This is to certify that the
dissertation entitled

THE RELATIONSHIP BETWEEN TEAM COMPOSITION AND
TEAM EFFECTIVENESS: A MULTI-LEVEL STUDY

presented by

JOSHUA M. SACCO

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Psychology

Neal Schmitt
Major Professor's Signature

12/11/02
Date

MSU is an Affirmative Action/Equal Opportunity Institution

THE RELATIONSHIP BETWEEN TEAM COMPOSITION AND TEAM EFFECTIVENESS: A MULTI-LEVEL STUDY

By

Joshua M. Sacco

A Dissertation

Submitted to
Michigan State University
In partial fulfillment of the requirements
For the degree of

Doctor of Philosophy

Department of Psychology

2002

Abstract

THE RELATIONSHIP BETWEEN TEAM COMPOSITION AND TEAM EFFECTIVENESS: A MULTI-LEVEL STUDY

By

Joshua M. Sacco

The research reported here examines from a personnel selection standpoint team composition and person-team fit in relation to aspects of team effectiveness in a sample of 3,454 fast-food restaurant customer service teams consisting of 190,156 employees. Team composition was examined with regard to ability, three dimensions of the Five-Factor Model of personality (Agreeableness, Extraversion, and Conscientiousness; Costa & McCrae, 1990; Digman, 1990) and demographic variables (race, sex, and age); person-team fit is examined with regard to all of these variables except ability. Team financial performance and team turnover rates were used as indicators of team effectiveness. Team composition and person-team fit were studied in relation to team performance, and turnover at both the team and individual level. The results indicated that mean levels of ability were positively, though weakly related to changes in controllable profit, though none of the ability or personality predictors had any meaningful relationship with turnover. Latent growth curve analyses indicated that racial diversity at the start of data collection was negatively associated with changes in profitability over time such that practically meaningful reductions in profitability are predicted even from even very small increases in racial diversity. A series of multilevel survival analyses indicated that the age, race, and sex-based fit of an employee within a store composition predicted the

likelihood of turnover, and that this effect was markedly stronger earlier on in employees' tenure. Similarly, age, race, and sex diversity was positively associated with restaurant-level turnover rates. The practical, theoretical, and methodological implications of these results are discussed.

Acknowledgements

This dissertation would not have been as successful as it was, or possibly completed at all were it not for the contributions of a number of people. First, Tara Lynn Victor provided an enormous amount of support and encouragement while I spent several weeks straight at the library night after night, and a similar amount of time wading through mounds of data and obscure statistical procedures. Second, Neal Schmitt has made monumental contributions to my professional growth and approach to problems in I/O psychology. His contributions to the thinking and skills demonstrated in this dissertation cannot be overstated, and range from my greatly improved ability to write clearly and concisely to the use of structural equation modeling to answer complex research questions. Much of this learning has taken place on Neal's time, in his office or working on various projects together throughout my career as a graduate student. Neal, being accessible, thoughtful and interested when I unexpectedly popped in to chat for an hour has had a huge impact on my development. Thanks are also due to Rick DeShon for his help throughout the years on various statistical issues that also helped lay the groundwork for the approaches studied here. Ann Marie Ryan provided me with a number of opportunities to conduct field research that helped me get interested in projects like this. Dan Ilgen, by reading an earlier version of this paper for his Motivation class, and by providing insightful comments during the proposal and final defense, made a number of important points that shaped my thinking about this project.

