47)	578.94	3.83*
Reflexivity	4.00	(0.51)	1.01	(1.21)	n/a	n/a	n/a	n/a	n/a	n/a	5.36	(1.78)	3.47	(0.47)	582.08	0.69
Psychological Safety	4.01	(0.52)	-0.33	(0.80	n/a	n/a	n/a	n/a	n/a	n/a	5.51	(1.82)	3.37	(0.47)	582.59	0.18
Step 4: Main Effects:																
Tenure Dissim. + Procedural Rationality	3.89	(0.47)	-0.01*	(0.00)	0.97*	(0.46)	n/a	n/a	n/a	n/a	4.72	(1.58)	3.28	(0.45)	573.28	4.05*
Educational Diversity + Proced. Rat.	3.96	(0.43)	7.56*	(4.04)	1.10**	(0.39)	n/a	n/a	n/a	n/a	3.43	(1.21)	3.47	(0.47)	573.17	3.25*
Tenure Dissim. + Edu. Div. + Pro. Rat.	3.95	(0.44)	-0.00	(0.00)	6.84	(4.21)	1.07*	(0.41)	n/a	n/a	3.80	(1.30)	3.28	(0.45)	569.17	6.00*
Tenure Dissim. + Mtgs.	4.09	(0.50)	-0.01*	(0.00)	-0.81*	(0.47)	n/a	n/a	n/a	n/a	4.99	(1.65)	3.28	(0.45)	574.52	4.42*
Educational Diversity + Mtgs.	4.18	(0.44)	8.35	(5.31)	-1.05*	(0.43)	n/a	n/a	n/a	n/a	4.15	(1.42)	3.47	(0.47)	575.76	2.18
Tenure Dissim. + Edu. Diversity + Mtgs.	4.13	(0.48)	-0.00	(0.00)	5.83	(4.65)	-0.70	(0.46)	n/a	n/a	4.50	(1.54)	3.28	(0.45)	573.01	2.16
Step 5: Interactions:																
Tenure Dissim. * Procedural Rationality	3.99	(0.48)	-0.00	(0.00)	0.93*	(0.46)	-0.00	(0.00)	n/a	n/a	4.66	(1.56)	3.21	(0.44)	570.93	2.35
Educational Diversity * Proced. Rat.	3.98	(0.42)	7.25*	(3.97)	1.10**	(0.39)	-3.18	(3.17)	n/a	n/a	3.27	(1.17)	3.46	(0.48)	572.18	0.99
Tenure Dissim. * Edu. Div. * Pro. Rat.	3.89	(0.45)	-0.01	(0.00)	7.39	(4.26)	1.15*	(0.41)	0.04	(0.03)	3.87	(1.32)	3.22	(0.44)	567.46	1.71
Tenure Dissim. * Mtgs.	4.09	(0.50)	-0.01*	(0.00)	-0.69	(0.48)	-0.00	(0.00)	n/a	n/a	5.08	(1.64)	3.21	(0.44)	572.43	2.11
Educational Diversity * Mtgs.	4.18	(0.44)	8.35	(5.31)	-1.05*'	(0.43)	-3.96	(3.94)	n/a	n/a	3.96	(1.37)	3.47	(0.47)	575.76	2.18
Tenure Dissim. * Edu. Diversity * Mtgs.	4.14	(0.48)	-0.00	(0.00)	5.39	(4.78)	-0.67	(0.47)	0.01	(0.03)	4.53	(1.53)	3.29	(0.45)	572.88	0.13

* significant at p <0.05; and ** significant at p <0.01, one-tailed. Steps 2: d/f 2,117 to 2,127; Step 3: d/f, 1,118 to 1,128; Steps 4 & 5: d/f 2,117 to 2,127 and 3,116 to 3,126 depending on missing values. Δ*2L_li for 3 way interactions at Step 5 are compared against main effects at Step 4 as a proxy for full formula of 3 main effects and 3 two-way interactions.

When the three predictors are added together, procedural rationality maintains its statistically significant effect. The next step is to investigate whether these observations are due to cross-level interactions.

9.4.5.3 Interaction Effects: Cognitive Dissimilarity Post-Discussion

The computation of interaction terms, as reported in the lower section of Table 9.7, shows that none of the potential interaction terms is statistically significant.

9.4.5.4 Summary: Team Processes and Cognitive Dissimilarity Post-Discussion

In summary, the results for this section provide contradictory support for H6 in that one of the five possible relationships conforms to prediction, and one does not. Specifically, frequency of meetings, generally speaking, reduces the levels of disagreement amongst team members' rankings post-discussion as predicted. Procedural rationality on the other hand, which is also expected to reduce disagreement, has the opposite effect in that it significantly exacerbates disagreement.

Reflexivity and psychological safety, which are widely held to engender constructive yet conflicting opinions, exhibit no influence at all.

9.4.6 Team Processes and Cognitive Cohesion Dissimilarity Post-Discussion

The final cognitive individual response variable to be investigated is cognitive cohesion dissimilarity post-discussion. As established in Chapter 8 (section 8.3.6), the variance in cognitive cohesion dissimilarity post-discussion is almost all at the individual level. Therefore, modelling of this response variable with team process variables is inappropriate.

9.4.7 Conclusions Regarding Hypothesis 6

Hypothesis 6 posed two similar but discrete questions. First, do team processes predict individual cognitive variation? Second, do team processes affect cognitive variation amongst the most dissimilar individuals in demographically diverse teams? That is to say, is there an interaction with demographic dissimilarity and diversity factors? The analyses for this dual faceted question employed multi-level modelling. Twenty core relationships were tested (five cognitive dissimilarity variables x four team process variables). Four statistically significant results emerged in support of the hypothesis, and another was contrary to hypothesis. This number is about what one would expect occur due to chance alone. All five results concerned the two processes hypothesised to reduce cognitive variation, procedural rationality and frequency of team meetings. There were no statistically significant results concerning reflexivity and psychological safety.

Specifically, and in support of H6, it was found that procedural rationality positively influences proximity to the team consensus pre-discussion. Frequency of team meetings reduces dissimilarity in rankings pre- and post-discussion and encourages people to change their minds more in the interests of agreement.

A further 32 relationships were tested with the team process variables in combination with demographic dissimilarity or diversity variables either as main effects (16 tests) or interaction terms (16 tests). No interactions were statistically significant, which means that the second strand of the hypothesis concerning the influence of team processes on dissimilar individuals in teams, is not supported.

The findings presented in this chapter provide an insight into TMT processes and their effect upon dissimilar individuals in diverse teams. Until the advent of multi-level modelling, the contextual approach to TMT research (i.e. individuals in the context of teams) has been unexploited. However, not all the findings, interesting as they are, conform to prediction. For example, procedural rationality is a positive predictor of cognitive dissimilarity post-discussion. This result was definitely not expected, yet one can see how a systematic analysis of the information by the team during discussion could convince an individual that their candidate selection was correct even if the team came to a different consensus. Furthermore, reflexivity and psychological safety have no effects at all, when it was expected that they would increase cognitive variation. As was explained in Chapter 5, predicting the direction of the association regarding team processes is challenging. First, team processes are the traditional 'black box' (Lawrence, 1997), which euphemistically means no-one knows what goes on inside it, because team processes are rarely studied at this level in organisations. Some of the counterintuitive findings for this hypothesis highlight the complexity and subtlety of what goes on in the black box.

Second, when team processes are investigated, they are usually studied by means of either retrospective recollections by selected TMT members (e.g. Dean & Sharfman, 1996), or retrospective self report perceptions (e.g. West & Anderson, 1996). Both methods are subject to biased recall of information. By way of contrast, this study used independent observers to measure team processes, thereby providing a more rigorous assessment of their effect.

As the three hypotheses reported on in this chapter were clustered together in the guiding conceptual model on page 64, the next section will provide an overview of the similarities and differences between the three sets of findings.

9.5 Conclusions re. Hypotheses 4-6

Three key findings have emerged from the results presented in this chapter. First, partial support was found for H4, which predicted a negative relationship between demographic diversity and team processes. Specifically, it was found that, age diversity is a negative predictor of reflexivity, whilst gender diversity is a negative predictor of psychological safety and procedural rationality.

The second key finding to emerge concerned the relationship between team processes and cognitive diversity. Hypothesis 5, had two strands, both of which received limited support. The first, which predicted a negative relationship between team processes supposed to engender agreement and less cognitive variation, was partially supported in that frequent team meetings reduce cognitive diversity post-discussion. The second, which predicted a positive relationship between team processes which are supposed to foster constructive controversy (Tjosvold, 1996), and make it acceptable to change one's opinion, is partially supported in that psychological safety predicts cognitive change diversity.

Third, as part of the aim of this thesis is to examine what happens to individuals within the context of teams, analyses showed that team processes do predict individual cognitive variation. Specifically, procedural rationality as expected, reduces individual cognitive variation, manifested by more individuals tending towards proximity to the team consensus pre-discussion.

Frequency of team meetings, as expected, reduces cognitive variation pre- and post-discussion, and encourages individuals to change their minds. The meaning of these findings, will be discussed in Chapter 11, together with some implications and contributions of this research to the TMT literature.

The next chapter will present the results and analysis for the final 3 hypotheses, concerning the relationships between diversity (demographic and cognitive), team processes and decision belief.

CHAPTER 10

Findings: Hypotheses 7 - 9

10.0 Overview

This chapter presents the findings for hypotheses 7 to 9, dealing with the relationships between diversity (demographic and cognitive), team processes, and decision belief all at the team level (as shown in the guiding conceptual model on p 76).

10.1 Introduction

Hypothesis 7 (H7) predicts a positive relationship between two team processes and decision belief. That is to say, teams which employ procedural rationality, and teams which tend to meet more often, will report greater satisfaction, confidence and effectiveness with their consensus ranking and their decision process. Reflexive teams on the other hand, and teams in which it is psychologically safe to express dissenting views, will report less satisfaction, confidence and effectiveness.

Hypothesis 8 (H8) predicts a negative relationship between demographic diversity and decision belief. It proposes that those teams characterized by greater variation in age, functional background, educational attainment, gender and tenure, will report less satisfaction, confidence and effectiveness with their consensus ranking and their decision process.

Finally, hypothesis 9 (H9) predicts a negative relationship between cognitive diversity and decision belief. In other words, it is argued that teams that show the greatest variation (i.e. least agreement) in how they rank the candidates will report less satisfaction, confidence and effectiveness with their consensus ranking and their decision process. For each of the hypotheses, statistical considerations are presented first, followed by zero-order correlations and multiple regression analyses. Within each relevant section, the two pre-discussion measures of cognitive variation are reported first, followed by cognitive change diversity, and the two post-discussion measures.

10.2 The Relationship Between Team Processes & Decision Belief (H7)

10.2.1 Statistical Considerations

Team process variables were normally distributed as earlier reported. There were three decision belief measures as reported in Chapter 6. These were perceived satisfaction with the decision-making process, confidence that the team had made the right decision (i.e. selected the most appropriate candidate for appointment) and perceived effectiveness in decision-making. The underlying distributions for all three decision belief variables were found to be sufficiently normally distributed with acceptable levels of skew and kurtosis to allow the use of parametric statistics.

10.2.2 Zero-order Correlations

Initial evidence supporting the hypothesis comes from the zero-order correlations reported in orange box of Table 9.1 (p 182). Only three of a possible 12 relationships are statistically significant, and they are all in the expected direction.

The first finding, is that frequency of team meetings and confidence are positively correlated (r = .28 p <0.10). This finding means that teams which meet more often are more confident in the decisions they make. The second finding, is that between frequency of team meetings and perceived effectiveness (r = .34 p <0.05). This finding means that teams which meet more often believe themselves to be more effective as decision-makers. As explained in Chapter 9, it is smaller teams which meet more frequently, so extrapolating on the finding, one might suggest that it is smaller teams that are more confident and believe themselves to be more effective. This issue is addressed later.

The third finding is that between reflexivity and confidence (r = -.40, p <0.05). This result, explaining 16% of the variance (neither variable is related to team size, hence no regression analysis is necessary), means that reflexive teams, that is those that air their problems in the interests of self improvement, are the least confident in their ultimate decision.

In order to investigate further the possible effect of team size in relation to frequency of team meetings and decision belief, the next section will report on regression analyses controlling for team size.

10.2.3 Multiple Regression Analyses for H7

The results of the multiple regression analyses for H7 are shown in Table 10.1. Team size was entered first followed by frequency of team meetings. In both cases, once team size is controlled for, the predictor is no longer statistically significant.

Table 10.1 Multiple Regression Analyses for the Effects of Team Processes on Decision Belief

Outcome Variable	Predictor Variable	β	$R^2\Delta$	F	df
Confidence	1. Team Size	-.04	.04	.85	1,21
	2. Frequency of Team Meetings	.26	.04	.87	2,20
Perceived Effectiveness	1. Team Size	-.02	.05	1.07	1,21
	2. Frequency of Team Meetings	.33	.12	1.31	2,20

N = 23; † = p <0.10; * = p <0.05; ** = p <0.01. One-tailed tests of significance levels for standardised β coefficients in final equation. $R^2\Delta$ refers to discrete steps.

In summary, these findings mean that neither team size nor frequency of meetings predicts confidence nor effectiveness. The next section reviews the implications of the findings in terms of H7.

10.2.4 Conclusions Regarding Hypothesis 7

As with earlier hypotheses, evidence in support of H7 is sparse. That there are no statistically significant effects for psychological safety nor procedural rationality suggests that these processes are not important for decision belief. On the other hand, the relatively small dataset may preclude the observation of effects in this case. For the two effects involving frequency of meetings which emerged in the preliminary analysis, with confidence and effectiveness respectively, the predictor is no longer statistically significant when subjected to multiple regression controlling for team size. The only other statistically significant finding concerns the effect of reflexivity and confidence. It is of interest to note that if one were to conduct a regression analysis to control for team size, the observed relationship is, once again, no longer significant. As discussed in Chapter 7, the loss of statistically significant relationships after multiple regression analysis must largely be due to the increase in degrees of freedom coupled with the very small effect size at the margins of statistical significance.

The key finding to emerge regarding H7, is that, as predicted, the association between reflexivity and decision belief is negative. It is argued that more reflexive teams do (and should) embrace conflict regardless of its impact upon morale, which will probably be negative (West, 1996). In support of this argument, the finding of this study is that reflexive teams experience a lack of confidence in the decision they made for candidate selection. West (1996) has consistently argued that reflexivity is particularly good for TMTs (West et al., 1997) and that the concept is symbiotic with psychological safety which is supposed to reduce the conflict inherent in reflexivity. There is no evidence in this study to suggest that reflexivity, and the conflict that it supposedly generates, is beneficial for TMTs, nor has there been a consistent link with psychological safety in any of the hypotheses tested. Whilst charitably one might argue (based on West 1996, and West et al. 1997) that lack of confidence is good, because it will lead to the introduction of improved processes, or a review of the decision so that a better one can be made, this has yet to be tested. Moreover, in teams in other organisational settings (i.e. not TMTs) it has been found that whilst teams characterised by conflict are the best generators of ideas for improvement (i.e. innovation), they are typically the worst implementers of the ideas generated (Agrell & Gustafson, 1996). Therefore, one should be cautious in advocating low levels of confidence arising from reflexivity as a beneficial function in TMTs.

The next section presents the analyses testing the relationships between demographic diversity and decision belief.

10.3 The Relationship Between Demographic Diversity and Decision Belief (H8)

H8 focuses on the team level, predicting a negative relationship between demographic diversity and decision belief.

10.3.1 Statistical Considerations

All five demographic diversity and three decision belief variables used for testing this hypothesis are normally distributed as reported in the previous section and in Chapter 6.

10.3.2 Zero-order Correlations

Initial supporting evidence for H8 comes from the bivariate correlations reported in the light yellow box in Table 9.1 (on p 182). Of the possible 15 relationships, six are statistically significantly associated in the expected direction, whilst two are statistically significantly associated in the opposing direction.

With regard to satisfaction with the decision process, four of the five demographic diversity variables are associated in the expected (negative) direction, of which one is statistically significant. Educationally diverse teams report less satisfaction with the decision process (r = -.41, p <0.05). This finding means that teams of varied educational backgrounds report less satisfaction than educationally similar teams.

Contrary to expectation, tenure diversity is positively related to decision satisfaction (r = .33, p <0.10), suggesting that teams characterised by greater variation in tenure experience greater satisfaction than teams made up of similar tenured members.

As educational and tenure diversity are both associated with team size, regression analyses controlling for team size will be conducted in the next section.

With respect to confidence in having made the right decision, relationships with functional diversity, educational diversity and gender diversity are substantial, and negative, as predicted (r = -.38, p <0.05; r = -.39, p <0.05; and r = -.33, p <0.10 respectively). This suggests that teams characterized by greater variation in functional background and educational attainment, and mixed gender teams, do not have as much confidence in their decision as more homogenous teams on these attributes. Opposing these findings is a further result concerning confidence. Again, tenure diversity is statistically significantly associated but in the wrong (positive) direction (r = .53, p <0.01), suggesting that teams characterised by greater variation in tenure have more confidence in the decision. With the exception of gender diversity, the other variables are positively associated with team size, so multiple regression analyses which follow in the next section will control for this potentially confounding factor. Gender diversity explains 11% of the variance in decision confidence.

With respect to perceived effectiveness, the expected negative relationship was observed with functional background diversity (r = -.34, p <0.10) and educational attainment diversity (r = -.39, p <0.05). This means that teams perceive themselves to be less effective when their members come from a variety of functional and educational backgrounds

10.3.3 Multiple Regression Analyses for H8

Bivariate correlations, as shown in Table 9.1 (p 182), show that some of the independent variables were related to team size. Therefore, team size was entered first in the regression equation. This was followed by entering those predictor variables found to be statistically significant related to the dependent variable in the bivariate analysis. As has been observed elsewhere in this thesis, the ratio of cases to the dependent variable is such that multiple predictors must be added in separate analyses with team size.

The regression analysis for satisfaction shows that after controlling for team size, educational diversity is no longer statistically significant.

The next regression analysis for satisfaction and tenure diversity, shows that team size by itself is not a statistically significant predictor, but that once team size is controlled for, tenure diversity is the primary statistically significant predictor (β = .50, p<0.01) explaining 20% of the variance in satisfaction. This result suggests that tenure diverse teams enjoy a shared interpretation of process which is satisfactory. This is opposite to prediction, and Pfeffer's (1983) argument that tenure diverse teams would be less satisfied.

The results for the three regression analyses for decision confidence are shown in Table 10.2. Team size was entered first. Once team size is controlled for, educational diversity is the only remaining statistically significant negative predictor of decision confidence (β = .77, p<0.05) explaining 13% of the variance. Functional background diversity is no longer statistically significant. It would seem that the effect of functional background diversity is a function of team size.

Again, tenure diversity remains strong in opposition to the hypothesis (β = .77, p<0.01) and once team size is controlled for, explains 49% of the variance.

With regard to the final expected negative relationships, educational attainment diversity and functional background diversity were regressed onto effectiveness. The variance explained by these variables is not statistically significant.

Although confidence and satisfaction were correlated highly (r = .53, p<0.01), the intercorrelation was not high enough to automatically suggest they were tapping the same construct. In order to ascertain whether the findings were distinct, a multiple regression analysis was conducted in which satisfaction was the dependent variable, confidence was controlled for in the first step and tenure diversity was entered in the second step. If tenure diversity had achieved a statistically significant beta weight, one could assume that tenure diversity was a predictor of satisfaction separate from the confidence effect. However, tenure diversity did not achieve significance (β = .06). An identical analysis was conducted, but this time using confidence as the dependent variable and controlling for satisfaction. In this case, tenure diversity achieved a statistically significant result over and above satisfaction (β = .40, p<0.05). This means that the finding with satisfaction is spurious, whereas the finding for tenure diversity as a predictor of confidence is robust, albeit against the hypothesis.

Table 10.2 Multiple Regression Analyses for the Effects of Demographic Diversity on Decision Belief

Outcome Variable	Predictor Variable	β	$R^2 \Delta$	F	df
Satisfaction	1. Team Size	-.08	.03	.55	1,21
	2. Educational Diversity	-.20	.03	.63	2,20
	1. Team Size	-.38	.03	.55	1,21
	2. Tenure Diversity	.50**	.20	2.90*	2,20
Confidence	1. Team Size	-.54**	.04	0.85	1,21
	2. Tenure Diversity	.77**	.49	10.72**	2,20
	1. Team Size	-.04	.04	.85	1,21
	2. Educational Diversity	-.40*	.11	2.09†	2,20
	1. Team Size	-.04	.04	.85	1,21
	2. Functional Diversity	-.36	.11	1.70	2,20
Effectiveness	1. Team Size	-.09	.05	1.07	1,21
	2. Functional Background Diversity	-.30	.07	1.40	2,20
	1. Team Size	-.08	.05	1.07	1,21
	2. Educational Attainment Diversity	-.36	.11	1.86	2,20

N = 23; † = p <0.10; * = p <0.05; ** = p <0.01. One-tailed tests of significance levels for standardised β coefficients in final equation. $R^2 \Delta$ refers to discrete steps. No entry for gender diversity and confidence as no relationship with team size.

10.3.4 Conclusions Regarding Hypothesis 8

Once again, as with previous hypotheses, results supporting H8 are meagre indeed. Of a possible 15 relationships, only two are statistically significant after being subjected to regression analysis, and a further one, also statistically significant, is in opposition to the hypothesis.

The key findings to emerge form H8 are that, as predicted, educationally diverse teams, and gender diverse teams are less confident in their ultimate decision. However, opposing H8, tenure diversity was found to be a positive predictor of confidence in the decision. In trying to explain this last result, one's own experience in TMTs suggests a possibility. That is, individuals who have been on the team a comparatively short time (i.e. less tenure) will follow the lead of those who have been on the team a comparatively long time (i.e. more tenure). These long-standing individuals, because of their tenure and experience with the TMT decision making process, are likely to have more confidence in their decisions and be content with the status quo (i.e. familiarity with the process). Collectively, then, tenure diversity predicts confidence in the decision. There is some support for this supposition from the literature. Variation in team tenure is reportedly responsible for power hierarchies and communication patterns in teams (Clark et al. 1997), although this has generally been about groups of longer serving managers vs. groups of shorter tenure managers, not diversity within the team.

The next section presents the analysis and findings for the final set of expected relationships, those between cognitive diversity and decision belief.

10.4 The Relationship Between Cognitive Diversity and Decision Belief (H9)

H9 predicts a negative relationship between cognitive diversity and decision belief.

10.4.1 Statistical Considerations

The underlying distributions of the five cognitive diversity variables and three decision belief variables have already been addressed in earlier sections and in Chapter 6, and allow the use of parametric statistical methods.

10.4.2 Zero-order Correlations

Initial supporting evidence for H9 comes from the bivariate correlations reported in the purple box in Table 9.1 (on p 182). Of the 15 possible relationships three were statistically significant and associated in the expected negative direction, and another was statistically significant but counter to expectations.

Cognitive change diversity was negatively associated with satisfaction (r = -.28, p <0.10), which means that the more variation in cognitive change there is within a team, the less satisfaction is reported. Cognitive diversity post discussion was negatively associated with confidence (r = -.46, p <0.01) and perceived effectiveness (r = -.33, p <0.10). These results suggest that teams with more diversity (i.e. least agreement) post-discussion are less confident in their decision, and believe they are less effective.

Contrary to expectation, cognitive cohesion diversity pre-discussion was positively associated with confidence (r = .30, p <0.10). This result suggests that the more variation there is around the team consensus pre-discussion, the more confident the team in their decision.

10.4.3 Multiple Regression Analyses for H9

Regression analysis for H9 confirms the finding for cognitive diversity post-discussion and lack of confidence in the team decision as reported in Table 10.3. After controlling for team size, cognitive diversity post-discussion explains 19% of the variance in confidence (β = -.57, p <0.05).

With regard to satisfaction, the predictor cognitive change diversity ceases to be statistically significant (β = -.37), with regard to confidence, the predictor cognitive cohesion diversity pre-discussion ceases to be statistically significant (β = .30), and similarly with regard to effectiveness, the predictor cognitive diversity post-discussion loses power (β = -.32) and ceases to explain any statistically significant variance.