Table of Contents

LIST OF TABLES vi

LIST OF TABLES viii

CHAPTER 1
INTRODUCTION 1
What Are Teams? 3
The Input-Process-Output Approach To Teams 5
Theoretical Model Of Team Effectiveness 7
KSAO – Performance Linkages at the Individual Level 8
Individual Level KSAOs, Team Composition, and Team Performance 18
Person-Environment Fit in Organizational Research 39
Teams as a Context 46
The Present Study 75

CHAPTER 2
METHOD 79
Sample and Procedures 79
Measures 81

CHAPTER 3
RESULTS 91
The Relationship Between Mean Levels of Ability and Personality with Changes in Controllable Profit 92
The Effects of Individual Standing on Ability and Personality Variables on Turnover 94
The Relationships Between Demographic Diversity and Profitability and Turnover Rates 97

CHAPTER 4
DISCUSSION 121
Team Ability and Personality Composition and Team Performance 122
Relationships Between Ability and Personality and Turnover 126
Effects of Demographic Similarity on Turnover 130
Methodological Implications for the Dynamic Studies of Teams 132
Theoretical and Methodological Implications for Relational Demography and PO Fit 134
Conclusion 135

REFERENCES 136

APPENDICES 234

List of Tables

Table 1. Summary of Hypotheses and Planned Analytic Methods 160
Table 2. Descriptive statistics, internal consistency estimates, uncorrected and corrected correlations between the CRI and IPI scales (N = 135) 163
Table 3. Test-retest correlations of CRI and IPI scales at a one-week interval (N = 29) 164
Table 4. Race and Sex Composition of the Sample. 165
Table 5. Age composition of the sample (Valid N = 251,355) 166
Table 6. Categories of Turnover and Incidence 167
Table 7. Descriptives statistics and intercorrelations among restaurant-level variables. 168
Table 8. Hierarchical regressions for mean ability levels as a predictor of controllable profit (N =2,206). 169
Table 9. Hierarchical regressions for mean conscientiousness levels as a predictor of controllable profit (N =2,207). 170
Table 10. Hierarchical regressions for mean agreeableness levels as a predictor of controllable profit (N =2,207). 171
Table 11. Hierarchical regressions for mean extraversion and extraversion^2 levels as a predictor of controllable profit (N =2,205). 172
Table 12. Hierarchical regressions for mean conscientiousness X manager tenure levels as a predictor of controllable profit (N =2,206). 173
Table 13. Hierarchical regressions for mean agreeableness X manager tenure levels as a predictor of controllable profit (N =2,206). 174
Table 14. Hierarchical regressions for mean ability X manager tenure levels as a predictor of controllable profit (N =2,206). 175
Table 15. Hierarchical regressions for coefficient of variability in conscientiousness as a predictor of controllable profit (N =2,029). 176
Table 16. Correlations among 11-month aggregates of restaurant level demographic diversity measures and turnover rate (N = 2496-2498). 177
Table 17. Multiple regression analyses for the prediction of 11-month aggregate of turnover rates with 11-month aggregates of age, sex, and demographic diversity. .. 178
Table 18. Means, standard deviations, and intercorrelations for racial diversity across the 11 time periods (N = 2,388). 179
Table 19. Means, standard deviations, and intercorrelations for age diversity across the 11 time periods (N = 2,388). 180
Table 20. Means, standard deviations, and intercorrelations for controllable profit across the 11 time periods (N = 2,388). 181
Table 21. Intercorrelations for controllable profit and racial diversity across the 11 time periods (N = 2,388). 182
Table 22. Intercorrelations for controllable profit and age diversity across the 11 time periods (N = 2,388) 183
Table 23. Example of the period-person data structure used for the multilevel survival analyses. 184

Table 24. Results of HnLM examining the effect of individual sex X restaurant sex composition interaction on turnover probabilities. 185
Table 25. Results of HnLM examining the effect of individual sex X restaurant sex composition X time interaction on turnover probabilities. 186
Table 26. Results of HnLM examining the effect of individual race X restaurant race composition interaction on turnover probabilities. 187
Table 27. Predicted turnover hazards as a function of employee race and store racial composition. 188
Table 28. Results of HnLM examining the effect of individual race X restaurant race composition X time interaction on turnover probabilities.. 190
Table 29. Predicted turnover hazards as a function of employee race and store racial composition and time. 191
Table 30. Results of HnLM examining the effect of individual age X restaurant age composition interaction on turnover probabilities. 193
Table 31. Predicted turnover hazards as a function of employee age X store age composition interaction. 194
Table 32. Results of HnLM examining the effect of individual age X restaurant age composition X time interaction on turnover probabilities. 196
Table 33. Predicted turnover hazards as a function of employee age X store age composition X time interaction.. 197
Table 34. Results of HnLM examining the effect of individual age X restaurant age composition X extraversion interaction on turnover probabilities. 199
Table 35. Results of HnLM examining the effect of individual race X restaurant race composition X extraversion interaction on turnover probabilities. 200
Table 36. Results of HnLM examining the effect of individual sex X restaurant sex composition X extraversion interaction on turnover probabilities. 201
Table 37. Predicted turnover hazards as a function of employee age X store age composition X extraversion interaction. 202
Table 38. Summary of Hypotheses and Results 204

List of Figures

Figure 1. Hackman's IPO approach to teams 207
Figure 2. Theoretical model underlying the prediction of team performance 208
Figure 3. Theoretical model underlying the prediction of turnover. 209
Figure 4. Representation of a simple latent growth curve structural equation model 210
Figure 5. Trends in racial diversity for a sample of 15 restaurants. 211
Figure 6. Latent growth curve model for racial diversity. N = 2388 restaurants 212
Figure 7. Plot of standardized LISREL estimates of racial diversity slope factor loadings 213
Figure 8. Trends in age diversity for a sample of 15 restaurants. 214
Figure 9. Plot of standardized LISREL estimates of age diversity slope factor loadings. 215
Figure 10. Latent growth curve model for age diversity. N = 2388 restaurants. 216
Figure 11. Trends in controllable profit for a sample of 15 restaurants. 217
Figure 12. Latent growth curve model for controllable profit. N = 2388 restaurants. 218
Figure 13. Plot of standardized LISREL estimates of controllable profit slope factor loadings 219
Figure 14. Combined latent growth curve model for racial diversity and controllable profit. N = 2388 restaurants. 220
Figure 15. Predicted trajectories for controllable profit based on several different levels of racial diversity 221
Figure 16. Predicted yearly total impact of racial diversity on controllable profit 222
Figure 17. Predicted yearly controllable profit for several hypothetical racial compositions. 223
Figure 18. Combined latent growth curve model for age diversity and controllable profit. N = 2388 restaurants. 224
Figure 19. Predicted cumulative survival function based on logistic regression analysis of person-period dataset with linear and quadratic time factors. 225
Figure 20. Predicted turnover hazard as a function of sex and two different store sex compositions. 226
Figure 21. Predicted turnover hazard as a function of sex and two different store compositions and time. 227
Figure 22. Predicted turnover hazard as a function of race and three different restaurant race compositions. 228
Figure 23. Predicted turnover hazard as a function of race, three different restaurant race compositions and time. 229
Figure 24. Predicted turnover hazard as a function of age and three different restaurant age compositions. 230
Figure 25. Predicted turnover hazard as a function of age, three different restaurant age compositions, and time 231
Figure 26. Predicted turnover hazard as a function of race, two different restaurant racial compositions, and two levels of extraversion.. 232
Figure 27. Schematic summary of results 233

INTRODUCTION

Over the last few decades there has been a dramatic shift from individual-based work towards the reliance on work teams to accomplish organizational goals (e.g., Ilgen, Major, Hollenbeck, & Sego, 1995). In fact, numerous authors suggest that successful teams are critical for achieving organizational objectives (Ilgen, 1999). This has generated a huge amount of theoretical and empirical literature on the nature of team effectiveness; however, from a personnel selection perspective we still know relatively little about the impact of team composition on team effectiveness. This is despite the fact that personnel psychologists exert a strong influence on team composition and the huge body of literature that has successfully identified methods for predicting individual outcomes (i.e., job performance; Barrick & Mount, 1991; Hunter & Hunter 1984).