Table 10.3 Multiple Regression Analyses for the Effects of Cognitive Diversity on Decision Belief

Outcome Variable	Predictor Variable	β	$R^2\Delta$	F	*df*
Satisfaction	1. Team Size	-.28	.03	.55	1,21
	2. Cognitive Change Diversity	-.37	.13	1.77	2,20
Confidence	1. Team Size	.18	.04	0.85	1,21
	2. Cognitive Diversity Post-Discussion	-.57*	.19	4.87	2,20
	1. Team Size	-.20	.04	.85	1,21
	2. Cognitive Cohesion Diversity Pre-Discussion	.30	.09	1.48	2,20
Effectiveness	1. Team Size	-.01	.05	1.07	1,21
	2. Cognitive Diversity Post-Discussion	-.32	.11	1.20	2,20

$N = 23$; * $p = <0.05$; ** $p = <0.01$. One-tailed tests of significance levels for standardised β coefficients in final equation. $R^2\Delta$ refers to discrete steps.

10.4.4 Conclusions Regarding Hypothesis 9

H9 finds paltry support in as much as, 15 relationships were investigated but only one, the finding that cognitive diversity post-discussion is a negative predictor of confidence, continues to be statistically significant after subjected to stringent regression analyses. Teams that experience the greatest variation in their rankings after discussion, lack confidence in their decision.

10.5 Conclusions re Hypotheses 7 – 9

Three key findings emerge from this chapter. The first, concerning H7, is that teams which are more reflexive report less confidence in their decision than less-reflexive teams.

The second key finding concerns H8, which predicted a negative relationship between demographic diversity and decision belief. This received partial support in that educationally diverse teams and gender diverse are less confident in their decision. Opposing the hypothesis, tenure diversity positively predicts confidence in the decision.

The third key finding is that H9, which predicted a negative relationship between cognitive diversity and decision belief, is barely supported by this study. Decision confidence is negatively associated with cognitive diversity post-discussion, meaning that teams that show the greatest variation in their rankings after discussion, have a lack of confidence in the decision.

The overriding finding, however, is that the observed effects are patchy and weak. The interpretation of these findings, together with those from Chapters 7 - 9 will be discussed in the concluding chapter.

CHAPTER 11

Discussion

11.0 Overview

This chapter evaluates and explains the findings in relation to the hypotheses generated in earlier chapters. Some of the propositions found support in the empirical study. The chapter concludes by summarising the findings in terms of what they tell about top management team decision-making. Limitations and potential criticisms of the research study's capacity to measure the constructs under discussion are addressed. The research framework model first proposed in Chapter 4 is revisited in light of the findings.

11.1 Introduction

This research sought to investigate the relationship between variation (demographic and cognitive) at the individual and team levels, team processes and decision quality. 'Upper echelons' theory asserts that managerial decision making is significantly affected by the viewpoints, values, beliefs and attitudes of TMTs, which in turn, are directly related to the immutable demographic attributes of team members (Hambrick & Mason, 1984). In Chapter 2 it was noted that there is as yet limited support for 'upper echelons' theory owing to disparities between studies with regard to the variables and the level (individual dissimilarity or team diversity) at which they are investigated. Moreover, social interaction processes are thought to moderate or mediate the relationship between demographic variation and outcomes such as decision belief. In Chapter 3 it was argued that procedural rationality and frequency of team meetings should lead to more consensus (less cognitive diversity), whilst reflexivity and participative safety should provide a healthy climate for constructive disagreement (more cognitive diversity). In Chapter 4, the relationships between the four groups of variables studied were crystallised into nine hypotheses, and graphically presented in a descriptive framework to guide the empirical research. This chapter is organised into five main sections according to the model presented in figure 4.1 (p 76).

Section one speaks to the fundamental premise underpinning 'upper echelons' theory (that there is a relationship between demographic and cognitive variation) and discusses the findings concerning Hypotheses 1 – 3. Section two interprets the findings of the study in terms of demographic diversity and team processes (Hypothesis 4). Section three reports on the findings concerning cognitive variation and team processes as posited in Hypotheses 5 and 6. Section four discusses and evaluates the findings in respect of team processes and decision belief (Hypothesis 7). Section five interprets the results of the study which relate to diversity (demographic and cognitive) and decision belief (Hypotheses 8 and 9). The chapter concludes by summarising the overall view that the results provide.

11.2 Demographic and cognitive variation (Hypotheses 1-3)

Three hypotheses focused on the relationship between demographic and cognitive variation. The first was at the individual level (130 executives), the second was at the team level (23 teams) and the third transcended both levels (130 executives in 23 teams).

Hypothesis 1 was that demographic dissimilarity would be positively associated with cognitive dissimilarity. Only three statistically significant results from 25 relationships tested at the $p<0.05$ level were observed in support of Hypothesis 1 (12%), basically one more than one could reasonably expect to occur by chance alone.

Specifically it was found that educational dissimilarity predicted cognitive dissimilarity post-discussion. That is to say, the most educationally dissimilar individual is more likely to hold a radical opinion to his peers post-discussion. Gender dissimilarity predicted zero cognitive change and age dissimilarity was a negative predictor of pre-discussion proximity to the team consensus.

Fundamental to 'upper echelons' is the concept of multiple idea generation, that is to say, demographic dissimilarity will manifest itself in a variety of different ideas that need to be melded into a team collective decision. This concept of multiple idea generation, which is akin to cognitive dissimilarity pre-discussion in this study (that is to say, the distances between pre-discussion candidate rankings of individuals), found no support in the empirical research conducted here. In other words, dissimilar individuals in terms of age, functional background, educational attainment, gender and tenure are not more likely (as the theory asserts), to have a more radical viewpoint to their colleagues prior to a discussion.

This finding means that recent advice for TMTs to appoint to their ranks more dissimilar persons (Carson et al., 2004), is likely to be of limited value in realizing assumed organizational performance benefits supposedly due to the causal link between demographic variation and multiple idea generation. Moreover, as wisely observed by Priem et al. (1999), demographic variables are very difficult to manipulate. *"Thus replacing one or more TMT members likely will not affect all demographic variable uniformly; a new member may increase tenure heterogeneity, decrease age heterogeneity, leave functional heterogeneity unchanged and so on"* (Priem et al., 1999, p 941).

As this study extended the concept of cognitive variation to include proximity to the team consensus, it was of interest to note that age dissimilar individuals are the most likely to hold a disparate pre-discussion ranking to the team consensus. As was explained in Chapter 6, age dissimilar individuals were not necessarily the oldest individuals, but the youngest individuals could attract a high dissimilarity score. This means that the youngest member of a TMT in a team of predominantly older colleagues could be more likely to arrive at a different pre-discussion conclusion to the eventual team consensus. At first blush this appears to support Hambrick & Mason's (1984) assertion that younger managers may be more innovative and prepared to take risks than older TMT members, and Hermann & Datta's (2005) observation that younger managers are more prepared to diversify, and Wiersma & Bantel's (1992) finding that younger managers are more likely to be involved in strategic change. However, these prior findings did not concern age dissimilarity, merely the central tendency of the TMT on age towards younger members as opposed to TMTs with predominantly older members. As noted in Chapter 2, only three previous TMT studies have investigated the effects of age dissimilarity. These found that age dissimilarity was a predictor of TMT turnover (Wagner et al. 1984; Jackson et al. 1991; and Wiersma & Bird, 1993). Thinking about the interpretation of the finding of this study in relation to those just cited, one might suggest that if age dissimilar persons consistently arrive at a pre-discussion judgement that is remote from the eventual team consensus, the individual could be disaffected to the point where he voluntarily exits the team.

With regard to the propensity of individuals to change their private opinions, it was found that gender dissimilar individuals maintain their pre-discussion views. In Chapter 7 it was suggested that it was mainly women (as they attracted highest dissimilarity values) were the least likely to change their minds. Whilst this statement is true, it should be remembered (as explained in Chapter 6) that men in mixed gender teams also attract a dissimilarity value, which is greater than men in all male teams (who would each attract a value of 0). So it would be inappropriate to conclude that it is solely women that do not change their minds. This topic and supporting literature will be discussed in relation to hypothesis 3, the results of which provide more clarity.

Whilst the number of statistically significant findings is encouraging, the effect sizes are all very tiny, meaning that practical implications for TMTs are uncertain. Moreover, as they are at the individual level across the whole sample, care needs to be taken in interpreting their relevance to executives within TMTs. This was the subject of Hypothesis 2.

Hypothesis 2 was that demographic diversity at the team level would be positively associated with cognitive diversity at the team level. Four statistically significant results arose from 25 relationships tested (13%) at the $p<0.10$ level, one more than expected by chance alone.

The findings, as reported in Chapter 7, are succinct. Demographic diversity is not associated with pre-discussion cognitive diversity, nor with pre-discussion proximity to the team consensus. Thus 'upper echelons' theory in its pure form is not supported. However, a negative relationship was observed, as expected between functional background diversity and cognitive change diversity, which suggests that teams characterised by a variety of functional backgrounds exhibit less proportion of cognitive change across the team than teams which share the same functional specialization. Support for this finding comes from Knight et al. (1999), who found a negative relationship with functional background diversity and strategic consensus.

As observed earlier, one of the ways this study sought to extend 'upper echelons' theory was to measure post-discussion executive cognition. At the team level it was found that functional diversity and gender diversity were positively associated with cognitive diversity post discussion. In other words, teams in which a variety of functional specialisms are represented, and mixed gender teams have a greater variety of opinions post discussion. Glick et al. (1993), Hambrick et al. (1996) and Knight et al. (1999) all found a similar relationship to this study between functional background diversity and levels of disagreement amongst TMTs. It could be argued that gender diversity is associated with cognitive diversity owing to the existence of strong demographic faultlines (Lau & Murnighan, 1998). A faultline refers to the representation of people with identical demographic attributes who tend to identify with each other to the point of becoming a sub-group embedded within the larger group. The assumption is that gender faultlines (i.e. opposing sub-groups of men and women in mixed gender teams) may hold different opinions to the larger groups in which they are embedded (Lau & Murnighan, 2005). In this study, all of the mixed gender teams could be considered to have strong male faultiness.

Moreover, this study went much deeper than the classic 'upper echelons' assumption of multiple idea generation pre-discussion by assessing the level of disagreement post-discussion; the degree to which individuals private opinions concurred with the team consensus post-discussion; and the extent to which individuals changed their minds. It is this extension to previous knowledge that offers the opportunity for rich insights into TMT functioning. No statistically significant associations were observed with regard to the pre-discussion measures. The finding that tenure diversity is positively associated with diversity around proximity to the team consensus post-discussion finds inferential support from Finkelstein & Hambrick (1990) who similarly found that shorter tenured teams could agree quickly to pursue novel strategies, but contrasts with Knight et al. (1999) who found that tenure diverse teams were more likely to agree on strategic issues. With regard to the latter, this was contrary to that hypothesised, and was contradictory to the other forms of diversity they tested, namely educational diversity and functional background diversity which were negatively associated with consensus as predicted. Hence, it can be concluded that this study supports the contention of other TMT researchers that demographic diversity is negatively related to consensus.

Consistent with previous research in this field, the findings in support of 'upper echelons' theory are scant and patchy (West & Schwenk, 1996; Priem et al., 1999), but generally seem to be in line with findings in prior research. It could be tempting at this stage to say that the observed relationships between demographic variation and measures of executive cognition validate an important assumption of 'upper echelons' theory. However, it would be incautious to do so as there does not appear to be a *systematic* link between aspects of demographic variation and the various measures of cognitive variation.

Hypotheses 3 was that demographically dissimilar individuals in demographically diverse teams would experience more cognitive dissimilarity. Fifty relationships were tested (25 at the $p<0.05$ level and 25 at the $p<0.10$ level), of which six were statistically significant (12%). This number is one more result than could have been expected to occur due to chance alone, so at best, one can say that the hypothesis received marginal support.

Specifically, it was found that tenure dissimilar individuals in teams, and individuals in teams which are tenure diverse, are more likely to have a pre-discussion opinion which is disparate to the eventual team consensus. This finding resonates with previous research which finds that tenure dissimilar individuals are more likely to leave the TMT before their fellow team members (Wagner et al., 1984; Jackson et al., 1991). One of the reasons tenure dissimilar individuals may be more likely to voluntarily exit the team may be due to their private pre-discussion opinions regularly proving to be inconsistent with team consensus decisions. Executives whose personal opinions are not compatible with team goals tend to experience negative conflict (Amason, 1996; Janssen et al., 1999), although the extent to which this influences their decision to leave the team has not yet been explored.

Individuals in educationally diverse teams were more likely to have a diparate opinion to their peers post-discussion. Knight et al. (1999) similarly found a negative relationship between educational diversity and consensus, suggesting that individuals with a different educational orientation are likely to have a different mental model to their peers (Janssen et al. 1999).

The current study also found that functional background diversity and educational diversity were positive predictors of individuals' zero cognitive change. This suggests individuals in teams that are characterised by individuals from different functional specializations and different educational ontologies tend to hold onto their own mental models, rather than enjoy a shared cognition with the rest of the team. Again, this is supported by Knight et al. (1999) who found that functional background diversity and educational diversity were negatively associated with strategic consensus in TMTs as did Glick et al. (1993). Glick et al. (1993) and Knight et al. (1999), are unique in the TMT literature in that they actually defined and measured cognitive diversity and strategic consensus. However, there was only one cognitive variation measure in each case, so inferences need to be drawn in relation to the multiple measures used in this study.

It was found that gender diversity predicts negative individual cognitive change. This means that men and women in mixed gender teams do not change their minds. This is borne out when the data are tested solely at the team level (hypothesis 2) there is no statistically significant relationship between gender diversity and cognitive change diversity. That is to say, it is not a team level phenomenon, a proportional spread of men and women in a team does not equate to a proportional spread of cognitive change. Or, to put it another way, an all male team for example is no more or less likely than a mixed gender team to be subject to cognitive change. Thus it is fair to say in reporting this finding, that it is gender dissimilar individuals, generally women, but possibly men too, that do not change their minds once made up. Furthermore, individuals in mixed gender teams show the least propensity for individuals to change their minds. This result implies that there is greater propensity to change personal opinions by gender similar individuals, which one would have to say would be generally men, as they were the majority in the sample.

It was outside the scope of this study to determine how or why gender diversity and dissimilarity would have such an effect, whether this could be due to populist notions of a mis-match between communication styles (CfWBR, 1994) or of 'yes-men' agreeing to follow a company line (Janis, 1977). This could be a legitimate line of enquiry for a further research project. As was noted in Chapter 2, gender has not been studied in TMTs before, so any inferential support for these findings needs to be drawn from the wider diversity literature, in which it is widely held that owing to different social experiences had by males and females, that gender will make a fundamental difference to individuals' perceptions, although the empirical evidence to support this argument is scant (Forte, 2004). Recent studies have found that no statistically significant differences exist between the judgments made by male and female executives (see Forte 2004 and Church, 1997 for examples), which Forte (2004) argues is due to them being 'corporate members'. That is to say, they are trained to think and judge strategic situations according to organisational orientations in an asexual way. Like other researchers before her, Forte (2004) found that women typically took higher values on her survey instrument, and attributed the lack of statistical significance to the low representation of women (39 out of 400 managers sampled), which meant that the statistical power of the test was weak. Concurring with Forte (2004), in this study, the rankings of candidates (judgments) by men and by women did not differ appreciably. The finding that women in mixed gender teams did not tend to change their minds is of interest to TMTs. One might reasonably extrapolate from the simulation and the results that a real situation might occur where a female executive had personally selected candidate A yet the team consensus was candidate B. She remains convinced that candidate A is the correct choice, however, candidate B is appointed. This may have ramifications for relational dynamics between her and candidate B or for her attachment to the team (see Pelled, 1996).

In summary, the implications of these findings concerning Hypotheses 1 – 3 for TMTs are that the received wisdom of getting a diverse team together to make a creative decision is not necessarily valid. Indeed, it is likely to make very little difference. However, TMTs should be aware that the stated consensus decision during discussion is not the one to which all members hold privately afterwards. It appears that certain demographic characteristics (functional background, educational attainment and gender) are associated with a recalcitrance to change, whilst the most educationally dissimilar person is more likely to hold the most dissimilar opinion to the rest of the team post discussion. However, caution is in order here as discussed in Chapters 7 and 8, the variance explained by these results only ranges from 1% to 14%, which means that for the most part, the relationships observed in this study are not likely to be problematic for TMTs.

For researchers in this arena, it can be said regarding the findings for Hypotheses 1 – 3 of this study, that although 'upper echelons' in its pure form is not supported, the study has the potential to extend theory in two ways. First, the results indicate that there are subtle nuances between individual versus team cognition. This means that in the same way as individual demographic dissimilarity is not the same as team diversity (discussed at length in Chapter 2), cognitive dissimilarity is not the same as cognitive diversity. Second, the findings suggest that there are differences between pre- and post-discussion cognition. This has implications in terms of the model first presented in Figure 4.1 (p 76), which suggested (based in most part on 'upper echelons' theory) that there was a direct link between demographic and cognitive variation that was not necessarily influenced by indirect factors such as team processes.

The vagaries of studying demographic variation were commented on in Chapter 2, particularly with regard to inconsistent and contradictory results. This means that 'demographic diversity', despite being a ubiquitous term, cannot be viewed as a gestalt of the component factors, as each type of dissimilarity or diversity needs to be treated separately. That is to say, age dissimilarity for example, is not the equivalent of gender dissimilarity. What this study shows is that cognitive variation is similar, in that pre- and post-discussion cognition, proximity to the team consensus pre- and post-discussion, and cognitive change do not constitute an amalgam called 'cognitive diversity'. The relationships with the cognitive variation variables, as with the demographic variables, are inconsistent and do not equate with each other. Hence, it may be helpful to reflect this in the model more explicitly.

Moreover, previous studies of strategic consensus have either focused on links to organizational performance (e.g. West & Schwenk, 1996) or on the process of consensus formation using mechanisms such as dialectical enquiry (e.g. Schweiger et al., 1986) or, consensus formation as an outcome (dependent) variable (e.g. Glick et al., 1993; Knight et al., 1999). Although team processes will be reviewed in more detail in the next section, what is of interest to note here, in relation to the model is the potential juxtaposition of the component parts with some of the cognitive variation variables.

A revised version of the model is presented in Figure 11.1 below which suggests that the processes surrounding the team discussion (including decision belief as this was a team level only phenomenon relating specifically to the team consensus reached during the discussion) may mediate the relationship between demographic and cognitive variation pre- and post-discussion.

In concluding this section it is appropriate to point out that the effect sizes for all of the statistically significant findings were very small, which limits the practical implications for real TMTs. Both the issue of chance results and effect sizes will be returned to later in this chapter. Suffice it to say at this point, the results arising from this study do not provide convincing validation for the assumption underpinning 'upper echelons' that demographic variation leads to cognitive variation.

The next section will recap on the findings concerning demographic diversity and cognitive variation (at the individual and team levels) with team processes (Hypotheses 4 – 6).

11.3 Demographic and Cognitive Variation and Team Processes (Hypotheses 4 – 6)

Hypothesis 4 was that demographic diversity (team level) would be negatively associated with the team processes, namely: procedural rationality, frequency of meetings, reflexivity and psychological safety.

Of the nine hypotheses in the study, it is fair to say that Hypothesis 4 received qualified support, but even so, only four statistically significant results arose from 20 relationships tested (slightly more than one would expect to occur by chance alone). In three of the four cases, gender diversity was found to be negatively associated with procedural rationality, frequency of team meetings and psychological safety. In the final case, age diversity was negatively associated with reflexivity.

Procedural rationality refers to the extent to which teams exhaustively search and assimilate information during the decision-making process (Fredrickson, 1984; Eisenhardt & Bourgeois, 1988; Dean & Sharfman, 1996). As most studies of procedural rationality comprise retrospective questionnaire surveys of decision-makers, little attention has been paid to the dynamic interactions in a team discussion. Of course, there is likely to be a social desirability factor where subjects respond that they do engage in procedural rationality. Yet, there could be a myriad of interpersonal factors that inhibit or facilitate the sharing and consideration of information. Even the process of verbalizing ideas and turn-taking during a discussion undoubtedly influences the amount of information considered by a team. The finding from this study is that mixed gender teams are less likely to engage in procedural rationality. As gender diversity was the only predictor, one has to look for interpersonal male/female aspects in interpreting this finding. The most likely explanation is the mismatch in communication styles between men and women during team discussions, as contrary to popular opinion, women tend to contribute more to discussions than men (Graves & Elsass, 2005). The effect size for this result was very tiny, so the practical dilemma for top management teams as to whether gender diversity is a problem for proper information search, is negligible.

The finding of this study, that gender diverse teams meet less frequently, is consistent with recent experimental work that finds that strong faultline groups (i.e. those dominated by one or more demographic sub-groups) have less frequent cross-gender communication (Lau & Murnighan, 2005). This means that the sub-groups communicate more frequently internally, that is to say, with persons of the same gender. The obvious practical implication of this finding is that mixed gender TMTs should consider their communication structures when approaching important strategic decisions. However, radical change to manipulate communication frequency in mixed gender teams may not achieve a marked difference as the effect size observed for this relationship in this study was tiny.

That gender diverse teams experience less psychological safety than all male teams appears to be consistent with theory which asserts that women are viewed as less competent by their male counterparts owing to the latter's higher social status relative to women (Wagner & Berger, 1997). It is further argued that the status differential is reflected in team interactions, such that high-status men make more task contributions, act more confidently and exert more influence (Carli & Bukatko, 2000), which in turn means they experience more psychological safety (Edmondson, 1999). Although the most recent research casts some doubt concerning the supposed relationship between the level of exchange and attachment to the team, (Graves & Elsass, 2005), there is an abundance of studies that confirm that gender dissimilarity has negative implications for team performance, and generates more conflict (see for example, Chatman & Flynn, 2001; Pelled, 1996).

This study contributes to this literature by demonstrating a negative effect between gender diversity at the team level and psychological safety. It is likely that the naturally occurring representation of men and women in the 23 teams was balanced in such a way as to facilitate strong faultlines (i.e. dominant male vs. tiny female sub-groups), whereas weak gender faultlines have been found to result in more psychological safety amongst experimental student teams (Lau & Murnighan, 2005). In any case, as discussed in Chapter 9, the effect size of this particular relationship is small (Cohen, 1988). This issue will be addressed again later when a commentary on the effect sizes of all results is discussed. Suffice it to say here that gender diversity explains 20% of the variance in psychological safety. This means that the relationship between gender diversity and psychological safety was not an issue in 80% of team discussions. Hence CEOs and boards considering the composition of their TMT need not overly concern themselves about the negative effect of gender diversity with regard to psychological safety.

With regard to reflexivity, it was found that mixed age teams are less reflexive. This finding is supported from the research literature which observes a potential for dissonance between older, more risk averse managers and their youthful, creative risk-taking counterparts (Hambrick & Mason, 1984), resulting in more internal conflict amongst teams characterised by age diversity (Williams & O'Reilly, 1998). In practical terms, this finding suggests that age diverse teams tend to suffer from 'biased closed-mindedness', and that they 'inadequately evaluate new information, simplify problems and place unwarranted confidence in their initial positions' (see Tjosvold, 1992). The practical implication of this study's finding is that age diverse TMTs should seek to articulate any generation gap in viewpoints in the interests of making an excellent rather than sub-optimal decision (Tjosvold, 1992; 1994; Tjosvold et al., 2004). West (1996, 1998) has promulgated reflexivity as the process that can be adopted by TMTs to induce such constructive controversy, particularly by using mechanisms such as devil's advocacy and dialectical enquiry (Schweiger et al., 1986). However, a word of caution is in order, the effect size of this result was tiny (R^2 <0.20), which means that one should not overstate the negative relationship between age diversity and reflexivity, nor the need to introduce compensatory mechanisms in real TMTs.

Hypothesis 5 was that two team processes (procedural rationality and frequency of team meetings) would decrease cognitive variation and that two other processes (reflexivity and psychological safety) would increase cognitive variation. Twenty relationships were tested at the $p<0.10$ level (team), of which, three (two supporting and one opposing) were statistically significant. This number of results is about what one would expect to occur by chance alone. Furthermore, for each type of process, only one of each was found to be statistically significant, these were frequency of team meetings and psychological safety.

Specifically it was found that TMTs which meet more often tend to experience less cognitive diversity after a discussion. This would suggest that in terms of attaining a shared cognition or mental model, frequency of team meetings is important as this is where the social norms of the group evolve and collective attitudes are formed such that decisions are made according to a consistent pattern

(Glick et al, 1993; O'Reilly et al., 1993). This finding concurs with previous research which finds that verbalization during discussion increases comprehension (Johnson & Johnson, 1987; Glick et al., 1993), speed of decision making (Eisenhardt & Schoonhoven, 1990) and collective commitment to decision-outcomes (Heath & Gonzalez, 1995).