The few studies that have examined the prediction of team effectiveness used small samples and have not taken full advantage of the more organizationally relevant criteria that teams often produce as compared to that of individuals. Teams often produce output that can be easily measured and has immediately obvious relevance to the organization (e.g., financial performance; Schneider, Smith, & Sipe, 2000); however, the only selection-oriented studies to date have focused on team performance ratings. The ability to predict more organizationally relevant criteria would substantially enhance the utility of personnel selection (Schneider et al., 2000). This is underscored by research indicating that performance ratings are subject to a host of factors not directly related to the construct of job performance including measurement instability, rater idiosyncrasies and rater perspective (e.g., Conway, 1996; Gregarus & Robie, 1998; Mount et al., 1998; Scullen, Mount, & Goff, 2000; Viswesvaran, Ones, & Schmidt, 1996).

A number of authors have recognized two additional important shifts in today's workplace. First, the United States economy is now largely service-oriented which represents a significant change from the manufacturing emphasis of only a few decades ago; service jobs are estimated to represent 41% of the American workforce (National Research Council, 1999). Researchers have acknowledged that this is likely to lead to an emphasis on a different skill set requirement for service workers (Schmitt & Chan, 1998), yet research on this issue has appeared only recently (Klimoski & Jones, 1995). Second, the workforce is becoming increasingly diverse with regard to race, sex, and age. Many authors have discussed this trend and its numerous practical, theoretical, and research implications. The research reported here examines these issues by studying the composition of diverse teams in a service setting from a personnel selection perspective.

In particular, the large sample research reported here looks at the relationship between team composition and team effectiveness in a customer service-oriented setting. Team composition is broadly defined in terms of ability and personality constructs typically used in personnel selection, as well as the demographic variables of race, sex, and age. Team effectiveness is measured by financial outputs and turnover within the team. The use of financial performance as criteria is especially important in the present research because each team works in a fast-service restaurant that can be considered its own organization. Hence, this research has the potential to answer many questions relating to whether the knowledge we have gained in personnel selection directly relates to organizational effectiveness (Ostroff & Bowen, 2000; Schneider, 1996; Schneider et al., 2000). Team composition is hypothesized as the main predictor of team performance, while the fit within a given personality or demographic composition (i.e. mix) is used as a

predictor of turnover as well as performance. In addition, two potential moderators of the effects of demographic and personality-based fit are studied: Time and Extraversion.

This paper is organized as follows. First, teams are defined, and the dominant perspective in teams research is discussed. Second, a theoretical model is described that underlies the proposed relationships examined in this research. This model serves as an organizing framework allowing us to understand how the constructs and variables described above might be related. Third, the theoretical and empirical underpinnings of the specific linkages in the model are discussed in turn, and finally, hypotheses tested in the present study are developed.

What Are Teams?

It is important to explain what constitutes a team and to differentiate it from similar concepts. For the purposes of this paper a team will be defined as a collection of individuals working in a differentiated manner on an interdependent task in a work setting towards an organizationally relevant goal. This definition includes elements from several influential definitions (Guzzo & Dickson, 1996; Guzzo & Shea, 1992; Hackman, 1987; Morgan, Glickman, Woodward, Blaiwes, & Salas, 1986; Salas, Dickinson, Converse, & Tannenbaum, 1992) and provides a straightforward way of determining whether something is a team. This definition is composed of five key points. First, a team consists of a collection of individuals; that is, one individual does not constitute a team. While pairs of individuals are often called dyads, and while there have been studies on the effects of size on group processes (see Morgan & Lassiter, 1992; Nieva, Flieshman, & Rieck, 1978), Guzzo & Shea (1992) note that size is not a primary concern in the team research. Second, a team's members must be working in a differentiated

manner; in other words, though there may be overlap in duties there must be some qualitative difference between the duties being performed by the team members. Third, the task required of the team members must be interdependent in that at least some aspect of each team member's functioning impacts the rest of the team. Fourth, in order to be a team these individuals must exist within a work setting. This is because the focal context of the present research is on work in an organization. Last, there must be an organizationally relevant goal. The team must exist to accomplish a specific task or set of tasks impacting the organization.

It is also instructive to differentiate teams from groups. Whereas teams exist at work and are generally studied by industrial/organizational or management scholars, the latter are generally studied by social psychologists (also note that researchers use the term 'workgroup' to refer to teams as well, and that the two terms will be used interchangeably here). Teams are a specific type of group; that is, groups can occur anywhere (i.e., not necessarily at work) and do not necessarily have specific goals. As such, groups are most frequently studied in laboratory settings and naturalistic settings other than at work. This reflects different theoretical and practical orientations of the two camps, as social psychologists are often concerned with more basic group process than are team researchers, who are interested in specific types of groups in work settings (i.e., teams; Driskell & Salas, 1992). This frequently results in substantively different research questions that are asked using groups versus teams. For instance, much of the group research examines collections of individuals that work together for only very short periods of time (generally only a few hours total) and on tasks that bear little resemblance to work tasks. A number of authors have indicated that this makes it difficult to

generalize from studies of groups to studies of teams because social processes that unfold over time are absent, and the nature of the tasks are often highly artificial. Further, groups often consist of college students participating in research for extra credit, a population that frequently is far narrower than the ones studied by team researchers. Thus, although group research can inform our knowledge of team functioning one must carefully consider whether the nature of a given group and the context in which it exists allows us to make valid inferences about teams (e.g., Cohen & Bailey, 1997; Sundstrom, McIntyre, Halfhill, & Richards, 2000).

The Input-Process-Output Approach To Teams

The prevailing approach to team research and theory development follows the input-process-output (IPO) model frequently attributed to McGrath (1984). This approach suggests that team process mediates the relationship between team inputs and outputs. Hackman's (1987; Hackman & Morris, 1975) elaboration on this notion is depicted in Figure 1. As can be seen in the figure, team inputs are thought to exist at multiple levels; that is, they can be properties of individuals, workgroups, or of the environment in which the team exists. Examples of individual level factors described in Hackman's original formulation include skills, personalities, and attitudes. Workgroup-level factors include attributes that can be ascribed to the team as a whole, including size, the structure of the group and cohesiveness. Environment-level factors are influences in the broader context in which the group operates (e.g., reward structure).

Process, which is responsible for converting inputs into outputs, was originally conceptualized as "all the observable interpersonal behavior" (Hackman & Morris, 1975, p. 49) that exists within the workgroup over a period of time. This notion has since been

expanded by a number of researchers. For instance, Gladstein (1984) broadened this view by incorporating prior thinking describing team process as having either a task or team preservation function. For example, weighing individual inputs appropriately would be considered a task function, while fostering smooth interpersonal communication and relationships would be considered a preservation function. Later work has focused on more specific processes as coordination (e.g., Guastello & Guastello, 1998; Zalesny, Salas & Prince, 1995), cooperation (Swezey & Salas, 1992; Tjosvold, 1995; Tjosvold & Tjosvold, 1994), task monitoring (see Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1992), strategy formation (e.g., Tesluk & Mathieu, 1999) and communication (e.g., Sanderson & Benda, 1998). Nonetheless, process is still frequently broadly conceptualized as anything that turns inputs into outputs.

Team outputs are multidimensional and consist of three components: (1) performance, (2) affective outcomes such as team member satisfaction, and (3) the continuance of the team (i.e., or capability of team members to work together in the future). These three outputs have been collectively termed team effectiveness. Sundstrom et al. (2000) indicate that about half of the studies they reviewed used a global indicator of effectiveness such as performance or productivity. Other outcomes that were used less commonly include cohesion, satisfaction, attitudes, and behavioral outcomes such as turnover, absenteeism, accidents, and prosocial behavior.

The IPO approach to teams has received considerable empirical support, both in terms of tests of the overall linkages as well as studies focusing on only one linkage (e.g., that between team process and effectiveness; see Cohen & Bailey, 1997; Sundstrom et al., 2000). Despite the fact that a number of researchers have expanded and/or modified

this conceptualization, the current thinking is that the basic model still holds and is a useful way of approaching teams theory and research. Indeed, a number of researchers have stated that this model is widely used or that it underlies all team research (e.g., Klimoski & Jones, 1995; Morgan & Lassiter, 1992; cf. Sundstrom, de Meuse, & Futrell, 1990).