Contrary to expectation, it was found that there was less proximity amongst team members' rankings compared to the team consensus after discussion in teams which meet more often. Similarly, Smith et al. (1994) were also surprised to find a negative relationship between frequency of meetings and organizational performance, which was opposite to their prediction. They concluded that whilst team meetings were beneficial for sharing and assimilating information, meetings absorbed valuable time that could instead be spent on task oriented activities. In this case, the proposed interpretation by Smith et al. (1994) is not entirely credible, as on the one hand it is found that frequency of meetings predicts less cognitive diversity post-discussion, whilst on the other, frequency of meetings predicts greater variation around the team consensus post-discussion. At face value, one can say that there is a convergence of rankings post-discussion, but this concurrence is not centered on the consensus decision. A plausible explanation is that teams will generally tend to agree on the first choice option (plan A) and perhaps the second choice option (plan B), but if a couple of members have an entirely different opinion (Plans C and D) then there will be diversity around the agreed decision-outcome.

It is worth restating the basis on which the cognitive diversity and cognitive cohesion diversity (proximity) measures were computed. Each participant was asked to rank order four candidates. What may underlie the results is that teams' opinions tended to converge on their number one candidate, and perhaps also their second candidate leading to a finding of less cognitive diversity. The second measure compares the proximity of all the individuals' rankings to the team consensus ranking. If say, two team members each chose a completely different rank order of the four candidates to the team consensus, this would result in there being diversity around the team consensus (See Chapter 6 for a detailed discussion of the rankings and computation of measures).

What tends to happen in real life (as borne out by discussions with the 23 teams in this study), is that candidates may be ranked in terms of their suitability as part of the short-listing process, but after the interview, executives will only have a first and second choice in mind. That is to say, the ranking of a third or fourth candidate is irrelevant, because if the first and second choice candidate refused appointment, the company would likely re-advertise. In teams which meet reasonably often, there is likely to be a shared appreciation of what kind of candidate is required for the job consistent with the circumstances of the company. However, the established norms and patterns of discussion, which as noted above, are a function of frequent team meetings may not give sufficient airtime to opposite viewpoints, hence not all people agree with the consensus at the end of the day. Hence the two statistically significant findings of this study, whilst opposing in terms of simple prediction, are consistent with reality.

What is of interest to TMTs is that just because a consensus decision is reached, in which all members publicly verbalise their assent, this does not mean that all members agree with the decision. This research consistently demonstrates that there can be a range of personal and indeed sub-group opinions that are still privately held despite discussion. What ramifications this may have for decision commitment and implementation is beyond the scope of this study.

It was also predicted that procedural rationality along with frequency of team meetings would reduce cognitive variation. However, there were no statistically significant relationships with procedural rationality. This lack of results may suggest that type of decision process (intuitive or comprehensive) is not important for reducing cognitive variation either pre- or post-discussion. This is a particularly interesting observation in as much as, procedural rationality is the most popular and arguably the most important decision process in 'upper echelons' research (Papadakis & Barwise, 2002; Hodgkinson, 2001). The practical implication for TMTs is that adopting a particular type of process is no more or less likely to ~~decrease~~ increase consensus, so they need not impose a type of decision process specifically to reduce disagreement. Other benefits attributed to procedural rationality, such as being more effective in achieving decision objectives over a 2 year period (Dean & Sharfman, 1996) were not tapped by the current study.

Two processes were argued to increase cognitive variation, these were reflexivity and psychological safety. As predicted, it was found that teams characterized by psychological safety experience more cognitive change. This is deemed to be beneficial as in order to implement a consensus decision, which has to be an amalgam of several individual opinions, which often differ considerably regarding what is considered to be an appropriate course of action (Forbes & Milliken, 1999), someone needs to change their minds in the interests of the organization. Psychological safety allows them to do so without feeling that they have capitulated or that they appear foolish to their peers (Edmondson et al., 2003; Janssen et al., 1999; Pelled et al., 1999). As discussed in Chapter 3, most of the research into psychological safety as a team level construct is with teams lower in the organizational hierarchy than TMTs. This research demonstrates that it is a feature of TMT dynamics also, and lends some support to Edmondson et al.'s (2003) argument that deliberately trying to engender more psychological safety in teams will make them more effective. However, the effect size of this result was tiny (R^2 <0.10). Moreover, it was the only time that psychological safety was found to be a predictor in the analysis, which tempers any prescription as to its importance.

The second process predicted to increase cognitive variation was reflexivity, however no statistically significant relationships were observed. This seems to indicate that reflexivity is likely to be a non-starter for increasing cognitive variation, despite theoretical aspirations as to its efficacy (West 1996; 1998).

In Chapter 2 the difficulty in gaining access to TMTs was discussed, which laid the groundwork for Chapter 3 in which it was observed that there are very few if any published works on observed processes within TMTs. Indeed, it was noted that most process studies of TMTs rely on MBA students in role play, or at best, retrospective self report questionnaires as to process. Even then, as reported in Chapter 3, the most consistent constructs in TMTs are procedural rationality and frequency of team meetings. This study took these two processes, which it was argued were beneficial for cognitive variation and juxtaposed them with two opposing processes about which much has been written, and indeed, studied at other levels in organizations. In so doing, the research sought to peer into the so-called 'black box' (Lawrence, 1997) that has thus far eluded 'upper echelons' researchers. What the findings from this study demonstrate is that the contents of the 'black box' continue to be hard to pin down and that much more work remains to be done to understand the complexities of TMT processes.

Hypothesis 6 was that team processes would influence individual cognitive dissimilarity, and particularly demographically dissimilar individuals embedded in demographically diverse teams. Fifty relationships were tested at the $p<0.10$ level, of which, five statistically significant results could be expected to occur due to chance alone. Five statistically significant associations were observed, four supporting the hypothesis and one opposing.

These five statistically significant associations were all pertinent to the first part of the hypothesis that there would be a relationship team processes and individuals' cognitive dissimilarity relative to their team. Similar to the team level proposition just discussed regarding hypothesis 5, it was posited that for teams that use procedural rationality and who meet often that there would be a corresponding reduction in dissimilar views amongst individual team members. Conversely, it was suggested that teams characterised by reflexivity and psychological safety would make individuals in those teams more radical in their viewpoints. However, no statistically significant associations were noted for reflexivity and psychological safety. The second part of the hypothesis, that team processes would have greater influence on the cognition of demographically dissimilar individuals was not supported.

Supporting the first part of the hypothesis, it was found that procedural rationality positively influences proximity to the team consensus pre-discussion. In practice, this means that in teams which regularly tend to be more exhaustive and analytical in their decision-making (Frederickson, 1984; Eisenhardt & Bourgeois, 1988), the individual members are more likely to arrive at an individual perception that is close to the eventual team consensus. Simons et al. (1999) in a study of 57 TMTs found that procedural rationality fomented constructive debate during team discussions, which ensured that TMT decisions covered a wide range of options. What the finding from this study suggests is that when teams consistently consider and debate a wide range of options during group decision-making, individual members probably privately explore and investigate the available information in much the same way (Mintzberg et al., 1976). This finding is consistent with the interpretative approach to decision-making, which argues that executives' as individuals and teams consciously structure information for decision-making consistent with what they believe to be relevant (Sparrow, 1994; Hodgkinson & Sparrow, 2002) and rely upon their analysis in making the choice (Dean & Sharfman, 1996). Procedural rationality is generally held to be the most beneficial decision-making process, is deemed to be consistent with decision-making experience (Hendry, 2000; Hodgkinson, 2001) and is thought to be the most straightforward research construct for understanding decision-making experience in TMTs (Papadakis & Barwise, 2002).

Should TMTs who employ less exhaustive and analytical decision-processes try to implement procedural rationality as a means of encouraging executives to think the same way such that their opinions converge around consensus even before discussion? It would be unwise to do so, on this basis alone, as the effect size for this result was trivial, explaining only 1% of the variance. Moreover, procedural rationality did not behave consistently in reducing cognitive dissimilarity as predicted, as the next result illustrates. Furthermore, the research failed to find a relationship between procedural rationality and three of the five cognitive dissimilarity measures.

Opposing the hypothesis, procedural rationality was found to be a positive predictor of cognitive dissimilarity post-discussion. This means that instead of reducing cognitive dissimilarity, procedural rationality actually increases cognitive dissimilarity amongst team members subsequent to a discussion. The research literature from which the proposition was derived offers little in the way of explanation of this contrary result. In Chapter 9 the interpretation was offered that if a team regularly uses a comprehensive approach to decision-making during team meetings, that an individual who rationalised their choice of candidate before a discussion might still be convinced that their first choice was right after the discussion. In any event, the effect size was marginal and the relevance for TMTs debateable.

With regard to the second process posited to reduce cognitive dissimilarity, three statistically significant associations were observed with frequency of team meetings in support of the hypothesis. Specifically, it was found that more frequent team meetings reduces dissimilarity in rankings pre- and post-discussion and encourages people to change their minds more in the interests of agreement. As discussed in Chapter 3, the research literature consistently attributes speed and efficiency in decision-making to the frequency of team meetings (Eisenhardt & Schoonhoven, 1990; Smith et al., 1994; Papadakis & Barwise, 2002). Moreover, it is argued that this is due to more frequent meetings socialising the team into procedural norms and consistent patterns of decision-making (Glick et al., 1993;

O'Reilly et al., 1993). This cluster of findings suggests that one of the reasons that speed and efficiency gains occur is due to the reduction in cognitive dissimilarity both pre- and post-discussion. In other words, the more often a team meet, the more similar individuals' selection of appropriate options will be. Furthermore, when teams meet between three and six times per month, individual executives show greater propensity to change their personal opinions, which is necessary to arrive at a team consensus.

For some decisions, one can imagine that such convergence of opinions and frequent team meetings in some circumstances might be particularly beneficial. For example, with relation to diversification posture, convergence of opinion is important in order to move speedily to take advantage of market conditions (Eisenhardt & Schoonhoven, 1990). Similarly with regard to this research, a TMT would want to confidently appoint a candidate to their ranks who was universally approved. Sometimes however, reduced cognitive dissimilarity is not always beneficial. This refers to circumstances where TMTs miss strategic opportunities, or misinterpret cues in the marketplace due to the fact that they have a singular way of looking at and interpreting the environment they operate in (see Hodgkinson & Sparrow, 2002 for examples). It was beyond the scope of this research to ascertain whether the reduction in cognitive dissimilarity arising from frequency of team meetings was beneficial or detrimental in certain circumstances and not in others. Future research would do well to consider this aspect of strategic cognition, paying particular attention to antecedent factors that could influence TMTs to take a particular world view which could result in either a market innovation or an industry blindspot.

The two team processes put forward as facilitators of cognitive dissimilarity amongst individuals, reflexivity and psychological safety, exhibited no statistically significant results. This suggests that neither has an influence on the way individuals think in relation to their peers. Although the lack of statistically significant results concerning these processes may seem disappointing, it should be remembered that this study is unusual in that it investigated the effects of team processes on individual cognition. This has never been done before at TMT level, and was made possible by the availability of multi-level modelling. As was mentioned with regard to the previous hypothesis, the 'black box' still remains an area of enormous potential in terms of research. Other processes which may be important will be covered in a later section.

11.4 Team Processes and Decision Belief (Hypothesis 7)

Hypothesis 7 was that teams which regularly employed procedural rationality and met more frequently would tend to report more confidence, satisfaction and effectiveness, whereas reflexivity and psychological safety would result in less confidence, satisfaction and effectiveness. Twelve relationships were tested, of which only one was statistically significantly correlated, but when subjected to a regression analysis controlling for team size, the relationship disappeared. This is less than one might expect to occur through chance alone.

Specifically, it was found that reflexive teams reported less confidence in their decision. Although not explicitly studied before, this result is in line with predictions from theory concerning reflexivity, constructive controversy and internal conflict (see Swift & West 1998; Tjosvold, 1995; Tjosvold et al., 2004; Agrell & Gustafson, 1996). It is very likely that the conflict between individuals putting forward opposing viewpoints during the team discussion means that there are simply too many options available and opinions to consider. So when an option is selected by the team as its consensus decision, individuals are still perplexed by 'yes, but…' and 'what if…' questions about some of the other options. Hence, when they are asked how confident are you that the team made the right decision?, they still have some nagging doubts based on the fact that the pros and cons of so many options were considered.

This obviously has ramifications in TMTs as to how the consensus decision will be implemented and the commitment of individuals who were not confident about the decision (Schweiger et al., 1986; Sniezek & Henry, 1990; Heath & Gonzalez, 1995; Agrell & Gustafson, 1996). It may also be the case that if a decision does not turn out well, that those individuals who were not confident in its efficacy will lack confidence in future decisions also, perhaps even having an 'I told you so attitude', which could impact upon future team discussions. They may begin to pursue private agendas and engage in political behaviour (Dean & Sharfman, 1996) that subverts team decisions and gives them more individual prominence, meaning that the team ceases to work together for the common good of the organisation (Nadler, et al., 1979).

The practical implication from this finding is that reflexivity inherently, engenders a lack of confidence in decisions reached. TMTs would do well to consider this aspect of the decision-making process, particularly as reflexive practices are recommended to take place at the end of a discussion, when a team decision has been reached. For example, West (1994) proposes that teams take time to reflect on the decision just made and try to point to all the pitfalls, errors, possible negative consequences and then ask themselves searching questions along the lines of 'are we still confident we have made the right choice?' One can imagine a TMT discussion, in which a team has taken considerable time and care to reach a consensus decision perhaps engaging in such reflexive questioning. Some on the team might feel irritated that this is an unnecessary adjunct to the discussion now that all the arguments have been worked through, perhaps they are eager to get back to more pressing activities. For others, it might be an opportunity to revisit their pet arguments in an attempt to overturn the decision, so as to favour their own preferred option, which was not adopted as the consensus. It is perhaps not surprising that such a practice would lead to dissonance amongst the team so that they report lack of confidence.

Although the finding that reflexivity was negatively associated with decision belief was in line with prediction and theory, the overall supposed beneficial nature of reflexivity for TMTs is hard to pin down. This was the only time in the study that reflexivity was found to be a predictor, and as mentioned above and in Chapter 9, although not strictly necessary, when one controls for team size the relationship disappears. This surely must cast doubt on the efficacy of reflexivity as a process specifically theorised to apply to TMTs and strategic decision-making (see West et al., 1997). It seems that the ability to 'overtly reflect upon objectives, strategies and processes, adapting them to current or anticipated circumstances' (West, 1996, p 559) is neither here nor there in the real world of TMT decision-making. In other words, the practical implication from this study is that TMTs should neither try to avoid reflexivity nor try deliberately to become more reflexive.

Before leaving a discussion hypothesis 7, one should point out that there were four processes posited to have an effect on decision belief. No relationships were noted with the other three. That is to say, procedural rationality, frequency of team meetings and psychological safety were not associated with perceived satisfaction, confidence or effectiveness as predicted. It could be argued that this was due to the relatively small dataset, however issues such as ratios to cases were discussed in Chapter 7, and it was clarified that the dataset was of sufficient size to elicit the relationships sought. Hence, alternative explanations must be offered. This issue will be addressed in more detail in the theoretical contribution section of this Chapter.

11.5 Demographic and Cognitive Variation and Decision Belief (Hypotheses 8 – 9)

Hypothesis 8 was that teams characterised by greater diversity in variation in age, functional background, educational attainment, gender and tenure would report less satisfaction, confidence and effectiveness. Fifteen relationships were tested, from which one or two statistically significant results could be expected to occur due to chance alone. Two results were statistically significant in support of the hypothesis, and one was opposing.

The key findings to emerge form H8 are that, as predicted, educationally diverse teams, and gender diverse teams are less confident in their ultimate decision. With regard to the educationally diverse teams, this negative finding is interesting in light of Simons et al.'s (1999) finding that educational diversity is associated with more debate during discussions. In terms of parsimony with Simons et al. (1999), it is likely that greater debate could lead to the polarisation of views, with belief in the efficacy of the decision being based on educational orientation. As discussed in Chapter 2, the literature argues that different educational backgrounds lead to different ways of approaching information and decision-making. Hence, it is consistent with experience that educational diversity is associated with more debate, and with less confidence in the consensus decision if this is in opposition to the educational orientation of some of the constituent parties in the team.

As to why gender diversity would be associated with less confidence in the decision, one can put forward similar arguments that sub-groups of men or women in the team have opposing viewpoints as to the efficacy of the consensus decision.

A number of controlled experiments (typically using students) have found that mixed gender 'teams' surpass all male teams in business decision-making scenarios (see for example, Karakowsky & Elangovan, 2001; Webber, 1987). Getting the right gender representation on such teams is a delicate balance. For example, Rogelberg & Rumery (1996) examined the effects of five gender compositions (all-male, lone female, balanced-gender, lone male and all female) and decision quality (as measured by the achievement of an a priori solution in a problem solving task). They found that as the number of males in a team increased, so did decision quality, but, lone-female teams outperformed all other gender compositions. According to Lau & Murnighan (2005), one female in a group constitutes a strong demographic faultline for male representation. It is reasonable to assume that the current dataset was comprised of strong gender (male) faultlines according to Lau & Murnighan's (2005) definition. The extent to which clusters could form within the 23 teams around demographic attributes was discussed in Chapter 7, and ICCs which showed that there was sufficient heterogeneity to study diversity in the dataset was presented in Table 7.3. The computation for faultlines (Lau & Murnighan, 2005) was not applied however, as this was not the feature of interest at the time. What is intriguing is that in an earlier work (Lau & Murnighan, 1998) it was posited that strong faultlines would facilitate more effective team work, but in the first experimental test, it was found, contrary to expectation, that strong faultlines have a deleterious effect (Lau & Murnighan, 2005). The findings of the current study are more in line with the latter.

Opposing Hypothesis 8, tenure diversity was found to be a positive predictor of confidence in the decision. It is likely that individuals who have been on the team a comparatively short time (i.e. less tenure) will follow the lead of those who have been on the team a comparatively long time (i.e. more tenure). These long-standing individuals, because of their tenure and experience with the TMT decision making process, are likely to have more confidence in their decisions and be content with the status quo (i.e. familiarity with the process). Collectively, then, tenure diversity predicts confidence in the decision. There is some support for this supposition from the literature. Variation in team tenure is reportedly responsible for power hierarchies and communication patterns in teams (Clark et al. 1997), although this has generally been about groups of longer serving managers vs. groups of shorter tenure managers, not diversity within the team.

Although statistically significant relationships were noted for three characteristics of diversity and one variable for decision belief, no relationships were observed with functional diversity or age diversity. No relationships were observed with regard to perceived satisfaction or effectiveness. This ratio of findings to relationships tested is similar to previous hypothesis and suggests that there is no consistent link between demographic diversity and decision belief. This subject is covered in more depth in the theoretical section of this chapter. What this means in terms of practical implications for TMTs is that diversity need not be avoided, nor actively sought in order to engender greater belief in the decisions made by TMTs.

Hypothesis 9 predicted a negative relationship between cognitive diversity and decision belief. Fifteen relationships were tested at the $p<0.10$ level, of which only one was statistically significant, a result which one could reasonably expect to occur due to chance alone.

In support of the hypothesis, it was found that TMTs that experience the greatest variation in their rankings after discussion, report the least confidence in their consensus decision. This finding is consistent with that of Amason (1996) who found that when tension existed in the TMT concerning the decision, there was a corresponding negative association with confidence that the decision was the right one. The problem coupled to lack of confidence, is lack of commitment by the team to follow through and implement the decision (Sniezek & Henry, 1990; Schweiger et al., 1986). Moreover, lack of confidence in the appropriateness of the decision is also linked to political and subversive behaviour in which participants pursue other agendas (Dean & Sharfman, 1996), and to dysfunctional dynamics amongst teams who stop working together for the common good of the organisation (Nadler et al., 1979). What this research shows is a plausible reason for lack of confidence in a consensus decision is the existence of a range of personal viewpoints being held by the team. Future research would do well to build on this, by taking forward the argument in the literature that lack of confidence equates to lack of commitment. It is likely that either a case study approach or a longitudinal study over the life-cycle of a decision would be necessary as many strategic decisions take months or even years to implement (Mintzberg et al., 1976).

In summary, the findings across all nine of the hypotheses are sparse and inconsistent, and the resulting prescriptions for real TMTs, are of necessity, tentative.

11.6 Continued Faith in 'Upper Echelons' Despite Lack of Results

The over-riding outcome of the study described in this thesis is the lack of strong or consistent support for 'upper echelons' theory. In point of fact, only 27 statistically significant effects were found from the 177 core relationships tested, before any tests for interactions (see Table 11.1). Moreover, these were derived using one tailed tests of statistical significance at the $p<0.10$ for the team level analysis and $p<0.05$ for the individual level analysis. Superficially, the number of results may seem promising. However, with so many relationships tested, some could be expected by chance. More specifically, as a rule of thumb, for relationships tested at the $p<0.05$ significance level, one can expect that at least one in twenty of the relationships tested to be statistically significant due to chance. As many of the relationships in this study were tested at the $p<0.10$ significance level (at the team level), one could expect that at least one in ten of these could achieve statistical significance due to chance. That means that overall, some 15 - 20 statistically significant results in the current study could have occurred due to chance alone (assuming that these were normally distributed). Thus, overall, the number of findings is not very much above that expected by chance. As discussed in the previous section, the number of findings per hypothesis is not particularly convincing. Hypotheses 2, 5, 6 and 8, receive no more support than could be achieved by chance alone, hypotheses 7 and 9 receive slightly less support than that expected to occur by chance, whilst hypotheses 1, 3 and 4 receive one or two more results each than could be expected by chance alone, but those for hypothesis 1 disappear if a control for team size is included.

Table 11.1 groups the dependent variables according to predictors. The first group, cognitive variation at the individual level, achieved 13 statistically significant results from 70 relationships tested. This is approximately three times more than one would expect by chance alone. The second group, cognitive variation at the team level, achieves one more than the number of relationships (6) expected by chance alone. The third group, team processes, achieves one less relationship than expected by chance alone, and the fourth group, decision belief, achieves exactly the number of results that would have been expected by chance. Superficially, this re-grouping appears to offer a more promising ratio of statistically significant findings to relationships tested.

As has been mentioned previously, such a small number of supportive findings is common in this literature, and has not deterred researchers in the past from accepting such low numbers as supporting their hypotheses. For example, Glick et al. (1993) found only three statistically significant results from a possible 49 relationships tested (which is the number one would expect to occur due to chance alone), yet felt comfortable in reporting sufficient support for 'upper echelons' theory, with respect to the single type of diversity (functional) with which the results were observed. Non-results for other types of diversity (i.e. age and tenure) were simply dismissed and given short shrift, whilst the supportive findings were expounded upon.

'Upper echelons' has long accepted that demographic characteristics are "rough surrogates" (Michel & Hambrick, 1993 p 16) for accessing cognition and that such measures "contain more noise than purer psychological measures" (Hambrick & Mason, 1984). However, as pointed out by Markóczy (1997), 'upper echelons' theorists have, conveniently, never "defined the level of 'roughness' that is acceptable" (Markóczy, 1997 p 1240). Furthermore, based on her equivocal results, in which demographic differences in total explained only 17.2 percent of the variance in one of two belief clusters tested, Markóczy (1997) put forward the following argument for her rejection of the 'upper echelons' position that demographic proxies are a good substitute for cognitive variables:

"People accept without much hesitation that the response to the question "How tall are you?" is a good substitute for actually measuring height for many purposes. Stretching the notion of substitution a bit, a sugar substitute people would judge to taste 17.2 percent like sugar on a rating scale would not be considered a substitute at all" (Markóczy, 1997 p 1240).