Theoretical Model Of Team Effectiveness

Now that we have clarified the nature of teams and team effectiveness, theoretical models underlying the research reported here are introduced in Figures 2 and 3 for team performance and turnover, respectively. As can be seen in the figures there are constructs and linkages that are thought to exist at both the individual and team level (represented by solid lines that go from the individual to the team level), as well as interactions between the two (i.e., cross-level interaction effects; Rousseau, 1977; 1985), and a relationship among non-equivalent constructs at two levels (represented by a dashed line). The following sections briefly describe these models. Although not all of the constructs in these models are operationalized in the present study (those that are measured are underlined) it is nonetheless important to clearly explicate the mechanisms via which hypothesized relationships supposedly occur. Similarly, these models are not presented as being exhaustive, but rather are meant to convey certain aspects of the relationships that are studied here.

Figure 2 presents the theoretical model of team performance underlying the research reported here. The first link (A) in the model suggests that individual performance is a function of KSAOs and motivation. Next, KSAOs form the basis of team composition (link B), and the <u>fit</u> between these KASOs and team composition, a cross-level effect, or

interaction between two variables at different levels of analysis (Rousseau, 1985), is thought to impact individual affective outcomes (link C). Link D suggests that these affective outcomes impact individual motivation, and thus, individual performance. Further, individual level motivation, in turn, should impact team processes (Link E). Moving to effects at the team level, composition is thought to impact team process (Link F). In other words, the nature of the team in terms of its members KSAOs should impact team processes. These team processes exert cross-level effects, termed transformation, the manner in which individual performance is converted into team performance (Link G).

Figure 3 presents the theoretical model of turnover underlying this research. First, note that two linkages are identical to those in Figure 2, namely the relationships between KSAOs and composition, and the fit cross-level effects of the two on affective outcomes (links B and C, respectively). The first new Link (I) simply suggests that there are individual KSAOs that predict individual level turnover. Link J indicates that turnover can be conceptualized at the team level as well as the individual level (i.e., individual turnover and team viability). The last link in the model, Link K, indicates that team composition should impact team viability.

This paper is organized around the predictor constructs of interest within the broader context of these two models. Taking each in turn, we review the relevant empirical and theoretical work to determine how each construct might fit into these models, formulating hypotheses along the way.

KSAO – Performance Linkages at the Individual Level

Personnel selection is devoted to the study of individual differences that can be used to select people for jobs (e.g., Murphy, 1996). A number of different constructs and testing methods are used in personnel selection including cognitive ability, personality, job preferences, biodata assessments, assessment centers, situational judgment tests, and work samples. The present study examines cognitive ability and Conscientiousness, two constructs frequently used in personnel selection. In addition, two other constructs, Extraversion, and Agreeableness are used because they have been shown to be especially pertinent in team-based customer-service environments. Accordingly, the following sections describe conceptualizations of individual job performance and explain how the predictors used in this study have been shown to fit into these models.

The Construct of Job Performance

Campbell, McCloy, Oppler, and Sager (1993) suggest that up until the 1990s individual job performance was viewed as a unitary construct, with any data indicating otherwise being attributed to statistical artifacts or unreliability. In response to this perspective, as well as the laments of many researchers about 'the criterion problem' (Austin & Villanova, 1992; Landy & Farr, 1983; Smith, 1976), Campbell et al. and Borman and Motowidlo (1993, 1997) introduced multidimensional conceptualizations of the job performance construct. Campbell et al.'s formulation consists of eight factors that can be thought of as facets of Borman and Motowidlo's two-factor model of task and contextual performance. Of these two models, the latter has by far earned the most attention in the empirical literature. Task performance is described as the technical core of job requirements for a specific job. For example, the specific task components of a software engineer's job might be writing efficient and correct software code, as well as

debugging existing code. Contextual performance, on the other hand, is seen as anything that contributes to organizational effectiveness aside from task performance. Thus, contextual behaviors are seen as less role-prescribed, relatively constant across jobs, as not contributing to the specific technical focus of the organization, and as being preceded by dispositional or volitional as opposed to ability-based constructs. Aspects of a software engineer's contextual performance might be organizational citizenship behaviors such as volunteering for organizational events, helping other coworkers, or behavior that helps to maintain cohesion within the organization, for instance.