Table 11.1 Summary of Statistically Significant Results

Dependent Variables	Independent Variables	*Direction as Predicted*	*P*
Cognitive Variation at the Individual Level: (n = 13 results from 70 relationships tested, 19%)			
Cognitive Dissimilarity Pre-Discussion	Frequency of Meetings (t)(m)	-ve	0.05
Cognitive Cohesion Dissim. Pre-Disc.	Age Dissimilarity (i) (c) Tenure Dissimilarity (i)(m) Tenure Diversity (t)(m) Procedural Rationality (t)(m)	-ve -ve -ve +ve	0.05 0.05 0.05 0.05
Cognitive Change Dissimilarity	Gender Dissimilarity (i)(r) Functional Diversity (t)(m) Educational Diversity (t)(m) Gender Diversity (t)(m) Frequency of Meetings (t)(m)	+ve +ve +ve +ve -ve	0.01 0.05 0.05 0.05 0.01
Cognitive Dissimilarity Post-Discussion	Educational Dissimilarity (i)(r) Educational Diversity (t)(m) Frequency of Meetings (t)(m)	+ve +ve -ve	0.01 0.05 0.05
Cognitive Cohesion Dissim. Post-Disc.	Nil		
Cognitive Variation at the Team Level: (n = 6 results from 45 relationships tested, 13%)			
Cognitive Diversity Pre-Discussion	Nil	n/a	n/a
Cognitive Cohesion Diversity Pre-Disc	Nil	n/a	n/a
Cognitive Change Diversity	Functional Diversity (t)(r) Psychological Safety (t)(r)	-ve +ve	0.05 0.10
Cognitive Diversity Post-Discussion	Functional Diversity (t)(r) Gender Diversity (t)(c) Frequency of Meetings (t)(r)	+ve +ve -ve	0.10 0.05 0.05
Cognitive Cohesion Diversity Post-Disc.	Tenure Diversity (t)(r)	+ve	0.05
Team Processes: (n = 4 results from 45 relationships tested, 9%)			
Procedural Rationality	Gender Diversity (t)(c)	-ve	0.05
Frequency of Team Meetings	Gender Diversity (t)(r)	-ve	0.05
Reflexivity	Age Diversity (t)(r)	-ve	0.05
Psychological Safety	Gender Diversity (c)	-ve	0.05
Decision Belief: (n = 4 results from 27 relationships tested, 14%)			
Satisfaction	Nil	n/a	n/a
Confidence	Gender Diversity (t)(c) Educational Diversity (t)(r) Reflexivity (t)(c) Cognitive Diversity Post (t)(r)	-ve -ve -ve -ve	0.10 0.05 0.05 0.05
Effectiveness	Nil	n/a	n/a

There are 27 statistically significant results of a possible 177 core relationships tested.
(i) = individual level predictor; (t) = team level predictor; (c) = result from zero-order correlations; (r) = result from multiple regression analysis; (m) = result from multi-level modeling.

Two types of reporting processes may account for authors making more of sporadic findings than is appropriate. The first masks the fact that there are inconsistent findings by focusing the publication on a single type of demographic variation when in fact other demographic factors were investigated. An example is West & Anderson (1996), who collected demographic data on age, functional background, educational attainment, gender and tenure (personal communication, Neil Anderson, 1997) to predict innovation, but only reported the statistically supportive findings which occurred, as it happened, for tenure. Although it is difficult to prove definitively, it certainly appears from many of the works reviewed in Chapter 2 that such 'cherry-picking' is not uncommon (see for example, Krishnan et al., 1997; O'Reilly et al., 1993). Such a practice continues to promote 'upper echelons' theory, masking the fact that propositions in general are not supported. Indeed, in the current study, relationships with regard to tenure diversity for example, did not behave as expected whilst other types of diversity were more in line with prediction. Despite the admonition that more demographic variables should be studied simultaneously in order to tease out the differences with respect to demographic variation (Jackson et al. 1991), researchers still continue to limit the number of demographic variables studied or at least reported upon (see for example, Kilduff et al., 2000; Krishnan et al., 1997; Finkelstein & Hambrick, 1990; Finkelstein, 1992; O'Reilly et al., 1993).

The second reporting process concerns the issue of conceptual slippage. For example, an author may report statistically significant but opposing findings with respect to the different types of diversity investigated. (A case in point is the study by Smith et al., 1994 which reported a positive relationship between tenure diversity and performance but a negative relationship (contrary to prediction) between educational diversity and performance). Yet, reviewers of the field often gloss over such inconsistencies giving an over positive view. Thus, in relation to Smith et al.'s (1994) study, Kilduff et al. (2000) concluded that 'diversity is a predictor of organizational performance'. At one level this is obviously true, but, such blanket statements tend to suggest that all types of diversity are beneficial for performance, that there is no difference between them, and that the effect is in the direction expected. Therefore, belief in 'upper echelons' theory continues to persist, despite unconvincing evidence to support it.

With regard to this study, as suggested by Table 11.1 above, some predictors occur several times in relation to the dependent variables, whereas others occur only once, and some none at all. As Table 11.2 below shows, gender diversity is the most frequently occurring predictor (6 times) followed by frequency of meetings (4 times), functional diversity and educational diversity (3 times each), tenure diversity (twice), and age dissimilarity, educational dissimilarity, gender dissimilarity, tenure dissimilarity, age diversity, procedural rationality, reflexivity and psychological safety (once each).

Table 11.2 N of Predictors by Variable Group Studied

Demographic Dissimilarity					Demographic Diversity				
AGE	PROF	EDU	SEX	TEN	AGE	PROF	EDU	SEX	TEN
1	0	1	1	1	1	3	3	6	2
Total: = 4					Total: = 15				
Cognitive Diversity					**Team Processes**				
PRE	CO. PRE	CHG	POST	CO. POST	PRO	MTGS	REF	SAFE	
0	0	0	1	0	1	4	1	1	
Total: = 1					Total: = 7				

Prof = functional background; edu = education; sex = gender; ten = tenure; pre = cognitive diversity pre-discussion; co. pre = cognitive cohesion diversity pre-discussion; chg = cognitive change diversity; post = cognitive diversity post-discussion; co. post = cognitive cohesion post-discussion; pro = procedural rationality; Mtgs = frequency of meetings; ref = reflexivity; safe = psychological safety.

If one were to engage in the kind of selective reporting that is so characteristic of the 'upper echelons' literature, one might take the strongest predictor, gender diversity, and knit the results together as follows: Mixed gender teams show the least propensity to change their individual opinions and experience greater cognitive diversity post-discussion than single gender teams. Moreover, gender diverse teams tend to be more intuitive than analytical in their decision-making, meet less frequently and do not feel as psychologically safe. They show less confidence in their decision than all male teams. Such an interpretation would no doubt be exciting to most 'upper echelons' researchers owing to its coherence and its novelty in as much as gender diversity has never been studied in TMTs before.

Yet, with so many individual statistical tests in the study, there is a danger of making an error in either rejecting or retaining the 'upper echelons' hypothesis. A type I error occurs if the researcher incorrectly declares a relationship to be true due to chance producing the observed results. A type II error occurs if researcher retains the null hypothesis when it should be rejected in favour of the alternative. As has been pointed out, the number of statistically significant results across all the relationships tested in the study are only slightly better than chance. There is clearly a potential to make a type I error and retain the hypotheses predicated on 'upper echelons' theory.

The aim of the next section is to decide, on the basis of the analysis presented in the previous four chapters, whether to accept or reject 'upper echelons' theory as being appropriate for real TMTs. A review of the results suggests that the global null hypothesis should be retained, that is to say, there is no systemic effect of demographic variation on individual and team cognition, team processes and decision belief.

11.7 Bonferroni Adjustment or Cohen's Standard?

Overall 'study-wise alpha levels' (Becker, 2000), using adjustment procedures, are becoming increasingly common as a way to mitigate type I errors, especially in clinical trials (Perneger, 1998). One can imagine a situation, for example, in which a clinical trial testing a new drug (independent variable) is tested using performance indicators, such as blood sugar levels, heart rate, blood pressure, et cetera are tested. A statistically significant result may be found with one or two outcomes, but not with the others. In such a case, should the drug be declared effective or not? This is a similar situation as that in this study, where multiple demographic diversity predictors were tested against multiple dependent variables. Occasional relationships were observed. Should the hypothesis be accepted or rejected on this basis?

This issue is often addressed by using corrective procedures, such as a Bonferroni adjustment, which involves adjusting the significance level downwards so as to compensate for the increased probability of error when multiple tests are performed on the same dependent variable (Perneger, 1998). A simple Bonferroni adjustment applied to this dataset would mean dividing the required p level by the number of observations. Those relationships that would be kept or rejected if a Bonferroni adjustment were made are shown in Table 11.3. Only two relationships survive from the original 27 statistically significant results reported in Chapters 7 through 10. These are the relationships which find that: women are the least likely to change their minds; and frequency of team meetings encourages individuals to change their own personal opinions.

However, scholars disagree as to the efficacy of the Bonferroni adjustment, pointing to two major problems. The first is that such a correction is inappropriate for small datasets as the reduction in power is too severe (Simon, 2005). Indeed, in relation to the current dataset it might seem counter-productive to deliberately restrict the alpha level at the conclusion of the study as it was amplified in Chapter 9 in relation to the team level tests. However, as Bonferroni is the default correction procedure of choice by most statisticians, it is appropriate to discuss its application to the current study.

The second issue is that, as the correction is used specifically to reduce a type I error, of necessity it inflates the possibility of a type II error (see for example, Perneger, 1998; Sankoh, Huque & Dubey, 1997). Furthermore, considerable debate surrounds the issue of whether the correction procedure should be applied to datasets such as the one in the current study. The reasoning is that if the research question is concerned with demographic variation at a general level, then a Bonferroni correction should be used on all the relationships tested, but, where the specific relationships between types of demographic variation and outcome variables are the level of interest *"then Bonferroni should not be used"* (SISA, 2005, italics in original). With regard to the current study, each type of demographic variation was tested and reported separately, but, the hypotheses were at a global level. That is to say, discrete hypotheses were not made, say, about age diversity as opposed to gender diversity. Hence, if the dataset were larger, a Bonferroni adjustment would probably be the corrective procedure of choice.

A danger even with the Bonferroni adjustment is that by selectively accepting and rejecting results based solely on significance levels, the accepted results may still be "statistically significant but realistically meaningless" (Tabachnick & Fidell, 2001, p 52). That is to say, the significance levels do not assess the degree to which the independent and dependent variables are related. A more robust test is the strength of association or effect size (Cohen, 1988; Lipsey & Wilson, 1993; Tabachnick & Fidell, 2001). An effect size measures the strength of association between two variables regardless of sample size, and several different indices are available. Most are similar in that they selectively reduce the number of observed associations considered based on explicit criteria so that only important associations are kept (Becker, 2000). One of the most widely regarded is Cohen's standard or Cohen's *d* (Cohen, 1988). The application of Cohen's standard does not artificially reduce or inflate the potential to make type I or type II errors.

Cohen (1988) defined effect sizes as small (d =0.2), medium (d =0.5), and large (d =0.8), based on the R^2 association between the variables. Cohen's *d* (1988) is a straightforward criterion to apply to most of the relationships tested in this study, and is much more meaningful in terms of understanding the relationships observed than Bonferroni. The slight challenge arises for the multi-level modeling which does not use R^2. For these results, one needs to calculate a "pseudo R^2"(see Thoresen, Bradley, Bliese & Thoresen, 2004), for which there is no accepted norm. Moreover, as the name implies, it does not offer exactly the same interpretation of R^2 in ordinary least squares regression, but it is analogous and is computed in order to explain variance approximate to R^2 (Bateman, Jones, Nishikawa & Brouwer, 2000). It is computed manually as the reduction in the deviance (noted after the addition of predictors), divided by the deviance established in the null (variance components) model (Bateman et al., 2000). Thus it can be used as an index of effect size for the present purpose.

Applying Cohen's *d* (1988) to the 27 statistically significant relationships observed (based on R^2 and pseudo R^2), the biggest effect size is in actual fact, small, and relates to the relationship shown in Table 11.3 between gender diversity and psychological safety. The rest of the results show lesser effect sizes between $d = 0.01$ and $d = 0.19$, that is, below even the level defined as small.

The dilemma is choosing the level of acceptable effect size, which ideally should be specified in advance rather than judged retrospectively (Cohen, 1988). If one were to return to the drug testing analogy, a very small effect size (if defined as improvement in symptoms) may be attractive for easing the pain of terminally ill cancer patients. But in terms of most behavioral science research (and indeed, medical research too), larger effect sizes are desirable. In terms of practical implications for TMTs, where interventions are usually measured in terms of financial impact on the bottom line, a minimum of a medium effect size would be desirable.

A comparison of Bonferroni and Cohen's effect size concerning the current results is shown in Table 11.3. If one accepts Bonferroni as the method of choice, two relationships remain. If one accepts that Cohen's standard is more appropriate, only one relationship remains. Moreover, the marginal and small effect sizes in relation to this study mean that at least 26 of the 27 statistically significant relationships observed, are trivial in real TMTs. Even then, the small effect size of gender diversity predicting less psychological safety is far from convincing as a stand alone result for a single aspect of demographic variation. Interestingly, the latter result, despite having the largest effect size in the study, would be rejected by a Bonferroni adjustment as not achieving enough significant statistical power.

Perhaps the difference between the two methods is best illustrated as a practical implication arising from the research. The statistically significant negative relationship between frequency of team meetings and cognitive dissimilarity pre-discussion (H6), which was interpreted in Chapter 9 to mean that the more often a team meets the less cognitive dissimilarity it will experience, achieves a Cohen's *d* of 0.01, but is no longer statistically significant after a Bonferroni adjustment. If, on the basis of initial statistical significance alone (consistent with most reported 'upper echelons' findings), one were to offer the prescription to TMTs that meeting more often is beneficial for reducing cognitive dissimilarity amongst members, TMTs are unlikely to achieve a practical benefit owing to the fact that the effect size is so small.

Having sought to overcome the numerous methodological constraints and difficulties associated with the sheer complexity of studying TMT demographic variation, the determination that the effect sizes are almost without exception, trivial, could be disappointing. However, from the outset, this study has been careful to remain centered on real TMTs, thus the effect size in terms of practical and theoretical implications is more in keeping with its ontology, as opposed to tests of statistical significance alone. Therefore, this study adopts Cohen's standard, and thus retains the global null hypothesis (Becker, 2000) in rejecting 'upper echelons' theory. This is undoubtedly a controversial stance to take, especially as research into 'upper echelons' theory is currently flourishing (Hodgkinson & Sparrow, 2002; Carpenter et al., 2004). The practical and theoretical implications for TMTs will be addressed in next.

Table 11.3 Summary of Results Accepted or Retained After Correction

Dependent Variables	Independent Variables	*Original p*	*Accept or Reject Bonferroni*	R^2 */Psuedo-* R^2	*Accept or Reject Cohen's d*
Cognitive Variation at the Individual Level:					
Cognitive Dissimilarity Pre-Discussion	Frequency of Mtgs. (t)(m)	0.05	×	0.01	<0.20 ×
Cognitive Cohesion Dissim. Pre-Disc.	Age Dissimilarity (i)(c) Tenure Dissimilarity (i)(m) Tenure Diversity (t)(m) Procedural Rat. (t)(m)	0.05 0.05 0.05 0.05	× × × ×	0.02 0.03 0.01 0.01	<0.20 × <0.20 × <0.20 × <0.20 ×
Cognitive Change Dissimilarity	Gender Dissimilarity (i)(r) Functional Diversity (t)(m) Education Diversity (t)(m) Gender Diversity (t)(m) Frequency of Mtgs. (t)(m)	0.01 0.05 0.05 0.05 0.01	✓ × × × ✓	0.01 0.04 0.02 0.04 0.05	<0.20 × <0.20 × <0.20 × <0.20 × <0.20 ×
Cognitive Dissimilarity Post-Discussion	Education. Dissim. (i)(r) Education. Diversity (t)(m) Frequency of Mtgs. (t)(r)	0.01 0.05 0.05	× × ×	0.08 0.02 0.01	<0.20 × <0.20 × <0.20 ×
Cognitive Cohesion Dissim. Post-Disc.	Nil	n/a	n/a	n/a	n/a
Cognitive Variation at the Team Level:					
Cognitive Diversity Pre-Discussion	Nil	n/a	n/a	n/a	n/a
Cognitive Cohesion Diversity Pre-Disc	Nil	n/a	n/a	n/a	n/a
Cognitive Change Diversity	Functional Diversity (t)(r) Psychological Safety (t)(r)	0.05 0.10	× ×	0.13 0.09	<0.20 × <0.20 ×
Cognitive Diversity Post-Discussion	Functional Diversity (t)(r) Gender Diversity (t)(c) Frequency of Mtgs. (t)(r)	0.10 0.05 0.05	× × ×	0.09 0.14 0.10	<0.20 × <0.20 × <0.20 ×
Cognitive Cohesion Diversity Post-Disc.	Tenure Diversity (t)(r)	0.05	×	0.13	<0.20 ×
Team Processes:					
Procedural Rationality	Gender Diversity (t) (c)	0.05	×	0.05	<0.20 ×
Frequency of Team Meetings	Gender Diversity (t)(r)	0.05	×	0.09	<0.20 ×
Reflexivity	Age Diversity (t)(r)	0.05	×	0.19	<0.20 ×
Psychological Safety	Gender Diversity (t) (c)	0.05	×	0.20	0.20✓
Decision Belief:					
Satisfaction	Nil	n/a	n/a	n/a	n/a
Confidence	Gender Diversity (t)(c) Educational Diversity (t) (r) Reflexivity (t) (c) Cognitive Diversity (t) (r)	0.10 0.05 0.05 0.05	× × × ×	0.11 0.13 0.16 0.19	<0.20 × <0.20 × <0.20 × <0.20 ×
Effectiveness	Nil	n/a	n/a	n/a	n/a

(i) = individual level predictor; (t) = team level predictor; (c) = result from zero-order correlations; (r) = result from multiple regression analysis; (m) = result from multi-level modeling. Small effect size d = 0.20; medium effect size d = 0.50; large effect size d = 0.80. ✓ = accept, × = reject.

11.8 Practical and Theoretical Implications of Retaining the Null Hypothesis

The discussion thus far has established that the statistically significant findings in this research may be largely attributable to chance alone, and are certainly based on small effects. It is of interest then, to consider how previous 'upper echelons' research has dealt with such a situation. Only one study reviewed in Chapter 2 has had the courage to report no support for 'upper echelons' theory, that conducted by West & Schwenk (1996). In their case, there were no statistically significant relationships with demographic variation, at all, and even the sub-title of their paper, "A report of resounding non-findings" (West & Schwenk, 1996) attests to their reasons to retain the null hypothesis.

West & Schwenk's (1996) study was discounted in Chapter 2 for incorrectly confounding all the types of diversity together, and then trying to measure them simultaneously. The reasoning was that there were no results as a matter of course, due to incorrect statistical techniques being applied. Indeed, they believed that there was a measurement issue: *"the complete nonsignificance of the regression results suggests strongly that the dependent variables were inadequately measured"* (West & Schwenk, 1996, p 574). Revisiting their paper in the light of the foregoing suggests that they may not have found any results even if they had correctly measured demographic variation, and addressed their self-identified problem with the dependent variables (global TMT consensus).

Furthermore, if one were to determine effect sizes based on R^2 statistics in West & Schwenk's (1996) study, they were, like this study, tiny. Notwithstanding the problems regarding their paper, it is of interest to note now, that there is a precedent in the demographic variation literature for retaining the global null hypothesis. As in the current study, West & Schwenk (1996), made a rare attempt to depart from sole reliance on public archival data and get closer to real TMTs (in their case through survey of at least three TMT members per team). O'Reilly, et al. (1993) also tried to get closer to real TMTs by conducting a repeated interview study with CEOs of 24 firms over an 18-month period. Although they only studied tenure diversity, the R^2 statistics they report range from .00 to .23. In terms of Cohen's *d*, three of their 10 results reported achieve a small effect size, the other seven do not. None were medium or large effects. It is also of interest to note the effect sizes arising from studies that rely on public archival data. Jackson et al. (1991) distinguished between types of diversity and types of dissimilarity in relation to predicting turnover. A review of the R^2 statistics they report finds a range of .00 to 0.35 for team level diversity indices, and .06 to .22 for dissimilarity indices. Similar to this study and those already cited, the effect sizes compared using Cohen's *d*, reveal that almost all are small or non-existent.

Given the ubiquitous persistence of 'upper echelons' theory, an obvious question arises as to whether the results and effect sizes are stronger in artificial or synthetic teams? That is to say, have previous supporting results been found in other types of TMT, especially those deliberately excluded from the review in Chapter 2? This does not appear to be the case. For example, Kilduff et al. (2000) conducted a business simulation with 35 synthetic TMTs made up of members on an executive training course. They found no statistically significant correlations, and *"there were no significant relationships between the demographic diversity variables and cognitive diversity in any of the regressions"* (Kilduff et al. 2000, p 27). Other studies often assume that demographic variation is a correlate or antecedent of cognitive variation, but do not actually measure it in synthetic teams (see for example, Mohammed & Ringseis, 2001; Heath & Gonzalez, 1995). There does not appear to be greater support for 'upper echelons' theory arising from synthetic teams, and the discussion above suggests that in authentic TMTs, statistically significant results are few, and effect sizes, are, almost without exception extremely small.

This has practical implications for TMTs, particularly where 'upper echelons' credo still persists. Take for example, the recent advice for TMTs to re-invent themselves to become more demographically diverse by appointing to their ranks more dissimilar persons (Carson et al., 2004). The premise for such a re-invention is, according to Carson et al. (2004), that organizational performance benefits accrue to demographically diverse teams, primarily through improved decision-making, based upon the classic 'upper echelons' tradition (Hambrick & Mason, 1984). As discussed in Chapter 2, researchers in related fields may not entirely agree with such prescriptions (see Adler, 1997). The point being made here however, is that the most recent TMT research is still persisting with the 'upper echelons' and 'value-in-diversity credo'.

Indubitably, such a move would be a serious intervention, one likely to cause considerable upheaval to the TMT and the organization as a whole. It may serve the purpose of furthering 'Equal Opportunities' policies. However, there is little empirical evidence to suggest a performance benefit. Moreover, Carson et al.'s (2004) contribution is one of a rash of similar prescriptions (see also Dewett, 2004; and Edmondson et al., 2003) that have drawn on the dominant position that has so long been held by Hambrick & Mason (1984). Given the inconsistencies with the research literature to date, and that 'upper echelons' theory has never really had much empirical support, together with this research which has systematically and meticulously sought to establish (albeit unsuccessfully) the required links to underpin it, such calls for radical reform to TMTs are premature at best, and foolish at worst. It would, however, seem eminently sensible to argue for future 'upper echelons' research to pay more attention to effect sizes in order to determine the strength of association between demographic variation and other study variables, and to report these in published works.

Another way of addressing the issue raised in this study, is to ask the question: Why would demographic differences have no systemic effect on team and individual cognition and team processes? Moreover, why is this observation the same in real and synthetic teams?

With respect to real teams, the favoured explanation offered here is that it is highly likely that demographic variation in TMTs gives way over time to familiarity. That is to say, individual differences based on attributes such as age, functional background, education, gender and tenure, only make a difference when a team is first formed or, at the individual level, when a new member joins the team. After a while, the novelty of demographic difference, for example, of appointing a young, female HR director to a TMT of all male engineers, wears off. The person becomes accepted in her own right, and blends into the rest of the team. This is likely to be the case in relation to all aspects of demographic variation. This scenario finds resonance in Wiersma & Bantel's (1992) study, reviewed in Chapter 2, in as much as the positive effects of demographic variation that they observed tended to taper off over time. With respect to synthetic teams, as was observed earlier, these are, in the main, made up of student cohorts who are in the process of studying for an MBA together. It is probable that the same familiarity posited to affect real teams is responsible for the non-supportive findings in synthetic teams. The students know each other and also work together on other projects in other subjects. The effects of demographic variation are not likely to be apparent in groups of people who are familiar with one another.

With regard to the issue concerning team longevity, as implied in Chapter 2 (p 33), this may be a particular facet of 'upper echelons' research that is as yet, untapped. Several studies have found that the effects of demographic variation differ over time (see for example, Murray, 1989; Hambrick et al., 1996). Indeed, Chapter 2 pointed to the discernible pattern across several studies in which demographic variation (particularly diversity) had a negative influence on short term performance, but appeared to be ultimately beneficial in the long term.

In order to test the validity of the argument just put forward concerning familiarity, two further types of analysis were conducted on the current dataset. In the first instance, an individual level approach was taken which involved conducting moderated regression analyses for all five of the cognitive dissimilarity dependent variables using the five demographic dissimilarity variables as predictors and tenure demography as the moderator (tenure in the TMT in months). Only one of 25 the moderated regression analyses showed an effect for the multiplicative term that was approaching statistical significance (that between functional diversity and cognitive change dissimilarity). This sole statistically significant effect could have arisen purely by chance. This means, that although the familiarity argument is commonsensical in its appeal, there is no evidence to support it (with regard to tenure in particular) from the current study.

The second set of analyses to test the familiarity argument focused on tenure at the team level as a moderator of the effect of the individual demographic dissimilarity variables on the cognitive dissimilarity variables. To this end, the multi-level models described earlier were re-run, with mean team tenure instead of the team diversity or team process variables. Not one analysis out of 20 (i.e. the 4 outcome cognitive dissimilarity variables[1] and five demographic dissimilarity variables) showed an interaction effect (the interaction between x and y predicting z). Therefore, the familiarity argument put forward at the beginning of this section, although plausible, is not supported by the proxy tests just applied. The application of these tests is somewhat of a double-edged sword. One the one hand, it could be argued that these tests are the obvious addendum to the familiarity argument. On the other hand, it could of course be argued that using tenure as a proxy of

---

[1] As noted earlier in Chapters 8 and 9, the final cognitive dissimilarity variable, cognitive cohesion dissimilarity post-discussion was not modelled owing to the minute amount of variance at the team level.

familiarity presents exactly the same issues as the general criticism of previous 'upper echelons' research as presented in Chapters 2 and 3! The point here is that given the question posed above as to why demographic attributes make no real difference to behaviour and cognition, alternative explanations must be sought. Outside-decision domain familiarity is one such alternative, but in light of the regression analyses testing the impact of tenure on cognitive dissimilarity, it would seem that this may have been a red-herring. Of course, as discussed in Chapter 2, and for the reasons outlined above, familiarity is a feature of interest in 'upper echelons' research, with studies being designed in such a way as to compare recently formed teams subsequent to mergers and acquisitions (i.e. those with little outside-decision domain familiarity) with mature teams (see Krishnan et al. 1997 for an example).

It is common practice in studies in which the results have been disappointing to consider the contribution of measurement issues. As has been repeatedly drawn attention to in this study, the measures used throughout were robust and thorough. In several cases bespoke or less common measures were used in order to tap the constructs. For example, the binary Euclidean distance measure was used to compute gender dissimilarity. A distance measure was created for functional background dissimilarity and for educational attainment dissimilarity. The Stride Index was used to compute diversity for continuous demographic data, an index that allowed direct comparison with Blau's (1977) index for categorical data. Moreover, unlike most other studies in this genre, the extent to which heterogeneity was evident in the data was assessed using ICCs before applying diversity indices. Rather than rely on self report questionnaires, the team processes were tapped using a bespoke coding guide, which was used by independent raters, ICCs again being used to validate the coding guide. Sophisticated multiple regression analyses were used in the multi-level modeling of the data and a Pseudo-R^2 was calculated in order to compare results across the study. Finally, correction procedures were applied to the results arising from the data. In short, accurate and robust measurement was a consistent feature of this study. Hence one cannot point easily to errors of measurement as being the cause of the paucity of results. Of course, it could be argued that in general terms, the dataset was small, although in real terms, 23 whole teams is a considerable achievement. That there were any results at all, attests to the fact that the dataset was adequate for the purpose of testing the hypotheses.

Having established that the 'how' of measurement was addressed properly, the next logical question concerns the 'what'. That is to say, were there unmeasured factors other than demographics that were affecting cognitive variation, team processes and decision belief? This latter question has recently given considerable pause for thought to 'upper echelons' researchers (Carpenter et al., 2004) so that the latest 'buzz-word' in this research arena is 'antecedents'.

It is worth re-stating at this point that 'upper echelons' theory, essentially is that demographics act as proxies for attitudes, beliefs, experience, thoughts and ways of thinking (see Hambrick & Mason, 1984, Hambrick et al., 1996). This often tends to get lost in discussions within and about 'upper echelons' research, so that discussion of any results tends to focus on the link simply between demographics and outcome, rather than the attitudes or beliefs supposedly tapped by demographic proxies that may account for observed relationships. This loose coupling has not escaped the notice of the critics however (Priem et al., 1999). Perhaps in response, there have been calls for 'upper echelons' researchers to acknowledge the broad array of antecedents other than demographics that influence TMTs (Hodgkinson & Sparrow, 2002), and for researchers to be much more specific about which antecedents they are testing (Carpenter et al., 2004). With regard to the current study for example, it was inherently assumed that demographic proxies would manifest themselves in different rankings of candidates. Transcending other studies in its design, this research actually tested relationships between demographics and rankings. However, by way of contrast with the antecedent argument, the study did not for example, specifically test psychographic factors such as attitudes. Talking in parentheses, it is interesting that the candidate consistently ranked as number 1 in the study, was the only external candidate (See Chapter 6), the other three candidates were seeking internal promotion. An antecedent study could have specifically tested attitudes to internal vs external appointment. As it was, the study took the classic 'upper echelons' line that demographics act as a broad proxy for non-specific and general attitudes, and thus it suffers from the inherent limitations, characteristic of TMT research, that such proxies present (Priem et al., 1999).

Another aspect of the antecedents argument which to date has not received as much currency as it probably deserves, is the idea of individuals influencing and persuading other individuals within the team towards a consensus. Chattopadhyay et al. (1999) for example, investigated the extent to which a variety of functional backgrounds and social factors in the TMT influenced each executive's beliefs within that team (as discussed in Chapter 2).

Clearly, the inclusion of antecedents offers much to the 'upper echelons' researcher. However, a word of caution is in order. The drive to include antecedents previously untapped by traditional 'upper echelons' research may simply be a case of rehashing well rehearsed claims and methods. This is because the current antecedents argument still promotes (see Carpenter et al., 2004), and uses (see Chattopadhyay et al., 1999), demographics as proxies. To take a case in point, in Chattopadyhay et al.'s (1999) study of focal executives, the proxy was functional background. That is to say, the relationship between the functional backgrounds represented on the team compared to that of the individual (focal executive), and the extent to which the individual's responses matched those of the team was tested using similar diversity indices to those used in this study. Chattopadyhay et al.'s (1999) study relied on secondary archival data, and questions as to how or why individuals' functional background would influence another person were not explored. Their study yielded few results, and concluded that social influence (communication, socialization, and social information processing) rather than functional conditioning (current and prior job position, feedback and rewards linked to experiences) was more important in shaping an individual's beliefs and propensity to agree with other members of the team. Furthermore, if one were to extrapolate on the example given above with regard to the current dataset, following Chattopadyhay et al. (1999), one would use organizational tenure as a proxy for the antecedent attitude to internal versus external promotion. There is clearly a tension between incrementally extending 'upper echelons' theory to include yet more features of interest by using demographics as proxies for antecedents, and focusing on antecedents to the exclusion of demographics, thereby, negating 'upper echelons' altogether. It remains to be seen whether the current treatment of antecedents will satisfy critics of the field, or whether a new generation of researchers will find ways to meaningfully and practically include the study of antecedents within 'upper echelons' research.

It is highly likely that external factors and situation specific factors could give rise to shared beliefs amongst TMTs that have not been covered either in the current research, or indeed in many other studies within the 'upper echelons' field. These could include (but are not limited to); industry blind spots, environmental turbulence, environmental munificence, decentralization of administration functions, organizational size and organizational effectiveness (Hodgkinson & Sparrow, 2002). With rare exception (see Chattopadyhay et al. 1999), such antecedents are hardly ever controlled for. Indeed, as previously mentioned, it is rare in 'upper echelons' research for TMT size to be controlled for (Carpenter et al., 2004), even though this would seem a fundamental element to understanding the impact of team processes. What the discussion in this section has highlighted, is that there are a myriad of factors and nuanced processes that may be influencing TMT decision-making which: (a) are not necessarily immediately apparent; (b) remain unmeasured; and (c) potentially mask the contribution of demographic factors to cognition, process and decision belief.

This chapter returns to a discussion of the theoretical implications of the study after reviewing the limitations and strengths, some of which have already been addressed throughout this discussion.

11. 9 Strengths and Limitations of the Current Study

Driving this thesis has been a fundamental desire to study authentic top management teams in real time. The research stayed true to this overarching requirement achieving a sample size comparable to interview based studies which range from 20 (Hambrick, 1981) to 24 (O'Reilly et al., 1993). The sample of 130 senior executives in 23 top management teams in this thesis is a respectable size considering that the study is with whole, real teams, studied intensively, over the life-cycle of making a realistic decision.

Within the context of real TMTs, the thesis had two specific aims. The first was to investigate the relationship between demographic and cognitive variation at the individual and team levels of analysis. This aim arose from a critique of the TMT demographic variation literature which showed that dissimilarity (individual) and diversity (team), though often treated as equivalent in the literature, should be treated as distinct constructs. One of the real strengths of the thesis is the fact that it focused very much on the TMT literature. This served to distil what is actually known about TMTs, which after all, is the point of 'upper echelons' theory (Hambrick & Mason, 1984). Much prior research has made links either to other levels in the organisation, the wider diversity literature, or the broad management literature in general. Such an approach has caused a muddle with regard to what applies to TMTs, and more fundamentally, what constitutes 'upper echelons' research and what is ordinary relational demography research as different findings arise from different types of groups (Williams & O'Reilly, 1998). Other research has tried to stay at the 'upper echelons' level, but has confounded the differences between TMTs and boards of directors (Flatt, 1996). By maintaining the spotlight on TMTs exclusively, this research is very much in line with current and anticipated directions for 'upper echelons' research (Carpenter, 2004).

To test the implicitly hypothesised link in 'upper echelons' theory that demographic variation predicts cognitive variation required that both cognitive dissimilarity and cognitive diversity were defined and directly measured, even though this has virtually never been done in previous TMT studies. For the purpose of this study, cognitive dissimilarity was defined and measured as the distances between individuals' rankings of candidates in an executive selection process; whereas diversity was a team level construct measured as the degree of variation.

Moreover, in meeting this aim, the research sought to overcome previously observed inadequacies in traditional decision-making research; failure to recognise the way in which individual decision choices change over time; and the neglect of the extent to which individuals' preferences match that of the team consensus.

A particularly important part of this research was to transcend two levels (individual and team) through the use of multi-level modelling. There is only one TMT study to date that has used a multi-level model, a study concerning demographic variation and turnover (Boone et al., 2004). Hence the current research is unique in having applied a multi-level model to understand the relationship between individual and team demographic factors, and individual level outcomes concerning cognitive dissimilarity. It is relevant to note that the greatest number of positive findings (13 from a possible 70 core effects tested) come from the tests that multi-level modelling uniquely allows, namely the effects of demographic dissimilarity, diversity and team processes on individual cognitive variation, which attests to the value of using such a tool in this type of research.

The second aim was to investigate team processes directly and to measure decision belief. It was predicted that procedural rationality and frequency of team meetings would lessen dissenting views, thereby enhancing consensus amongst TMTs and positively influencing decision belief. Two further process variables, reflexivity and psychological safety were predicted to allow greater freedom of expression and hence more cognitive disagreement within teams, which it was thought would lessen decision belief. In attempting to meet the aim of directly investigating processes, the research presented in this thesis went beyond conventional demographic variation research which uses demographic attributes as proxies for understanding process (Priem et al., 1999).

The potential limitations of the study have been referred at appropriate points throughout the thesis (see for example, the section on internal and external validity threats in Chapter 6), however they can be summarised as follows:

1. non-traditional outcome measures;
2. use of simulation as opposed to real decisions germane to organisations; and
3. the 'decision' in the study was not 'strategic' in the sense of the strategic cognition literature.

In Chapter 2, it was observed that most of the outcome measures in this literature are financial measures of performance, despite these being problematic (Murray, 1989). However, some studies, particularly those that aspire to understand cognitive variation (e.g. Glick et al., 1993; Chattopadhyay et al., 1999), use measures other than aspects of financial performance, such as strategic consensus. It was decided in this study that it was better to similarly concentrate on the processual issues rather than simply seek to validate 'upper echelons' theory by investigating yet another outcome variable. Whilst this approach is advocated by some scholars in the field, they also observe the difficulties in publishing studies that do not fit the traditional 'upper echelons' mode (Carpenter et al., 2004). With the benefit of hindsight, it would have been better to include a financial measure of performance (e.g. Smith et al., 1994; Knight et al., 1999). At the very least, this would have provided a richer context for the non-findings of the study with regard to cognitive variation, team processes and decision belief, especially if relationships had been observed between demographic variation and financial outcomes. This would have been even stronger evidence for Pfeffer's(1983) argument that the study of process is not necessary. If on the other hand no relationships had been observed between demographic variation and performance, then it would mean that 'upper echelons' theory should most definitely be rejected. It would be silly to hypothesise what the results may have been had such a measure been included. The point here, is that if a financial performance measure had been included, any observed relationships with that variable would have given more clarity with regard to the theoretical implications of the study.

In Chapter 6, it was argued that the simulation did not unduly influence the decision-making process under investigation, nor did it pose substantial threats to the internal or external validity of the study. Indeed, as commented on in Chapter 6, the simulation was so realistic that many TMTs were convinced that the fictitious company had been based upon their own organisation. It was noted above that the use of the simulation provided a precision which is one of the greatest strengths of the study. However, to purists, a simulation by its very nature compromises authenticity. Whilst the ideal would have been to observe real decisions in action (Pettigrew, 1992), the study goes a long way in terms of method and the development of measures which would be profoundly necessary in the observation of real decisions.

The simulation provided the ability to measure rankings of participants, which was a way of tapping their cognition. Thus it acted as a proxy for the fact that individual's attitudes and perceptions about candidates would be different. Hence the research went well beyond previous research in this area by actually measuring cognition, but did not specifically measure attitudes or beliefs. No pretence or assertion was made with regard to the findings being related to particular attitudes. For example, the HRM literature suggests that TMTs tend to select persons to their ranks that are similar to themselves (Jackson et al., 1991), and have a preferential attitude to candidates that are physically similar etc. The research did not specifically attempt to tap these kinds of constructs, and so is open to criticism that differences between cognitions are still obscure. However, in keeping with the research aims, the study did monitor the extent to which viewpoints were maintained or changed, and the proximity to the team consensus pre- and post-discussion and related these to demographic variation, a key tenet of 'upper echelons'.

Despite any criticisms that may occur due to the use of a simulation, what should be remembered is that the teams were real, and the discussions they had were real. The simulation provided the content for the discussion, but the behaviour in the teams was authentic and true. The real problem occurs when researchers try to study processes in artificial teams, an impossibility according to some (see Williams & O'Reilly, 1998).

In Chapter 3 (and as discussed above), it was noted that there are a variety of factors that can affect a TMT judgment, some are situationally specific, that is they relate only to the organisation involved, whereas others can be related to the industry or indeed the domestic or international economy (Adler, 1995). As discussed above, with reference to Chattopadhyay et al. (1999), Carpenter et al. (2004), Lau & Murnighan (2005), factors such as environmental turbulence, munificence, communication, social integration and demographic faultlines can all mean that teams have an ingrained and perhaps unconscious shared attitude that precedes the research.

The current study (like the majority of 'upper echelons' studies) did not control for such factors. Indeed, it took a typical small group research approach, that ceteris paribus, all TMTs had the same starting point. One of the strengths of the current study is that it investigated teams across organisations, which are representative of companies in the UK manufacturing sector, made up of five industries. Both Pettigrew (1992) and Lawrence (1997) observe that many studies of TMTs are within a particular company or within a very narrow sector. One of the major reasons such controls were not included of course, is that there was no organisational performance measure such as ROA or ROI, ROCE etc. Given the retrospective observation above concerning the realism of the simulation, it may have been possible to design the study to include some of these control measures. Indeed, these could have been specifically linked to the simulation exercise of appointing a new TMT member, such as organisation size, team turnover, market share, social integration etc.

In terms of processes, many could have been selected. Indeed, as discussed in Chapter 3, previous 'upper echelons' studies have suggested various processes such as receptivity to change, risk taking, creative innovation, diversity in information search etc. Similarly, the four chosen in this study had a theoretical or research precedent for application to TMTs. Whilst there could be debate as to the salience of these or other processes not investigated, one of the real strengths of this study is that it has at least made a start by observing processes in vivo which have only previously either been completely ignored in deference to demographic proxies, or have been captured retrospectively by self-report questionnaire. There is clearly much more work yet to be done in the field to understand the processes that are in the 'black box' (Pettigrew 1993; Lawrence 1997).

The next section seeks to position the theoretical contribution of the study and reflect this in a revised presentation of the research model first offered in Chapter 4.

11.10 The Research Model Re-visited: Implications for 'Upper Echelons' Theory

Although it is evidently appropriate to retain a global null hypothesis in this study given the unconvincing number of statistically significant results and more important, the tiny effect sizes of the predictors in relation to the dependent variables, one could be guilty of tossing out the proverbial baby with the bath water. The fact that there are results, even if these are small and inconsistent, requires further investigation. Indeed, Sparrow (1994, p 158) argues that "*the presence of any statistically significant relationship between specific demographic aspects of top teams (who frequently represent 0.1% of total organizational membership) and aspects of the organization's performance is remarkable*".

In concluding that the global null hypothesis should be retained, the question arises as to the theoretical contribution of the study. In Chapter 4 the research hypotheses were presented in a schematic illustrating the various relationships hypothesized. It is appropriate now to return to this schematic in light of the findings and non-findings. It is also appropriate at this juncture to re-visit the Lawrence (1997) critique of 'upper echelons' theory.

Lawrence (1997) is most often given superficial, cursory citation in the 'upper echelons' literature as arguing for the inclusion of process in 'upper echelons' research. Indeed, it has been latterly observed that in the self-serving interests of getting published, many researchers include a paragraph lamenting the lack of process studies, even when they do not research process (Carpenter et al., 2004).

Whilst it is true that Lawrence (1997) challenged the congruence assumption that demographic variation leads to organizational performance as espoused by 'upper echelons' (as discussed in Chapter 2), her analysis of the underlying tenets went much deeper. Lawrence's (1997) primary concern was the interdependence between theory and method. She argued that as the congruence assumption was based solely on the convenience of method (i.e. using public archival data), the basic theory was weak. Similarly, Pettigrew (1992) believed that new questions and methods needed to emerge to guide and complement research and theory if 'upper echelons' was to avoid being simply "a triumph of method over substance" (Pettigrew, 1992, p 174).

Lawrence (1997) classified three possible approaches to studying 'upper echelons'. The first, as mentioned above is the classic congruence assumption, that is to say, it treats demographic variables as predictors in instrumental theory (i.e. demography leads to outcome). Almost all research in the 'upper echelons' tradition could be classified as using this approach (Lawrence, 1997; Pettigrew, 1992; Carpenter et al., 2004).

In order to 'prove', instrumental theory Lawrence (1997) argued that studies would need to consistently find the same results. Take for example, the widely held prediction that tenure predicts turnover. If the fundamental theory is correct, this predictor should be proved over and over again in various studies, and should explain a comparable amount of variance in studies of similar units. Lawrence (1997) tested her criteria by reviewing the studies to date for the number of findings and their significance levels compared with the number of relationships tested, and found that the congruence assumption should be rejected in about 60% of predictions. Lawrence (1997) conceded that some of these non-findings may be clarified in future by researchers being precise about the level being studied (i.e. dissimilarity or diversity), but that overall, there was not convincing support for the congruence assumption that demographic factors are instrumental in predicting organisational performance.

The second approach put forward by Lawrence (1997) was to view processes as predictors in an explanatory theory. That is to say, demographic variables are indicators of a process that explains a particular outcome. So for example, there may be a negative relationship between demographic variation and communication, which leads to conflict (Lawrence, 1997). The criteria she established for acceptance of indicator theory concerned statistical reliability and validity (high proportions of variance explained) which she then applied to highly regarded research in the 'upper echelons' field. Lawrence's (1997) review of studies showed that the results across studies do not consistently meet the reliability and validity criteria. Moreover, she noted that results differed depending upon which demographic indicator was used, and lamented a plethora of single demographic indicators being used. Lawrence (1997) also criticised the common practice amongst researchers of not exploring why a particular demographic indicator might be linked to a particular process. "*Thus enormous interpretative leaps are made from distant demographic surrogates of team characteristics such as homogeneity and heterogeneity, through unobserved and remote intervening processes such as information processing, conflict resolution and problem solving, to outcome variables such as team effectiveness or organisational performance*" (Pettigrew, 1992, p 176).

The third methodological approach espoused by Lawrence (1997) was to investigate intervening or mediating processes. From this theoretical perspective, processes should be related both to demography and to outcomes. This means, when illustrated by the current study, an intervening process would be one that when entered to a regression equation where a relationship had already been noted between demographic variation and decision belief, the process variable would account for the variation and the original relationship would disappear. This did not happen with regard to the current study (See Chapter 9). Hence it can be said that the relationships observed between demographic variation and decision belief are not mediated by the processes investigated (frequency of meetings, procedural rationality, reflexivity and psychological safety).

With regard to intervening process explanations, Lawrence (1997) argued that "*because the final test of an intervening process explanation requires measuring the subjective concept, it is not possible to provide evaluation criteria for such explanations under the congruence assumption. However, intervening process explanations do not rule out alternate interpretations or situations that might explain a null result as do instrumental theories*" (Lawrence, 1997 p 10).

Clearly, it was assumed that the most promise for developing theory and research into 'upper echelons' was to investigate the intervening processes. The schematic of the research model in Figure 4.1 was responsive to all three theoretical approaches. First, notwithstanding the lack of a financial performance measure, the congruence assumption was tested with regard to demographic and cognitive variation (H1- H3), and with regard to demographic variation and decision belief (H8). Similar to Lawrence's (1997) review of the available literature, and as shown by the discussion of 'study-wise alpha levels' (Becker, 2000) above, this study does not find convincing support for the congruence assumption. Second, the indicator process theoretical approach was tested (H4 – H6), but as discussed above, the variance explained was not sufficiently high in any of the relationships tested to endorse this perspective. Third, the intervening process model was tested (H7 and H9) without success. This may be due to the wrong processes being selected for investigation in this study. The burgeoning interest in 'upper echelons' research (Hodgkinson & Sparrow, 2002) and the rich findings of the studies considered in Chapters 2 and 3 mean that 'upper echelons' is not likely to be superseded as the theory of choice in TMT research for some time to come. As this study did not include an organisational performance measure, one could argue that it was not a complete test of 'upper echelons' theory. However, as established in Chapter 2, unless the relationships between demographic variation and intervening processes can be determined, the relationships between demographic variation and outcomes cannot be attributed to process. Hence, any theoretical contribution arising from the study needs to extend rather than replace 'upper echelons'. The contribution of this study to theory is likely to have descriptive relevance (Priem et al., 1999). That is to say, it describes the day-to-day reality of diversity encountered by TMTs, which is an important societal value (Williams & O'Reilly, 1998), but, it does not encourage TMTs to include demographic composition as part of their decision-making.

That there were so few relationships observed with regard to process could mean that (a) the processes investigated are not going on in TMTs; or (b) these particular processes are not important for cognition and decision belief (the two outcome measures hypothesised to relate to process). An argument as to both is somewhat plausible. As to the first, Table 11.2 shows that team processes were only predictors in seven instances in the analysis. These seven were from a possible 55 relationships tested in which process could have been a predictor. Specifically, frequency of team meetings occurs four times, whilst procedural rationality, reflexivity and psychological safety only occur once each. Notwithstanding the caveats in this discussion with regard to chance occurrences, that there are any relationships with process at all suggests that there is some evidence of their existence in TMTs, but that this is not very strong.

With regard to the second explanation, the team processes which were examined are not stable predictors of cognition or decision belief, even the influence of frequency of meetings (the most occurring predictor) is not particularly great. Turning attention particularly to decision belief (as this is the outcome of interest for hypotheses 7, 8 and 9), one notes that of the three variables, satisfaction, confidence and effectiveness, statistically significant associations are observed only with respect to confidence. This would suggest that perceived satisfaction with the decision-process and perceived effectiveness as a team, are concepts which are on the periphery of strategic decision-makers' thinking. From a practical perspective influenced by experience at this level in organisations, one may argue that executives tend to focus on performance related activities, often working within demanding time constraints. Such pressures leave little time or mental space for reflection about whether one is satisfied with the process, or whether the team is working effectively or not. Many times during the conduct of this research, the author was told by TMTs that they simply did not believe such considerations to be relevant, indeed, this kind of reflection was deemed by many to be 'self-indulgent navel gazing'. The main priority they said, was to make the best decision possible, and then 'get on with it'.

In terms of theoretical extension with regard to processes, it is clear from this study that the effects of process are much more subtle than previously thought within the TMT research field. What is important is that the current study 'has helped to sort out what could be considered to be spurious or less influential processes from those which are substantive with regard to TMT decision-making' (see Priem et al., 1999 p 949-950).

In the revised model in Figure 11.1, team processes and decision belief are grouped together with the discussion, as it is apparent from this study that social norms within team meetings are much more subtle than bold prescriptive processes like psychological safety and reflexivity suggest. Indeed, communication accommodation theory could be salient for understanding the dynamics within such discussions for this type of study. This subject will be returned to in the next chapter. However, suffice it to say here that attention could be given to understanding affective conflict (Amason, 1996), political behaviour (Dean & Sharfman, 1996), dissent (Dooley & Fryxell, 1999) or creativity (Dewett, 2004), to name a few, as part of understanding the black box of team processes during decision-making. These could be investigated as either indicators or intervening processes as discussed above (Lawrence, 1997).