The existing literature indicates that this approach to understanding job performance has empirical and theoretical merit as well as practical utility. Empirically, a number of studies support the distinction between contextual and task performance (e.g., Conway, 1996, 1999; Johnson, 2001; Motowidlo & Van Scotter, 1994; Van Scotter & Motowidlo, 1996). Second, these two domains of performance have been used by many researchers to make theoretical and practical contributions in personnel selection (e.g., Borman & Motowidlo, 1997; Hattrup, O'Connell, & Wingate, 1998; Hattrup, Rock, & Scalia, 1997; Motowidlo, Borman, & Schmit, 1997). This distinction is pertinent to the present study because the precursors of the two aspects of job performance are thought to be different, as is the relative importance of the two (LePine & Van Dyne, 2001). Specifically, cognitive ability is thought to be the main predictor of task performance. On the other hand, contextual performance should be very important in team-based work environments in general, as well as those emphasizing customer service interactions in particular. This is because behaviors that are critical in teamwork such as interpersonal facilitation and cooperation are thought to be among the core components of contextual

performance (e.g., LePine, Hanson, Borman & Motowidlo, 2000; Van Scotter & Motowidlo, 1996). In sum, this suggests that cognitive ability should be helpful in predicting aspects of team (task) performance, though other constructs that predict team process (i.e., contextual) performance should be examined as well. The existing literature suggests that personality is likely to be an important predictor of contextual performance (e.g., Hogan, Rybicki, Motowidlo, & Borman, 1998; Lepine & Van Dyne; Murphy & Shiarella, 1998; Hurtz & Donovan, 2000). Accordingly, the following sections describe in more detail the nature of cognitive ability and several personality constructs used in the present research.

Cognitive Ability – Performance Linkages. Cognitive ability represents the totality of intellectual resources an individual can bring to bear. It is well known that specific abilities have a strong tendency to exhibit positive manifold; that is, they are all positively correlated and the nature of these correlations suggests a single underlying factor (Ackerman & Humphreys, 1990; Carroll, 1993; Hunter, 1986; Jensen, 1980). There is a vast literature supporting the notion that cognitive ability consistently predicts individual job performance for all jobs (Hunter & Hunter, 1984; Hunter, 1986; Ree & Earles, 1992; Schmidt & Hunter, 1998; Schmidt, Ones, & Hunter, 1992). The main mediator of this relationship is thought to be job knowledge (e.g., Hunter, 1986). Although the prediction of job performance is a function of job complexity, with better prediction seen in more complex jobs, the true relationship between these two constructs is estimated to be substantial (ρs=.56, .51, .40, & .23 for high, medium, semi-skilled, and unskilled jobs, respectively; Schmidt & Hunter, 1998). Thus, with regard to the

theoretical model presented in Figure 2, it is easy to see how individuals higher in cognitive ability will achieve better individual performance outcomes.