What is striking about the results arising from this study in terms of theoretical extension, concerns cognitive variation. Traditionally, 'upper echelons' has been associated with pre-discussion cognitive diversity only at the team level, and then, only superficially as requisite mention in the congruence assumption between demographic proxies and organisational performance. As discussed in earlier sections of this, and previous chapters, in total, there were five predictors of pre-discussion cognitive variation but only at the individual level and none at the team level. However, there were seven predictors of post-discussion cognitive variation across both the individual and team levels, and also seven predictors of cognitive change across both levels.

This suggests that more recognition needs to be given to the role of individual opinions and the effect this may have when cognitive dissimilarity (individual) and cognitive diversity (team) persist subsequent to a team decision-making discussion.

The schematic presented below shows that in order to meaningfully extend 'upper echelons' based on this study, cognitive change and cognitive variation post-discussion need to be more prominent in understanding the relationship between demographic proxies, team dynamics and organisational performance. The question driving theoretical extension of 'upper echelons' has to be: What is it about being part of a team discussion that polarises individual viewpoints, but still allows for a collective consensus?

Furthermore, antecedents, situation specific factors, demographic variation and cognitive variation pre-discussion are grouped together. This grouping reflects the symbiotic relationship in the literature between demographic variation and cognitive variation (Hambrick et al., 1996), and more recent theoretical work around antecedents (Carpenter et al., 2004).

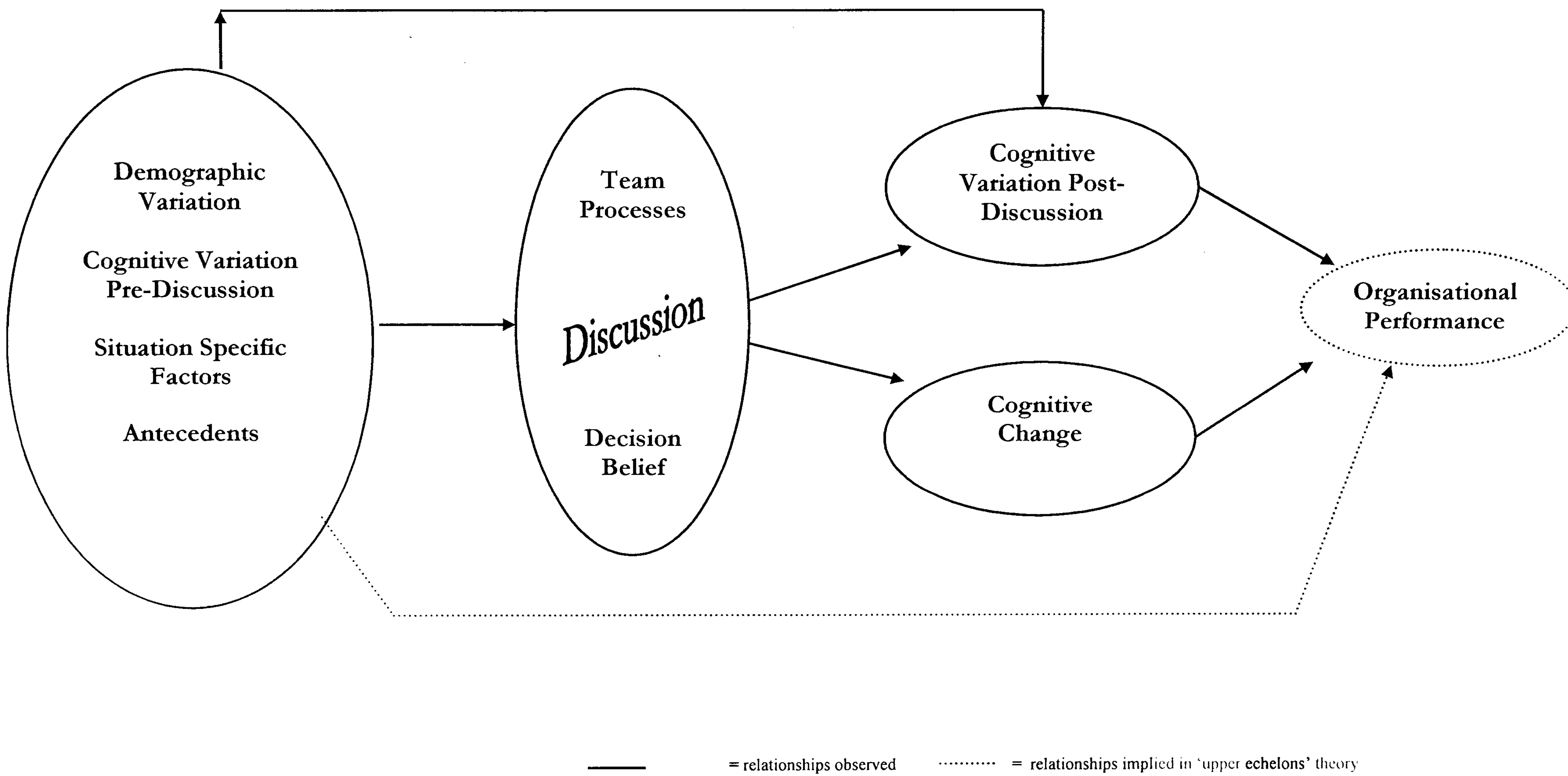
Finally, in deference to 'upper echelons' theory, it is necessary to continue seeking to understand what makes TMTs effective with particular reference to measurable organisational performance. The congruence assumption has been systematically discredited by various sources (see Lawrence 1997; Carpenter et al., 2004) (although it continues to persist), so it would be important to pursue investigations that include both intervening processes and organisational performance. Moreover, consideration must be given to the theoretical and methodological difficulties discussed in Chapter 2 concerning a cause and effect time lag with reference to financial measures (see Murray, 1989).

The current study has much to recommend it in terms of investigating cognitive variation pre- and post-discussion. That is to say, the research captured individuals' personal viewpoints prior to a team discussion which could positively or negatively influence those opinions, resulting in cognitive variation and/or cognitive change post-discussion. A criticism of previous research is that hypotheses relate to cognitive variation pre-discussion, but it is post-discussion which is measured (see Priem et al., 1999 re. Glick et al., 1993). This study measured both pre- and post-discussion consensus. Any subsequent study of these concepts would similarly need to secure such data in order to meaningfully extend the boundaries of this type of research.

As discussed in Chapter 3, this has not been a feature of TMT research, but emergent work using the simulation of a farmers market similarly measures cognitive diversity pre- and post-discussion (Mohammed & Ringseis, 2001). The use of a simulation as the content matter for the real discussion allows the researcher to set the parameters of the amount and type of information available to participants. In this case, it was rankings of candidates. In Mohammed and Ringseis's (2001) study, sub-groups of students in role play were given a set position from which to debate in the discussion. Obviously, there are problems of generalisability with their level of artificiality. However, the point being made here is that a simulation offers a mechanism for capturing individuals' choices both pre- and post-discussion.

When the research for this study was conducted some eight years ago, business simulations were somewhat novel, now they are commonly used in business, management education and play. Moreover, executives' familiarity with computers and technology in general has increased too. It is highly likely that as TMT research develops in the area of strategic cognition that simulations will become increasingly sophisticated and progressively more realistic. This bodes well for future research.

Figure 11.1 Research Model Re-visited



11.11 Conclusion

This chapter began by interpreting the findings for the hypotheses analysed in Chapters 7 to 9. Practical implications for TMTs were offered, but with caution as the number of results per hypothesis were very close to what one could expect due to chance alone. On balance, it would be fair to say that there is no systemic effect of demographic variation on cognitive variation, team processes or decision belief. Two correction techniques were applied, Bonferroni and Cohen's Standard, and only one or two results survived. Cohen's Standard reviewed effect sizes, which in each instance in this study, were very small, meaning that the practical implications for TMTs were negligible.

A discussion of measurement issues concluded that the methodology and the measures devised were robust enough to tap the constructs under investigation. Moreover, a raft of measures designed for the study were innovative and rigorous. Major strengths of the study were discussed such as the use of real, authentic TMTs, processes observed and measured in real time and the use of multi-level modeling. Limitations were also highlighted such as the lack of an organizational performance measure, use of a non-strategic decision and the omission of antecedent factors.

Although one must, on the basis of the sparse results and tiny effect sizes reject the hypotheses based on 'upper echelons' theory, the rigour with which the research was conducted do allow some fascinating insights into aspects of TMT functioning untapped in previous research. For example, it is clear from this study that demographic variation has more of an effect on post-discussion cognitive variation than on pre-discussion variation. Therefore, one must conclude that the discussion and the socialisation processes actually accentuate the demographic and cognitive variation effects. Furthermore, it is apparent through the use of multi-level modeling showed that team factors predicted individual cognitive variation.

Recognising these insights and acknowledging that 'upper echelons' remains the theory of choice for researchers in the TMT field, the discussion turned to a review of the Lawrence (1997) critique which put forward three ways in which the theoretical underpinnings of 'upper echelons' theory could advance. How the current research met these criteria was addressed, then, the research model first put forward in Chapter 4 was revisited in order to graphically illustrate the extension to 'upper echelons' theory made by the current study.

The next chapter explains in more detail how researchers can take forward the investigations of this study.

CHAPTER 12

Conclusions and Future Directions

12.0 Overview

This chapter provides the overall conclusion of the research, and suggests five directions emergent from the discussion in Chapter 11 for future research.

12.1 Introduction

The overall conclusion of this research is that there are not enough statistically significant findings, nor are the results of sufficient effect size to support the hypotheses based on 'upper echelons' theory. In terms of practical implications for TMTs, the conclusion afforded by this study is that they should neither avoid demographic variation nor actively seek it in order to achieve the widely supposed benefits of increased cognition, more effective decision-making and enhanced organizational performance.

However, the fact that there are any results at all, and that the findings concerning cognition suggests that there are differences pre- and post-discussion, and at individual and team levels, augurs well for research in this area to continue. Based upon the current study, there are several ways in which 'upper echelons' theory can be enhanced and extended. These are presented in the next section.

12.2 Towards an Agenda for Future Research

From the discussion in the Chapter 11, five key areas emerge as an agenda for future research. These are: (1) more work using effect sizes; (2) exploration of 'upper echelons' theory in relation to team longevity; (3) investigation of 'upper echelons' theory in circumstances of change; (4) investigation of whole life-cycles of decision-making; and (5) focus on team level predictors of individual cognition.

12.2.1 Effect Sizes

With regard to the first, it would seem that effect sizes could be particularly useful in establishing the relative value of practical interventions arising from 'upper echelons' research. From the current research and other studies cited earlier, it is of interest to note that effect sizes using Cohen's (1988) standard are all small. Cohen (1988) himself was hesitant in establishing the small, medium and large levels as a universal standard for all social science research, arguing that effect size levels should be determined on a basis germane to particular research streams. More work needs to be done concerning how small an effect size can be to have any theoretical or real practical value in TMTs. Researchers planning to use multi-level modeling would also do well to investigate effect sizes. As was noted earlier in this chapter, a pseudo-R^2 statistic currently needs to be computed manually for each model and even then, it is only an approximation of the R^2 variance in ordinary least squares regression. More work needs to be done to establish a normative standard for computing effect sizes in multi-level modeling that is analogous and therefore comparable to R^2 in regression analyses.

Further to the analysis of the Lawrence (1997) critique of 'upper echelons' theory discussed in Chapter 11, it is appropriate to give attention to effect sizes in order to establish which TMT processes are consistently operating within the TMT context.

12.2.2 'Upper Echelons' Theory and Communication Accommodation Theory

When seeking to extend 'upper echelons' theory, theoreticians may do well to pursue a connection with communication accommodation theory, which refers to the patterns of interaction that develop between people as they respond to others' cultural and demographic differences (Coupland & Giles, 1989). Originally a sociopsychological model of speech-style modifications it is now a way of interpreting relational processes in communicative interaction in any given situation (Giles, Coupland & Coupland, 1991). Although much of the work around communication accommodation theory has been in the area of health psychology, a small number of studies are beginning to investigate the communicative behaviours and strategies employed in the workplace as a response to demographic variation (see for example, Ayoko, Haertel & Callan, 2002; and Boggs, 1999). Communication accommodation theory has not yet been applied to TMTs. However, as this research finds that TMTs tend to assimilate individual demographic differences so that they have little observable effect, it seems that communication accommodation theory may hold the key to understanding how TMTs transcend differences to arrive at consensus decisions. This would be particularly interesting in terms of punctuated equilibrium when a new, perhaps dissimilar person joins a team, and the length of time and the mechanisms used by the team to assimilate such a person to the point where the dissimilarity is no longer noticed. This would require longitudinal studies of TMTs.

The present study deliberately went some way to examining processes, and similar questions to those posed above can be related to processes. For example, do TMTs use particular types of processes (e.g. procedural rationality) when they are newly formed? How long does it take before a team feels psychologically safe? Or if other processes are selected, how long does it take before a team engages in risk taking? These are questions that naturally arise from the current study in which the lack of highly visible results may suggest that TMTs have lost a certain consciousness or awareness as to these processes.

12.2.3 'Upper Echelons' Theory and TMT Contextual Changes

With regard to studying 'upper echelons' theory in the context of changing TMTs, what really needs to happen is that researchers become much more specific about the circumstances in which demographic variation and cognitive variation matter. For example, the study by Krishnan et al. (1997) reviewed in Chapter 2 investigated the influence of demographic variation in TMTs undergoing mergers and acquisitions. Krishnan et al. (1997) solely investigated functional background diversity, and found that this was detrimental for performance following a merger. In circumstances in which TMTs are merging, it is not hard to imagine that the influence of demographic factors, cognitive variation and team processes are more visible as people get to know one another, establish a pecking order and become familiar with each other's way of working.

12.2.4 'Upper Echelons' Theory and Life-Cycles of TMT Decision-Making

As most of the studies reviewed in Chapter 2 used a financial outcome measure of performance, the link between demographic variation and cognitive variation, and to decision-making over time, has not been established. For example, Knight et al. (1999) found that age diversity was associated with disagreement in TMTs. Whether such an effect is a constant (i.e. age diversity always means disagreement), or whether disagreement as a result of age diversity is stronger for certain types of decision, or indeed lessens over the time a team works together, is an unexploited avenue for future research.

With regard to studying whole life-cycles of decision-making, the current study utilized a business simulation in order to overcome some of the criticisms noted in Chapter 3 concerning the limitations of indirect, retrospective studies of strategic decisions. Further research needs to find ways of getting close to 'upper echelon' decision-makers, perhaps using a shadowing approach such as the 'boards-in-action' research of Samra-Fredricks (2000). What the current study points to is that individual executives may have very different private views even though they have ostensibly agreed to a team consensus. Future research would do well to investigate how such a disconnect may influence a person's commitment to implementation of the team consensus.

12.2.5 Team Level Predictors of Individual Cognition

With regard to team level predictors of individual cognition, a review of Table 11.1 shows that of the 13 results found to predict individual cognitive variation, 9 were team level predictors. Indeed, frequency of team meetings reduces cognitive variation in three of the five dependent variables. This suggests that much more work needs to be done to understand the how TMT processes affect cognitive variation amongst individuals in TMTs. Although some previous work has been conducted regarding individual cognition (concerning the extent to which individuals perceive that a team discussion has influenced their private opinions) this has been done with synthetic teams and has not taken a demographic variation approach (Heath & Gonzalez, 1995). Clearly, more work needs to be done in order to understand the contextual influence of real TMTs upon their individual members. Multi-level modeling provides a sophisticated means of understanding such relationships and needs to be used by 'upper echelons' researchers much more in the future.

12.3 Conclusion

In conclusion, although a thorough understanding of the relationships between demographic and cognitive variation, team processes and decision making requires further refinement, the present thesis has demonstrated that research no longer needs to rely on proxy variables for such insights. In so doing, the thesis has achieved the aims of: (1) defining and measuring cognitive variation; and (2) directly investigating social interaction processes within the context of authentic top management teams. The use of existing top management teams makes the research presented in this thesis particularly exciting. It also challenges 'upper echelons' researchers to move away from secondary archival data and into the real world. Furthermore, it behooves researchers to seek out the particular types of circumstances in which demographic variation may make a difference to authentic TMTs. This study concurs with the observation: *"Finally, the extra efforts are particularly necessary simply because the demographic alternative is so unacceptable"* (Priem et al., 1999).

REFERENCES

Adler, N. J. (1997). International Dimensions of Organizational Behavior. 3rd Ed. Cincinatti, South Western College Publishing.

Agresti, A., & Agresti, B. (1978). Statistical analysis of qualitative variation. Sociological Methodology, 9, 204-237.

Alexis, M., & Wilson, C. Z. (1967). Organizational Decision Making. New York: Prentice Hall.

Allison, P. D. (1978) Measures of inequality. American Sociological Review, 43, 865-880.

Amason, A. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic decision-making: Resolving a paradox for top management teams. Academy of Management Journal, 39, 123-148.

Amason, A. C., & Sapienza, H. J. (1997) The effects of top management team size and interaction norms on cognitive and affective conflict. Journal of Management, 23, 495-516.

Anderson, N. R., & West, M. A. (1994). The Team Climate Inventory: Manual and User's Guide. Windsor, England: ASE/NFER-Nelson Press.

Ayoko, O., Haertel, C., E., J., & Callan, V. J. (2002). Resolving the puzzle of productive and destructive conflict in culturally heterogeneous workgroups: A communication accommodation theory approach. International Journal of Conflict Management, 13, (2), 165-195.

Bales, R. F. (1949). Interaction process analysis: A method for the study of small groups. Cambridge, MA: Addison Wesley.

Bantel, K., & Jackson, S. (1989). Top management and innovations in banking: does the composition of the team make a difference? Strategic Management Journal, 10, 107-124.

Bateman, I. J., Jones, A. P., Nishikawa, N., & Brouwer, R. (2000). Benefits Transfer in Theory and Practice: A Review. CSERGE Working Paper, University of East Anglia. ISSN 0967-8875.

Bateman, T. S., & Zeithmal, C. P. (1989). The psychological context of strategic decisions: A model and convergent experimental findings. Strategic Management Journal, 10, (1), 59-74.

Bazerman, M. H. (1998). Judgement in Managerial Decision Making, (4th ed.) Wiley, New York.

Beach, L. R.. (1997). The Psychology of Decision Making: People in Organizations. Thousand Oaks, California: Sage.

Becker, L. (2000). Effect Size. http://web.uccs.edu/lbecker/Psy590/es.htm accessed 22/04/2005.

Blau, P. M. (1977). Inequality and heterogeneity. New York: Free Press.

Boggs, C. (1999). Equal employment opportunity, communication accommodation climate and the gender structuration of organizations: Development and test of a model for understanding the dynamics of male resistance to workplace gender integration. Unpublished PhD thesis.

Boone, C., Olffen, W. V., Wittleoostuijn, A. V., & Brabander, B. (2004). The genesis of top management team diversity: Selective turnover among top management teams in Dutch newspaper publishing, 1970-1994. Academy of Management Journal, 47, 633-656.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. In N. L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally.

Carli, L. L., & Bukatko, D. (2000). Gender, communication and social influence: A developmental perspective. In T. Eckes & H. M. Trautner (Eds), The Developmental Social Psychology of Gender. Mahwah, NJ: Earlbaum, 295-331.

Carpenter, M. A. (2000). The implications of strategy and social context for the relationship between top management team heterogeneity and firm performance. Strategic Management Journal.

Carpenter, M. A. Geletkanycz, M. A. & Sanders, W. G., (2004). Upper Echelons research revisited: Antecedents, elements and consequences of top management team composition. Journal of Management, 30, 748-778.

Carroll, G. R., & Harrison, J. R. (1998). Organizational demography and culture: Insights from a formal model and simulation. Administrative Science Quarterly, 43, 737-767.

Carson, C. M., Mosley, D. C., & Boyar, S. L. (2004). Performance gains through diverse top management teams. Team Performance Management, 10, (56), 121-126.

Carter, A. J. (2000). Team viability, group precesses and well-being in secondary health care teams. Working Paper #266, Institute of Work Psychology/ESRC Centre for Organisation and Innovation, University of Sheffield, U.K.

Chatman, J. A., & Flynn, F. J. (2001). The influence of demographic heterogeneity on the emergence and consequences of cooperative norms in work teams. Academy of Management Journal, 44, 956-974.

Chattophadhyay, P., Glick, W. H., Miller, C. C., & Huber, G. P. (1999). Determinants of executive beliefs: Comparing functional conditioning and social influence. Strategic Management Journal, 20, 763-789.

Chemers, M. M., Oskcamp, S., & Costanzo, M. A. (Eds.). (1995a). Diversity in Organizations: New Perspectives for a Changing Workplace. Thousand Oaks: CA: Sage.

Church, A. H. (1997). Managerial self-awareness in high-performing individuals in organisations. Journal of Applied Psychology, 82, 2, 281-292.

Clark, K., Smith, K., G, Sims, H. P. J., Flood, P., Moore, S., Morley, M., & O'Regan, P. (1997). Configuration of Top Management Teams: Structure, composition, process, power and leadership. Paper presented at the Annual meeting of the Academy of Management.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, L., Broschak, J., & Haveman, H. (1996). And then there were more?: The effect of organizational sex composition on hiring and promotion : Working Paper, Haas School of Business, University of California, Berkley.

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. Administrative Science Quarterly, 17(1), 1-25.

Cohen, M. D., March, J. G., & Olsen, J. P. (1976). People, problems, solutions and the ambiguity of relevance. In J. G. March & J. P. Olsen (Eds.), Ambiguity and choice in organizations (pp. 24-37). Norway: Harald Lyche & Co.

Collins, B. E., & Guetskow, H. (1964). A Social Psychology of Group Processes for Decision-Making. New York: John Wiley & Sons.

Connelly, T., & Wagner, W. G. (1988). Decision cycles. Advances in Information Processing in Organizations, 3, 183-205.

Cook, T. D., Campbell, D. T., & Peracchio, L. (1990). Quasi-experimentation. In M. D. Dunnett & L. M. Hough (Eds). Handbook of Industrial and Organizational Psychology. Chapter 9. Palo Alto, CA: Consulting Psychologists Press.

Coupland, N., & Giles, H. (1989). The communicative contexts of accommodation. Language and Communication, 8, (3-4), 175-182.

Cox, T., Lobel., S., & McLeod, P. (1991) Effects of ethnic group cultural differences on cooperative and competitive behaviour research. Academy of Management Journal, 34, 827-847.

CWBR, (1994). Styles of Success: The Thinking and Management Styles of Women and Men Entrepreneurs. Washington, DC: Centre for Women's Business Research (CWBR).

Cyert, R., M, & March, J. G. (1963). A Behavioral Theory of the Firm. Englewood Cliffs, New Jersey: Prentice-Hall.

Das, T. K., & Teng, B. (1999). Cognitive biases and strategic decision processes: An integrative perspective. Journal of Management Studies, 36, (6), 757-778.

Davis, J. H. (1969). Group Performance. Reading, MA: Addison-Wesley.

Dean, J. W., & Sharfman, M. P. (1996). Does decision process matter? A study of strategic decision-making effectiveness. Academy of Management Journal, 39, (2), 368-396.

Dewett, T. (2004). Creativity and strategic management: Individual and group considerations concerning decision alternatives in top management teams. Journal of Managerial Psychology, 19, (2), 156-169.

Dewey, J. (1933). How we think. Boston, DC: Heath & Company.

Dooley, R. S., & Fryxell, G. E. (1999). Attaining decision quality and commitment from dissent: The moderating effects of loyalty and competence in strategic decision-making teams. Academy of Management Journal, 42, (4), 369-401.

Downie, N. M., & Heath, R. W. (1970). Basic Statistical Methods. (3rd Ed)., New York: Harper & Row.

Edmondson, A. C. (1998). Psychological safety and learning behavior in work teams. Unpublished Working Paper.

Edmondson, A. C., Roberto, M. A., & Watkins, M. D. (2003) A dynamic model of top management team effectiveness: Managing unstructured task streams. Leadership Quarterly, 14, 297-325.

Eisenhardt, K. M. (1989). Making fast decisions in high-velocity environments. Academy of Management Journal, 32, 543-577.

Eisenhardt, K. M., & Schoonhoven, C. B. (1990). Organizational growth: Linking founding team strategy, environment, and growth among U.S. semiconductor ventures, 1978-1988. Administrative Science Quarterly, 35, 504-529.

Eisenhardt, K. M., Bourgeois, T. L. (1988). Politics of strategic decision making in high-velocity environments: Towards a midrange theory. Academy of Management Journal, 32, 543-576.

Ely, R. J. (1994). The effects of organizational demographics and social identity on relationships among professional women. Administrative Science Quarterly, 39, 203-238.

Etzioni, A. (1967). Mixed-scanning: A "third" approach to decision-making. Public Administration Review, 27, 385-392.

Finkelstein, S. (1992). Power in top management teams: Dimensions, measurement and validation. Academy of Management Journal, 35, 505-538.

Finkelstein, S., & Hambrick, D.C. (1990). Top-management-team tenure and organizational outcomes: The moderating role of managerial discretion. Administrative Science Quarterly, 35, 484-503.

Fisher, B. A. (1970). Decision emergence: Phases in group decision-making. Speech Monographs, 37, 53-66.

Flatt, S. (1996). Developing innovative strategies: How top management teams bring creativity and implementation to the firm. Paper presented at the 16th Annual Strategic Management Society Conference, Pheonix, Arizona.

Forbes, D. P., & Milliken, F. J. (1999), Cognition and corporate governance: Understanding boards of directors as strategic decision-making groups. Academy of Management Review, 24, (3), 489-505.

Forte, A. (2004). Antecedents of managers moral reasoning. Journal of Business Ethics, 51, 315-347.

Foushee, M. C. (1984). Dyads and triads at 35,000 feet: factors affecting group process and aircrew performance. American Psychologist, 39, 885-893.

Frederickson, J. W. (1984). The comprehensiveness of strategic decision-processes: Extension, observations, future directions. Academy of Management Journal, 27, 399-423.

Fripp, J. W. W. (1993). Learning Through Simulations: Guide to the Design and Use of Simulations in Business and Education. St Ives: McGraw-Hill.

George, J. M. (1990). Personality, affect and behaviour in groups. Journal of Applied Psychology, 75, (2), 107-116..

Gigone, D., & Hastie, R. (1997). Proper analysis of the accuracy of group judgments. Psychological Bulletin, 121, (1), 149-167.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context and consequence. In H. Giles & J. Coupland (Eds). Contexts of Accommodation: Developments in Applied Social Linguistics. Studies in Emotion and Social Interaction. New York: John Wiley

Glick, W., Miller, C., & Huber, G. (1993). The impact of upper echelon diversity on organizational performance. In G. Huber and W. Glick (Eds.), Organizational change and redesign. 176-224. New York: Oxford University Press.

Golden, B. R.. (1992). The past is the past – or is it? The use of retrospective accounts as indicators of past strategy. Academy of Management Journal, 35, (4), 848 – 860.

Goldstein, H. (1995). Multilevel Statistical Models. London: Edward Arnold/New York: Halstead.

Gouran, C. D. (1973). Group communication: Perspectives and priorities for future research. Quarterly Journal of Speech, 59, 22-29.

Graves, L. M., & Elsass, P. M. (2005). Sex and sex dissimilarity effects in ongoing teams: Some surprising findings. Human Relations, 58, 191-221.

Green, S. B. (1991). How many subjects does it take to do a regression analysis? Multivariate Behavioral Research, 26, 449-510.

Guzzo, R. A. (1982). Improving Group Decision Making in Organizations: Approaches From Theory and Research. New York: Academic Press.

Guzzo, R. A. (1996). Fundamental considerations about work groups. In M. A. West (Ed.), Handbook of Workgroup Psychology (pp. 3-24). Chichester: Wiley.

Hackman, J. R., & Morris, C. G. (1978). Group tasks, group interaction process and group performance effectiveness: A review and proposed integration. In L. Berkowitz (Ed.), Group Processes (pp. 1-55). New York: Academic Press.

Halebian, J., & Finkelstein, S. (1993). Top management team size, CEO dominance and firm performance: The moderating roles of environmental turbulence and discretion. Academy of Management Journal, 36(4), 844 - 863.

Hambrick, D. C., & D'Aveni, R. A. (1992). Top team deterioration as part of the downward spiral of large corporate bankruptcies. Management Science, 38, 1445-1466.

Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: the organization as a reflection of its top managers. Academy of Management Review, 15, 514-535.

Hambrick, D. C., (1981). Strategic awareness within top management teams. Strategic Management Journal, 2, 263-279.

Hambrick, D. C., Cho, T., & Chen, M. (1996). The influence of top management team heterogeneity on firms; competitive moves. Administrative Science Quarterly, 41, 659-684.

Hayduk, L. A. (1987). Structural Equation Modelling with LISREL: Essentials and Advances. Baltimore, MD: Johns Hopkins University Press.

Hays, J. R.. (1981). Ethics of prediction of future dangerousness. Psychological Reports, 49, (2), 593-594.

Heath, C., & Gonzalez, R. (1995). Interaction with others increases decision confidence but not decision quality: Evidence against information collection views of interactive decision making. Organizational Behavior and Human Decision Processes, 61, (3), 305-326.

Heinicke, C., & Bales, R. F. (1953). Developmental trends in the structure of small groups. Sociometry, 16, 7-38.

Hendry, J. (2000). Strategic decision-making, discourse, and strategy as social practice. Journal of Management Studies, 37, (7), 955-977.

Hermann, P., & Datta, D. K. (2005). Relationships between top management team characteristics and international diversification: an empirical investigation. British Journal of Management, 16, 69-78.

Hewes, D. E. (1986). A socioegocentric theory of group decision-making. In R. Y. Hirokawa & M. S. Poole (Eds.), Communication and group decision-making . Beverly Hills, CA: Sage.

Hickson, D. J., Butler, R. J., Cray, D., Mallory, G. R., & Wilson, D. C. (1986). Top decisions: Strategic decision making in organisations. Oxford, UK: Basil Blackwell.

Higgins, J. C. (1980) Strategic and Operational Planning Systems: Principles and Practice. Prentice-Hall International, London.

Hirokawa, R. Y. (1983). Group communication and problem solving effectiveness: An investigation of group phases. Human Communication Research, 9(4), 291-305.

Hirokawa, R. Y. (1990). The role of communication in group decision-making efficacy. Small Group Research, 21(2), 190-204.

Hirokawa, R. Y., & Pace, R. (1983). A descriptive investigation of the possible communication-based reasons for effective and ineffective group decision making. Communication Monographs, 50, 363-382.

Hodgkinson, G. P. (2001). Cognitive processes in strategic management: Some emerging trends and future directions. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds). Handbook of Industrial, Work and Organizational Psychology, (vol 2), 416-440.

Hodgkinson, G. P. (2003). The interface of cognitive and industrial, work and organizational psychology. Journal of Occupational and Organizational Psychology, 76, 1-25.

Hodgkinson, G. P., & Sparrow, P. R. (2002). The Competent Organization: A Psychological Analysis of the Strategic Management Process. Buckingham, Open University Press.

Hodgkinson, G. P., & Thomas, A. B. (1997). Editorial introduction to the special issue: Thinking in organizations. Journal of Management Studies, 34, (6), 845-850.

Hodgkinson, G. P., Bown, N. J., Maule, A. J., Glaister, K. W., & Pearman, A. D. (1999). Breaking the frame: An analysis of strategic cognition and decision making under uncertainty. Strategic Management Journal, 20, 977-985.

Hoffman, L. R. (1979). Applying experimental research on group problem solving to organizations. Journal of Applied Behavioural Science, 15, 375-391.

Howell, D. C. (2002). Statistical Methods for Psychology. (5th Ed)., Pacific Grove, CA: Duxbury.

Hox, J. J. (1995). Applied Multilevel Analysis. Amsterdam: TT-Publikaties.

Hrebiniak, L. G., & Snow, C. C. (1982). Top management agreement and organizational performance. Human Relations, 35, (12), 1139-1158.

Huber, G. P., & Power, D. J. (1985). Retrospective reports of strategic level managers: Guidelines for increasing their accuracy. Strategic Management Journal,6, 171-180.

Huff, A. S. (1997). A current and future agenda for cognitive research in organizations. Journal of Management Studies, 34, (6), 947-952.

Huff, A. S. (ed.) (1990). Mapping Strategic Thought. Wiley, Chichester.

Hurst, D. K., Rush, J. C., & White, R. E. (1989). Top management teams and organisational renewal. Strategic Management Journal, 10, 87-105.

Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. Administrative Science Quarterly, 37, 422-447.

Ilgen, D., LePine, J., & Hollenbeck, J. (1999). Effective decision making in multinational teams. In P. C. Earley, & M. Erez (Eds.), New Approaches to Intercultural and International: Industrial/Organizational Psychology. 377-409. San Francisco, The New Lexington Press.

Isabella, L. A., & Waddock, S. A. (1994). Top management team certainty: Environmental assessment, teamwork, and performance implications. Journal of Management, 20, (4), 835-854.

Jackson, S. E. (1992). Team composition in organizational settings: Issues in managing a diverse work force. In S. Worchel, W. Wood, and J. Simpson, (Eds.) Group processes and productivity. Beverly Hills, CA: Sage.

Jackson, S. E., Brett, J. F., Sessa, V. I., Cooper, D. M., Julin, J. A., & Peyronnin, K. (1991). Some differences make a difference: Individual dissimilarity and group heterogeneity as correlates of recruitment, promotions and turnover. Journal of Applied Psychology, 76, 675-689.

Jackson, S. E. & Joshi, A. (2001). Research on domestic and international diversity in organizations: A merger that works? In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds). Handbook of Industrial, Work and Organizational Psychology, (vol 2), 206-231.

James, L. R. (1982). Aggregation biase in estimated perceptual agreement. Journal of Applied Psychology, 67, (2), 219-229.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. Journal of Applied Psychology, 69, (1), 85-98.

Janis, I. L. (1972) Victims of Groupthink: A Psychological Study of Forgign Policy Decisions. Boston: Houghton Mifflin.

Janssen, O., Van de Vliert, E., & Veenstra, C. (1999). How task and person conflict shap the role of positive interdependence in managmenet teams. Journal of Management, 25, (2), 117-142.

Jenkins, M., & Johnson, G. (1997). Entrepreneurial intentions and outcomes: A comparative causal mapping study. Journal of Management Studies, 34, (6), 895-920.

Johnson, D., & Johnson, F. (1987). Joining Together: Group Theory & Group Skills. Englewood Cliffs, NJ: Prentice-Hall.

Jones, K., & Duncan, C. (1998). Modelling context and heterogeneity: Applying multilevel models. In E. Scarborough and E. Tannenbaum, (Eds)., Research Strategies in the Social Sciences: A guide to new approaches. England: Oxford University Press.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. Econometrica, 47, 263-291.

Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgment under uncertainty: Heuristics and biases. Cambridge, UK: Cambridge University Press.

Kanter, R. M. (1977). Some effects of proportions on group life: Skewed sex ratios and responses to token women. American Journal of Sociology, 82, 965-990.

Kanter, R. M. (1983). The Change Masters: Corporate Entrepreneurs at Work. New York: Simon & Schuster.

Karakowsky, L., & Elangovan, A. R. (2001). Risky decision making in mixed-gender teams – whose risk tolerance matters? Small Group Research, 32, 1, 94-111.

Kilduff, M., Angelmar, R., & Mehra, A. (2000). Top management team diversity and firm performance: Examining the role of cognitions. Organization Science, 11, 21-34.

King, N., Anderson, N. R., & West, M. A. (1992). Organizational innovation: A case study of perceptions and processes. Work and Stress, 5, 331-339.

Kirchmeyer, C. (1985). Demographic similarity to the work group: A longitudinal study of managers at the early career stage. Journal of Organizational Behavior, 16, 67-83.

Knight, D., Pearce, C. L., Smith, K. G., Olian, J. D., Sims, H. P., Smith, K. A., & Flood, P. (1999). Top management team diversity, group process and strategic consensus. Strategic Management Journal, 20, 445-465.

Kozlowski, S. W., & Hattrup, K. (1992). A disagreement about within-group agreement: Disentangling issues of consistency versus consensus. Journal of Applied Psychology, 77, (2), 161-167.

Krishnan, H. A., Miller, A., & Judge, W. Q. (1997) Diversification and top management team complementarity: is performance improved by merging similar or dissimilar teams? Strategic Management Journal, 18, (5), 361-374.

Lau, D. C., & Murnighan, J.K. (2005). Interactions within groups and sub-groups: The effects of demographic faultlines. Academy of Management Journal, 48, 4, 645-659.

Lawler, E. E., & Hackman, J. (1969). Impact of employee participation in the development of pay incentive plans: A field experiment. Journal of Applied Psychology, 53, 467-471.

Lawrence, B.S. (1988). New wrinkles in the theory of age: Demography, norms, and performance ratings. Academy of Management Journal, 31, 309-337.

Lawrence, B.S. (1997). The black box of organizational demography. Organizational Science, 8, 1-22.

Lindblom, C. E. (1959). The science of 'muddling through'. Public Administration Review, 17, 78-88.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment. American Psychologist.48, 1181-1209.

Lord, F. M., & Novick, M. R. (1968), Statistical Analysis of Mental Test Scores. Reading, MA: Addison Wesley.

March, J. G., & Simon, H. A. (1958). Organizations. New York: John Wiley.

Markóczy, L. (1997). Measuring beliefs: Accept no substitutes. Academy of Management Journal, 40, 5, 1228-1242.

Maule, A. J., & Hodgkinson, G. P. (2003). Re-appraising mangers personal perceptual errors: A behavioural decision-making perspective. British Journal of Management, 14, 33-37.

McCullagh, P., & Nelder, J. A. (1989). Generalized Linear Models. London: Chapman & Hall.

McGrath, J. E. (1991). Time, Interaction and Performance (TIP): A theory of groups. Small Group Research, 22, 147-174.

Mezias, J. M., & Starbuck, W. H. (2003). Study the accuracy of managers' perceptions: A research odyssey. British Journal of Management, 14, 3-17.

Michel, J. G., & Hambrick, D. C. (1992). Diversification on posture and top management team characteristics. Academy of Management Journal, 35, 9-37.

Miller, C. Burke, L., & Glick, W. H. (1998) Cognitive diversity among upper-echelon executives: Implications for strategic decision processes. Strategic Management Journal, 19, 39-58.

Milliken, F., & Martins, L. (1996) Searching for common threads: Understanding the multiple effects of diversity in organizational groups. Academy of Management Review, 21, 402-433.

Mintzberg, H., Raisinghani, D., & Théorêt, A. (1976). The structure of 'unstructured' decision processes. Administrative Science Quarterly, 21, 246-275.

Mohammed, S., & Ringseis, E. (2001). Cognitive diversity and consensus in group decision-making: The role of inputs, processes and outcomes. Organizational Behavior and Human Decision Processes, 85, (2), 310-335.

Murray, A. (1989) Top management group heterogeneity and firm performance. Strategic Management Journal, 10, 125-141.

Nadler, D. A., Hackman, J. R., & Lawler, E. E. (1979) Managing Organizational Behavior. Boston, Little, Brown & Co.

Nemeth, C., & Owens, P. (1996). Making work groups more effective: the value of minority dissent. In M. A. West (Ed.), Handbook of Workgroup Psychology, 125-142. Chichester: Wiley.

Nieva, V., & Gutek, B. (1980). Sex effects on evaluation. Academy of Management Review, 5, 267-276.

Nisbitt, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.

Norburn, D., & Birley, S. (1988). The top management team and corporate performance. Strategic Management Journal, 9, 225-227.

O'Reilly, C. A., & Flatt, S. (1989). Executive team demography, organizational innovation, and firm performance. Paper presented at the Forty-ninth Annual meeting of the Academy of Management, Washington, D.C.

O'Reilly, C. A., Snyder, R., & Boothe, J. (1993). Effects of executive team demography on organizational change. In G. Huber and W. Glick (Eds.), Organizational change and redesign. 147-175. New York: Oxford University Press.

Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision Making in Action: Models and Methods (pp. 3-21). Norwood, New Jersey: Ablex Publishing Corporation.

Orasanu, J., & Salas, E. (1993). Team decision making in complex environments. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision Making in Action: Models and Methods (pp. 327-345). Norwood, New Jersey: Ablex Publishing Corporation.

Orlando, R. C., & Shelor, R. M. (2002). Linking top management team age heterogeneity to firm performance: juxtaposing two mid-range theories. International Journal of Human Resource Management, 13, (6), 958-974.

Papadakis, V. M., Lioukas, S., & Chambers, D. (1998) Strategic decision-making processes: The role of management and context. Strategic Management Journal, 12, (2), 115-147.

Papadakis, V. M., & Barwise, P. (2002) How much do CEOs and top managers matter in strategic decision-making? British Journal of Management, 13, 83-95.

Parker, S. K. (1998) Enhancing role-breadth self-efficacy: The roles of job enrichment and other organisational interventions. Journal of Applied Psychology, 83, (6), 835-852.

Pelled, L. H. (1996) Relational demography and perceptions of group conflict and performance: A field investigation. International Journal of Conflict Management, 7, 230-246.

Pelled, L. H., Eisenhardt, K., & Xin, K. (1999) Exploring the black box: An analysis of work group diversity, conflict and performance. Administrative Science Quarterly, 44, 1-28.

Perneger, T. V. (1998). What is wrong with Bonferroni adjustments. British Medical Journal, 136, 1236-1238.

Peterson, C. R., Schneider, R. J., & Miller, A. J. (1965). Sample size and the revision of subjective probabilities. Journal of Experimental Psychology, 69, 522-527.

Petzall, S. B., Selvaragjah, C. T., & Willis, Q. F. (1991) Management: A Behavioural Approach. Longman Cheshire, Melbourne, Australia.

Pettigrew, A. (1992). On studying managerial elites. Strategic Management Journal, 13, 163-182.

Pfeffer, J. (1983). Organizational demography. In B. Staw and L. Cummings (Eds.), Research in Organizational Behavior, 5, 299-357, Greenwich, CT: JAI Press.

Priem, R.. L., Lyon, D. W., Dess, G. G. (1999) Inherent limitations of demographic proxies in top management team research. Journal of Management, 25, 935-953.

Rasbash, J., Prosser, R., & Steele, F. (2004). Multi-level Modelling Using MLWin. Institute of Education, London, UK.

Rasmussen, J. (1993). Deciding and doing: Decision making in natural contexts. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision Making in Action: Models and Methods (pp. 158-171). Norwood, New Jersey: Ablex Publishing Corporation.

Rogelberg, S. G., Leach, D. J., Warr, P. B., & Burnfield, J. L. (2005) "Not another meeting!" Are meeting time demands related to employee well-being? Journal of Applied Psychology, in press.

Rogelberg, S. G., & Rumery, S.M. (1996). Gender diversity, team decision quality, time on task and interpersonal cohesion. Small Group Research, 27, 1, 79-90.

Ruble, T., Cohen, R., & Ruble, D. (1984). Sex stereotypes. American Behavioral Scientist, 27, 339-356.

Salancik, G. R. & Pfeffer, J. (1977) Who gets power – and how they hold on to it. Organization Science, Winter, 3-21.

Samra-Fredricks, D. (2000). Doing 'Boards-in-Action' Research – an ethnographic approach for the capture and analysis of directos' and senior managers' interactive routines. Corporate Governance, 8, (3), 244-257.

Sankoh, A. J., Huque, M. F., & Dubey, S. D. (1997). Some comments on frequently used multiple endpoint adjustments methods in clinical trials. Statistics in Medicine, 16, 2529-2542.

Schön, D. A. (1983). The reflective practitioner: How professional think in action. New York: Basic Books.

Schweiger, D. M., Sandberg, W. R., & Ragan, J. W. (1986). Group approaches for improving strategic decision making: A comparative analysis of dialectical inquiry, devil's advocacy and consensus. Academy of Management Journal, 29, (1), 51-71.

Schwenk, C. R. (1982). Why sacrifice rigour for relevance? A proposal for combining laboratory and field research in strategic management. Strategic Management Journal, 3, 213-225.

Schwenk, C. R. (1984). Cognitive simplification processes in strategic decision-making. Strategic Management Journal, 5, 111-128.

Schwenk, C. R. (1985). The use of participant recollection in the modelling of organizational processes. Academy of Management Review, 10, 496-503.

Schwenk, C. R. (1985). The use of participant recollection in the modelling of organizational decision processes. Academy of Management Review, 10, (3), 496-503.

Shadish, W. R., Cook, D. T., & Campbell, D. T. (2001). Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin.

Simon, H. A. (1957). Models of man: Social and rational. New York: John Wiley.

Simon, H. A. (1961). Administrative behavior. New York: Macmillan.

Simons, T. (1995). Top management team consensus, heterogeneity, and debate as contingent predictors of company performance. The complimentarity of group structure and process. Academy of Management Best Papers Proceedings, 62-65.

Simons, T., Pelled, L. H., & Smith, K. A. (1999). Making use of difference: Diversity, debate and decision comprehensiveness in top management teams. Academy of Management Journal, 42, (6), 662-673.

Simple Interactive Statistical Analysis (SISA). Bonferroni Correction Online: Adjustment for Multiple Comparisons. http://home.clara.net/sisa/bonhlp.htm accessed 22/04/05.

Smith, K., Smith, K., Olian, J., Sims, H., O'Bannon, D., & Scully, J. (1994). Top management team demography and process: the role of social integration and communication. Administrative Science Quarterly, 39, 412-438.

Sniezek, J. A. (1992). Groups under uncertainty: An examination of confidence in group decision making. Organizational Behavior and Human Decision Processes, 52, 124-155.

Sniezek, J. A., & Henry, R. A. (1989). Accuracy and confidence in group judgement. Organizational Behavior and Human Decision Processes, 43, 1-28.

Sniezek, J. A., & Henry, R. A. (1990). Revision, weighting, and commitment in consensus group judgement. Organizational Behavior and Human Decision Processes, 45, 1, 66-84.

Sorenson, J. B. (1999). The ecology of organisational demography: Managerial tenure and organizational competition. Industrial and Corporate Change, 8, (4), 713-744.

Sparrow, P. R. (1994). The psychology of strategic management: Emerging themes of diversity and cognition. In C. L. Cooper & I. T. Robertson (Eds.), International Review of Industrial and Organizational Psychology (Vol. 9, pp. 147-181). Chichester: John Wiley & Sons Ltd.

SPSS. (1993). SPSS for Unix: Advanced Statistics, Release 5.0. Chicago, IL.

Steiner, I. D. (1972). Group Process and Productivity. New York: Academic Press.

Stride, C. B., Swift, T. A., & Wall, T. D. (2000). Team diversity in equal measures? Working Paper, Institute of Work Psychology, University of Sheffield, Sheffield, UK.

Svenson, O. (1979). Process descriptions in decision-making. Organization and Human Decision Processes, 23, 86-112.

Svenson, O. (1992). Differentiation and consolidation theory of human decision making: A frame of reference for the study of pre-and post-decision processes. Acta Psychologica, 80, 143-168.

Swift, T. A., and West, M. A. (1998) Reflexivity and Group Processes. Sheffield, UK: ESRC Centre for Organisation and Innovation. ISBN: 0 9533504 1X

Swift, T. A., Humphrey, C., & Gor, V., (2000) "Great Expectations?: Quality Audits inherit a Dubious Financial Legacy", British Journal of Management, 11, 34 – 49.

Swim, J., Borgida, E., Maruyama, G., & Myers, D. (1989). Joan McKay versus John McKay: Do gender stereotypes bias evaluations? Psychological Bulletin, 105, 409-429.

Tabachnick, B. G., & Fidell, L. S. (2001) Using Multivariate Statistics. (4th Ed). Needham Heights, MA. Allyn & Bacon.

Taylor, D. W., & Faust, W. L. (1952). Twenty questions: Efficiency in problem-solving as a function of group size. Journal of Experimental Psychology, 44, 360-368.

Tetlock, P. E. (1979). Identifying victims of Groupthink from public statements of decision makers. Journal of Personality and Social Psychology, 37, (8), 1314-1324.

Thomas, J. B., & McDaniel, R. R. (1990). Interpreting strategic issues: Effects of strategy and the information-processing structure of top management teams. Academy of Management Journal, 33, (2), 286-306.

Thoresen, C. J., Bradley, J. C., Bliese, P. D., & Thoresen, J. D. (2004). The big fiver personality traits and individual job performance growth trajectories in maintenance and transitional job stages. Journal of Applied Psychology, 89, (5), 835-853.

Tjosvold, D. (1992). Team organization: An enduring competitive advantage. New York: John Wiley.

Tjosvold, D. (1995). Cooperation theory, constructive controversy, and effectiveness: Learning from crisis. In R. A. Guzzo, E. Salas, & a. associates (Eds.), Team Effectiveness and Decision Making in Organizations (pp. 79-112). San Francisco: Jossey Bass.

Tjosvold, D., Yu, Z., Hui, C. (2004). Team learning from mistakes: The contribution of cooperative goals and problem solving. Journal of Management Studies, 41, (7), 1223-1245.

Trochim, W. M. (2002). The Research Methods Knowledge Base. http://www.socialresearchmethods.net/kb/

Tsui, A. S., & O'Reilly III, C. A. (1989). Beyond simple demographic effects: The importance of relational demography in superior-subordinate dyads. Academy of Management Journal, 32, 402-423.

Tsui, A. S., Egan, T. D., & O'Reilly III, C. A. (1992). Being different: Relational demography and organizational attachment. Administrative Science Quarterly, 37, 549-579.

Tsui, A. S., Porter, L. W., & Egan, T. S. (2002). When both similarities and dissimilarities matter: Extending the concept of relational demography. Human Relations, 55, (8), 899-929.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 221, 1124-1131.

Tversky, A., & Kahneman, D. (1981). The framing of decsions and the psychology of choice. Science, 221, 1124-1131.

Tversky, A., & Kahneman, D. (1982). Causal schemas in judgments under uncertainty. In D. Kahneman, S. P, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and Biases (pp. 117-128). Cambridge, UK: Cambridge University Press.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty, 5, 297-323.

Volkema, R. J., & Gorman, R. H. (1998). The influence of cognitive-based group composition on decision-making process and outcome. Journal of Management Studies, 35, (1), 105-121.

Wagner, D. G., & Berger, J. (1997). Gender and interpersonal task behaviours: Status expectation accounts. Sociological Perspectives, 40, 1-32.

Wagner, W., Pfeffer, J., & O'Reilly, C. (1984). Organizational demography and turnover in top management groups. Administrative Science Quarterly, 29, 74-92.

Webber, R. A. (1987). Changes in perception in mixed gender teams. Human Resource Management, 26, 4, 455-467.

Weick, K. (1979) The Social Psychology of Organizing. Addison-Wesley, Boston.

Weingart, L. R. (1997). How did they do that: The ways and means of studying group process. Research in Organizational Behavior, 19, 189-239.

West, C. T., & Schwenk, C. R. (1996). Top management team strategic consensus, demographic homogeneity and firm performance: A report of resounding non-findings. Strategic Management Journal, 17, 571-576.

West, M. A. (1994). Effective Teamwork. Leicester, UK. BPS Books.

West, M. A. (1996). Reflexivity and work group effectiveness: A conceptual integration. In M. A. West (Ed.), Handbook of Workgroup Psychology (pp. 555-579). Chichester: Wiley.

West, M. A., & Anderson, N. R. (1996). Innovation in top management teams. Journal of Applied Psychology, 81(6), 680-693.

West, M. A., Garrod, S., & Carletta, J. (1997). Group decision-making and effectiveness: Unexplored boundaries. In C. L. Cooper & S. E. Jackson (Eds.), A Handbook for Future Research in Organizational Behaviour (pp. 293-317). Chichester, UK: Wiley.

Wiersma, M., & Bantel, K. (1992). Top management team demography and corporate strategic change. Academy of Management Journal, 35, 91-121.

Wiersma, M., & Bird, A. (1993). Organizational demography in Japanese firms: Group heterogeneity, individual dissimilarity, and top management team turnover. Academy of Management Journal, 36, 996-1025.

Williams, K. Y., & O'Reilly III, C. A. (1998). Demography and diversity in organizations: A review of 40 years of research, Research in Organizational Behavior (Vol. 20, pp. 77-140), JAI Press.

Williges, R. C., Johnston, W. A., & Briggs, G. E. (1966). Role of verbal communication in teamwork. Journal of Applied Psychology, 50, 473-478.

Wilson, F. (2003). Organizational Behaviour and Gender. (2nd Ed.). Aldershot, Ashgate.

Winter, S. G. (2003). Mistaken perceptions: Cases and consequences. British Journal of Management, 14, 39-44.

Zenger, T. T., & Lawrence, B. S. (1989). Organisational demography: the differential effects of age and tenure distributions on technical communication. Academy of Management Journal, 32, 353-376.

A P P E N D I C E S

Team Diversity in Equal Measures?

Stride, C. B., Swift, T. A., & Wall, T. D.

Summary

It is argued in this paper that, despite being regarded as one of the premier indices of heterogeneity in demographic diversity research, the coefficient of variation is actually unsuitable for use in small groups and top management research. For the past two decades, a considerable amount of research in the upper echelons tradition as well as in small group research more generally has relied upon this measure. Yet, using examples from real top management teams, this paper exposes the limitations of the coefficient of variation for investigating team diversity, and proposes a new index that satisfactorily addresses the measurement challenges inherent in small group research.

Introduction

Over the last two decades there has been much interest in the impact of team member diversity on team effectiveness. This is based on the assumption that diversity with respect to factors such members' age, experience, tenure, gender, race and area of expertise, leads to better outcomes (Cox, 1993; Jackson et al., 1995). A significant proportion of the research into demographic diversity has concentrated on top management teams and follows an upper echelons tradition which links demographic attributes of senior executive teams to aspects of organisational performance (Hambrick & Mason, 1984, Finklestein & Hambrick, 1996).

Research in this domain obviously depends upon adequate ways of measuring team diversity. It is in this respect that problems are evident, of which two are focused upon here. The first stems from the fact that background demographic variables are generally of two types, categorical and continuous, and that different indices of diversity have been used for each. Diversity on categorical variables is typically represented by Blau's index (Blau, 1977), whereas for continuous variables the coefficient of variation (Allison, 1978) is used. These give measures on very different metrics, with Blau's index ranging from 0 to 1, but the coefficient of variation able to take any value greater than or equal to 0. Given the current indices, there is no way of comparing team diversity on continuous variables with a theoretical maximum, or measuring this diversity on a scale with the same limits as that used for categorical variables.

The second problem is that although the coefficient of variation is the most widely used index for continuous variables (and indeed the premier index), it is often unsuitable in practice, owing to its sensitivity to non-normal distributions and outliers which are inherent in the small samples typical of small group and team research.

Our expressed aim in this paper is to describe a measure of diversity for continuous variables which can cope satisfactorily with these problems and which can be expressed as a coefficient ranging from 0 to 1, thus making it comparable to Blau's index for categorical variables in that it has the same maximum and minimum limits.

Diversity Measures

Traditionally, demographic (or indeed any) data can be classified in two ways. Variables such as gender and ethnicity are defined as categorical variables whereas age, organisational tenure and team tenure are continuous variables. The statistics used to measure diversity within the two classes of variables are necessarily different. In a review of 15 demographic diversity studies, Tsui et al., (1995) found that all those that included a demographic attribute measured on a continuous scale employed the coefficient of variation. Historically, both Blau's index, used to measure categorical variables, and the co-efficient of variation (Allison, 1978) were described as measures of inequality reflecting their basis in sociology and economic theory.

Blau's Index of Heterogeneity

Blau's index for categorical variables (e.g., ethnicity and gender) uses the proportion p_i of population members in a particular category i and is expressed as

$$(1 - \sum_i p_i^{\,2}).$$

Blau's index can take values from 0, meaning that all team members are the same on a particular category, to a theoretical high of 1, which represents maximum diversity within the team. This index has found ready application to social units such as work teams.

Coefficient of Variation – Allison (1978)

Allison (1978) was primarily concerned with how sociologists measured dimensions of social welfare and inequity between nation states, cities and 'other social units'. The measure proposed was appropriate for measuring variables such as income inequity and was first used with extremely large data sets. This measure has since been widely adopted by social psychologists to assess individual differences in organisations and work teams (see for example, Wagner et al., 1984; Bantel & Jackson, 1989; Michel & Hambrick, 1992; Jackson et al., 1991; Pelled et al., 1999). Moreover, it is often asserted in demographic diversity research that it is the best and "most direct measure of scale-invariant dispersion" (Tsui et al., 1996; see also Pelled et al., 1999). Similarly, Kilduff et al. (2000) judged the coefficient of variation as "superior in its psychometric properties to other measures such as the standard deviation".

The coefficient of variation is the standard deviation divided by the mean. As a measure of inequality it has the advantages of being both easy to compute and it provides a scale invariant measure of dispersion. However, it is sensitive to outliers, making it particularly inappropriate for use with non-normal data and unstable when the sample size is small. These negative considerations are particularly important to organisational team research where the number of people in teams is likely to be low. One extreme value within a small team can affect the measurement of diversity disproportionately by inflating the standard deviation and hence the coefficient of variation. It can also result in a mean team score that differs dramatically from those of other teams, hence making comparisons of the coefficient of variation across the sample suspect.

The coefficient of variation also requires the variable to have a theoretically fixed zero point (i.e., to be a ratio scale), which is not true for many attributes of interest in team research. Moreover, it does not have an upper limit and hence cannot indicate how close any sub-sample is to the maximum achievable level of diversity given the range of values within whole sample. This is in contrast to Blau's index for categorical variables, which is bounded by 0 and 1.

The Present Study

The proposed index was developed in the context of analysing data from 23 top management teams from the UK manufacturing sector. The teams are fully functioning, existing and intact teams in real organisations, as opposed to being assembled purely for the purposes of research. The teams range in size from 3 to 8 members (overall sample contains 130 individuals), some have mixed gender representation, some contain members of similar ages, some have a range of functional backgrounds whereas others do not. The purpose of the research project was to investigate the extent to which all the naturally occurring demographic differences (diversity) within teams contribute to team processes such as decision making. However, the analysis was somewhat limited owing to the problems concerning measurement as outlined above. To overcome these limitations, and provide the foundation for work including the measurement of diversity across dimensions, we sought to develop a new measure for continuous variables, and propose the following index as a result of those efforts.

Proposed Index

The proposed index for continuous team level variables is based on established statistical ranking theory. It overcomes the challenge of dealing with small sample sizes, the restriction of assuming a theoretically fixed zero-point and the normality issues by taking a non-parametric approach. It also gives us boundaries of maximum and minimum diversity relative to the whole sample from which our subgroups were drawn. Unlike the coefficient of variation, it does not use raw data values, but instead measures the diversity of each team by comparing the rankings of the team members over the whole sample on the chosen variable. The index is computed as follows:

Step 1: Rank variable X of interest within the whole sample. For example, if the total sample size is n, then each case within the sample is given a new variable *rank_X*, which will take a value between 1 and n.

Step 2: Separate the data into teams. For each team, calculate the absolute value of the difference in ranks between each pair of subjects within the team, and add a correction value of 1 to each difference score. For a team of size *m* there should be P_m of these difference scores, where

$$P_m = \frac{m(m-1)}{2}$$

Take the natural logarithm of each of these difference scores (the reason behind the correction value, since the logarithm of 0 is undefined), and then take the mean of these log(difference scores) over the team.

Step 3: Divide through by $K_{m,n}$, the maximum possible mean(log(difference scores)) obtainable for a team of size *m* drawn from a sample of size *n* to get the final score.

This coefficient can be expressed as

$$K_{m,n} = \frac{1}{P_m}\left(\log_e\left[\prod_{i=1}^{m-1}\left(1+i\left(\frac{n-1}{m-1}\right)\right)^{m-i}\right]\right)$$

This three-step process gives us a value between 0 and 1, where 0 indicates no diversity and 1 represents the maximum diversity possible within the sample.

A comparison with the coefficient of variation

The following two examples taken from real top management teams within our database, illustrate how the proposed index out-performs the coefficient of variation when mean scores differ substantially between teams or when the data for one team contains an outlier.

Table A1:
Measuring the diversity of 'Industry Tenure (months)' within two teams of 6 people.

Cases	1	2	3	4	5	6	Team Mean	Team SD	Co-efficient of variation	Proposed Index
Team 1	9	14	14	72	116	367	98.7	138.1	1.40	0.84
Team 2	6	108	132	180	312	360	183.0	132.3	0.72	0.91

The scores of the two teams have a similar range, and the teams have similar standard deviations, but there is clearly greater diversity in team 2. However, the concentration of low scores in team 1 gives it a much lower mean score, resulting in it having a higher coefficient of variance in comparison to team 2. The proposed index does not suffer from such problems and, unlike the coefficient of variation, indicates that team 2 is actually more diverse in terms of time worked in the industry.

Table A2:
Measuring the diversity of 'Organisational Tenure (months)' within two teams of 6 people.

Cases	1	2	3	4	5	6	Team Mean	Team SD	Co-efficient of variation	Proposed Index
Team 1	6	9	14	14	19	451	85.5	179.1	2.09	0.71
Team 2	6	41	49	91	101	240	88.0	82.1	0.93	0.88

In this example the team mean scores are almost identical, but team 2 is clearly the more diverse. However the whole sample distribution of 'organisational tenure' is highly positively skewed, with 10% of the 130 cases having worked for over 300 months compared to a sample median of 90. One of these extreme values occurs in team 1 resulting in a very large team standard deviation, and consequently a higher coefficient of variation than that of the more diverse team 2. Since the current index uses ranks rather than raw scores it reduces the effect of the extreme value in team 1, and takes a lower value for team 1 than for team 2.

Advantages and limitations of the proposed index

As well as its superiority over the coefficient of variation in dealing with outliers and varying subgroup mean scores, another advantage of the proposed method is that it gives a theoretical index with a fixed range. Like Blau's index it takes values from 0 (no diversity) to a theoretical high of 1 (maximum diversity) on continuous variables. Therefore we can use it to assess the diversity of a continuous variable with reference to the maximum and minimum diversity achievable.

We caution against promoting a direct comparison between Blau's index and the proposed index. Although both take values between 0 and 1, the underlying statistical computation is necessarily different, making a direct comparison non-advisable

Despite the accepted wisdom being that demographic diversity in top management teams will predict better outcomes, the research findings have been disappointingly equivocal (see West & Schwenk, 1997; Norburn & Birley, 1988; Priem et al., 1999). Moreover, patchy results are often found for different dimensions of diversity (see for example, Knight et al., 1999). We believe that the problems we have uncovered with the coefficient of variation may have influenced some of these uneven findings. In the light of our exposé, prior work that used the coefficient of variation may need to be treated with caution. We suggest that the proposed index makes a significant advance in measuring and accurately representing demographic patterns of organisational work teams and holds considerable promise for diversity research in the future.

References:

Allison, P. D. (1978) Measures of Inequality, American Sociological Review, 43, 6, 865-880.

Bantel, K.A., & Jackson, S.E., (1989). Top Management and Innovation in Banking: Does the composition of the top team make a difference? Strategic Management Journal, 10, (Special Issue), 107-124.

Blau, P. M. (1977) Inequality and Heterogeneity. Free Press: New York.

Cox, T.H. (1993). Cultural diversity in Organizations: Theory, Research and Practice. Barrett-Koehler, San Francisco, CA.

Finklestein, S., & Hambrick, D. C. (1996). Strategic Leadership: Top executives and their effects on organisations. West, St Paul, MN.

Hambrick, D.C., & Mason, P. A. (1984). Upper echelons: the organisation as a reflection of its top managers. Academy of Management Review. 9, (2), 193-206.

Jackson, S.E. (1992) Team Composition in Organisational Settings: Issues in Managing an Increasingly Diverse Work Force. In S. Worchel, W. Wood & J. A Simpson (eds.). Group Process and Productivity. Newbury Park, CA: Sage. 138-176.

Jackson, S. E., Brett, J. F., Sessa, V. I., Cooper, D. M., Julin, J. A., & Peyronnin, K. (1991). Some differences make a difference: Individual dissimilarity and group heterogeneity as correlates of recruitment, promotions and turnover. Journal of Applied Psychology, 76, 675-689.

Jackson, S. E., May, K.E., & Whitney, K. (1995). Understanding the dynamics of diversity in decision making teams. In R.A Guzzo & E. Salas (eds.), Team Effectiveness and Decision Making in Organizations. Jossey Bass, San Francisco, CA. 204-261.

Kilduff, M., Angelmar, R. & Mehra, A. (2000). Top management-team diversity and firm performance: Examining the role of cognitions. Organization Science. 11, (1), 21-34.

Knight, D., Pearce, C.L., Smith, K.G., Olian, J.D., Sims, H.P., Smith, K.A., Flood, P. (1999). Top Management Team Diversity, Group Process and Strategic Consensus. Strategic Management Journal, 20, 445-465.

Michel, J. G., & Hambrick, D. C. (1992). Diversification on Posture and Top Management Team Characteristics. Academy of Management Journal, 35, 9-37.

Norburn, D., & Birley, S. (1988) The top management team and corporate performance, Strategic Management Journal. 9, (3), 225-237.

Pelled, L.H., Eisenhardt, K.M., Xin, K.R. (1999) Exploring the Black Box: An Analysis of Work Group Diversity, Conflict and Performance. Administrative Science Quarterly. 44, 1-28.

Priem, R. L., Lyon, D.W., Dess, G.G. (1999). Inherent limitations of demographic proxies in top management team heterogeneity research. Journal of Management. 25, (6), 935-953.

Sessa, V.I., & Jackson, S.E., (1995) Diversity in Decision-Making Teams. In M.M. Chemers, S. Oskamp, and M.A. Costanzo (Eds) Diversity in Organisations: New Perspectives for a Changing Workplace. Thousand Oaks: Sage. 133-156.

Tsui, A.S., Egan, T.D., Xin, K.R. (1995) Diversity in Organisations: Lessons from Diversity Research in M.M. Chemers, S. Oskamp, and M.A. Costanzo (Eds) Diversity in Organisations: New Perspectives for a Changing Workplace. Thousand Oaks: Sage. 191–219.

Wagner, W., Pfeffer, J., & O'Reilly, C.A. III. (1984). Organisational demography and turnover in top-management groups. Administrative Science Quarterly, 29, 74-92.

West, C.T., & Schwenk, C.R. (1997). Top management team strategic consensus, demographic homogeneity and firm performance: A report of resounding nonfindings. Strategic Management Journal. 17, (7), 571-576.

A P P E N D I X   B

Flyers and letters used in attracting participants to the research.

_ _ _ _D

nstitute of Work Psychology *Incorporating* the ESRC Centre for Organization & Innovation

Mushroom Lane
Sheffield
S10 2TN
Tel: +44 (0) 114 222 3258
Fax: +44 (0) 114 272 7206
http//www.shef.ac.uk/~iwp

08 February 1998

«company»
«Address1»
«Address2» «Address3»

Attention: «Title» «FirstName» «LastName» «JobTitle»

Dear «Title» «LastName»

re: Top Management Team Development Workshop

I am writing to you concerning a unique opportunity for Top Management Teams in the UK manufacturing sector. As you may be aware, the work of the Institute is primarily concerned with discovering the factors which contribute to organizational effectiveness and innovation.

Increasingly our research is demonstrating that teams which take the time out to reflect upon their performance are more creative and effective at dealing with varied demands in uncertain and complex environments. However, with all the pressures and demands upon their time, many teams do not take the time to reflect upon their work methods and so do not appropriately modify their approaches to become even more effective.

To provide UK top managers with an opportunity for team reflection, we are conducting a custom-designed workshop with single teams, with particular reference to effective decision-making. Using a number of simulated real-world decision-making tasks, presentations and feedback from experienced facilitators, your team will gain a deeper understanding of decision-making processes and learn creative techniques for dealing with in-team conflict. Each team member also receives a copy of the book *Effective Teamwork.*

We believe that your team's participation in the workshop will be highly beneficial for both your company and the IWP research team. Because of the individual nature of the workshop, (only your team attends on a given day), we are able to arrange a mutually convenient date and location. We appreciate the challenge of getting a whole top management team together on any one day, teams which have participated usually found it best to nominate a date far enough in advance for each of the managers to diary it ahead.

I would like to phone you in a couple of days to answer any questions you may have about the workshops.

Yours faithfully

Tracey Swift
PSS Project Co-ordinator

Effective Decision-Making in Top Management Teams

This Development Workshop has been designed exclusively for Top Management Teams in the UK manufacturing sector by the ESRC Centre for Organization and Innovation.

This workshop aims to:

- provide Top Management Teams with the opportunity to **reflect on** their decision-making processes.
- share **insights** from research on Top Management Team functioning with UK managers.
- provide **immediate feedback** to teams on observed behaviour by experienced facilitators.

Benefits for your team:

- helps members to **understand team decision-making** processes
- creative techniques for **constructively managing in-team controversy** are learned
- provides a forum for **discussing team functioning**
- reflection leads to action planning for conduct of future meetings

Comments from participants in the 1997 summer series:

"We all enjoyed the day spent with you - and have already implemented some of the points learned". MD Plastics Firm 24/09/97

'We found the exercises and the feedback most interesting. From a team building point of view the day was excellent, and for the future we have taken away some practical and worthwhile ideas.' CEO, Engineering Firm 20/06/97

"Far more interesting than I thought it would be". Design Director, Engineering Firm, 27/10/97

"We would like to participate in future workshops". CEO, Electronics Firm, 12/06/97

"Extremely practical and worthwhile". MD, Engineering Firm, 22/10/97

Your investment:

You are able to choose the most convenient venue - Facilitators can bring the workshop to you, alternatively the team can visit the Institute of Work Psychology, University of Sheffield or arrangements can be made for an external regional venue. There is a nominal cost of £50.00 per head.

To nominate a date for your team, please contact Tracey Swift. Phone 0114 222 3276,

Fax: 0114 272 7206, Email: T.Swift@Sheffield.ac.uk

'HE UNIVERSITY OF SHEFFIELD
ıstitute of
'ork Psychology
Incorporating **the ESRC Centre for Organization & Innovation**

Mushroom Lane
Sheffield
S10 2TN
Tel: +44 (0) 114 222 3258
Fax: +44 (0) 114 272 7206
http//www.shef.ac.uk/~iwp

I
W
P

08 February 1998

«company»
«Address1»
«Address2» «Address3»

Attention: «Title» «FirstName» «LastName» «JobTitle»

Dear *FirstName*

re: Top Management Team Development Workshop

I am really pleased that we have been able to arrange a date for us to come and deliver the Top Management Team Development Workshop.

I confirm the date and venue as follows:

Date:

Venue:

Time: 9:00 am - 4:30 pm

I am enclosing a programme for your perusal. As you will see, we have an action-packed day ahead of us. Before the workshop, each nominated member of your team will receive an information pack in preparation for the programme.

Would you please arrange for completion and return of the enclosed form to assist us with our arrangements.

We are sure that the programme will be stimulating and beneficial for your team. We anticipate a lively interaction and look forward to meeting with you.

Kind regards

Tracey Swift

PSS Project Co-ordinator

Institute of Work Psychology

Incorporating **the ESRC Centre for Organization & Innovation**

Mushroom Lane
Sheffield
S10 2TN
Tel: +44 (0) 114 222 3258
Fax: +44 (0) 114 272 7206
http//www.shef.ac.uk/~iwp

I W P

08 February 1998

«company»
«Address1»
«Address2» «Address3»

Attention: «Title» «FirstName» «LastName» «JobTitle»

Dear «Title» «LastName»

Dear *FirstName*

re: Top Management Team Development Workshop

We are very much looking forward to visiting with you and your top management team on the *(insert date)*.

Enclosed are *(insert number)* pre-work packets for you to distribute to members of your team prior to the workshop.

The pre-work packets contain information which will orient the participants for the Peak Selection Simulation tasks during the day. The theme of the workshop is: Effective Decision Making in Top Management Teams and the focus of all the exercises will be to contribute to gaining a deeper understanding of decision-making styles.

Kind regards

Tracey Swift

PSS Project Co-ordinator

D

nstitute of Work Psychology *Incorporating* **the ESRC Centre for Organization & Innovation**

Mushroom Lane
Sheffield
S10 2TN
Tel: +44 (0) 114 222 3258
Fax: +44 (0) 114 272 7206
http//www.shef.ac.uk/~iwp

I
W
P

08 February 1998

«company»
«Address1»
«Address2» «Address3»

Attention: «Title» «FirstName» «LastName» «JobTitle»

Dear *FirstName*

re: Top Management Team Development Workshop

It was a real pleasure to visit with your top management team for the workshop yesterday.

I certainly enjoyed the day and the interaction provided by your team on the subject of effective decision-making. I hope that you and your team have gained from both the report and the reflection exercises on your decision-making processes. Feedback to us from the teams which have participated to date has been extremely positive. All have found the workshop practical and indicate that ideas have been generated by it for future use.

The Institute appreciates your continued support and aims to provide further opportunities for collaboration, possibly in a similar workshop format.

Many thanks for your involvement.

Tracey Swift
PSS Project Co-ordinator