Personality. Highly influential work by Mischel (1968) and Guion and Gottier (1965) substantially curtailed research on personality in Industrial/Organizational psychology for nearly two decades. Mischel called into question the consistency of behavior across situations whereas the latter authors concluded that personality had very little or no relationship to outcomes of interest to I/O psychologists (cf. Ghiselli, 1966). Hough & Schneider (1996) point out that by the 1980s there was a revival of interest in the study of personality (e.g., Kenrick & Funder, 1998) among psychologists in general and among I/O psychologists in particular. This has been attributed to the increased consensus surrounding the dimensionality of normal (i.e., nonclinical) personality (e.g., Digman, 1990). In particular, a five-factor model (FFM) of personality has emerged that many researchers accept as adequately capturing normal personality variation (see Costa & McCrae, 1995; Digman & Inouye, 1986; Digman, 1990; Goldberg, 1990). The positive poles of these five constructs are commonly referred to as Conscientiousness, Emotional Stability, Agreeableness, Extraversion, and Openness to Experience. While there is still some debate as to the appropriateness of this generalization (e.g., Hough, 1992; Block, 1995), it has undoubtedly focused research in personnel selection that was previously plagued by the 'shotgun approach' to research that used dozens of labels for different personality traits that typically yielded near-zero relationships with job performance. More recently researchers have tended to focus more on the five-factor model of personality constructs to test relationships driven by theory (Driskell, Salas, & Hogan, 1987) and careful job analysis (Barrick et al., 2001). This is important because it is

thought that relationships between 'non-cognitive' predictor constructs such as personality are thought to be context specific (e.g., Murphy, 1996).

Social and legal pressures to reduce racial differences in selection rates have also likely impacted the importance of personality. Because racial minorities have been found to score considerably lower on cognitive ability tests (e.g., Black-White standardized difference $\approx$ 1.00; Roth et al., 2001), there has been an increasing emphasis on other predictors such as personality because they frequently exhibit differences that are either near zero or substantially smaller (Schmitt et al., 1997). Similarly, because cognitive ability is so well studied, many researchers have also examined the validity of alternative predictors after controlling for the effects of ability (i.e., their incremental validity; e.g., Clevenger et al., 2001). This reflects the practical necessity of demonstrating the benefits of adding constructs to a test battery containing cognitive ability tests. Consequently, the hypotheses in this research that focus on team performance will examine the incremental validity of personality as well.

Conscientiousness. Despite some disagreement, conscientiousness is generally thought to consist of two facets, dependability and perseverance. Aspects of dependability include the tendency to be careful, responsible, orderly, cautious, organized and efficient (Digman, 1990; Goldberg, 1990; 1993). The other facet, perseverance, has led some to characterize conscientiousness as the 'will to achieve' (e.g., Digman, 1989). Perseverance consists of facets that come to mind when one thinks of achievement motivation (Weiner, 1979; 1985); that is, diligence, hardworking, and showing a general interest in achievement and performing tasks well. Although a conceptual distinction exists between these two facets, Mount and Barrick's (1995) meta analysis found that the

two had very similar patterns of relationships across a range of different types of job performance criteria.

Of the five factors, conscientiousness has received the most attention in the personnel selection literature. The attention conscientiousness has received is likely due to several large-scale meta-analyses conducted in the last decade reporting that conscientiousness consistently predicts job performance across a range of job types and criteria (Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001; Hurtz & Donovan, 2000; Mount & Barrick, 1995; Tett, Jackson, & Rothstein, 1991). Despite widespread concerns about the impacts that faking might have on validities (e.g., Barrick & Mount, 1996; Ellingson, Sackett, & Hough, 1999; McFarland & Ryan, 2000; Rosse, Stecher, Miller, & Levin, 1998, Viswesvaran & Ones, 1999), the estimated mean corrected validity across all criteria in Barrick and Mount's (1991) first two meta analyses was .22 and .31, respectively. The traits described above imply that Conscientiousness is a motivational rather than ability-based construct in that it captures what an individual wants or prefers to do, rather than what an individual can do (i.e., 'will-do' versus 'can-do'; also, see Barrick, Stewart, & Piotrowski, 2002). Consistent with this notion, the estimated true validities of conscientiousness in predicting criteria Mount and Barrick (1995) labeled as 'can-do' versus 'will-do' were .22 and .45, respectively. Further, measures of conscientiousness are typically uncorrelated with cognitive ability measures, meaning that the former's variance in predicting a criterion of interest is incremental. This speaks to the separate criterion space assessed by measures of conscientiousness as opposed to ability measures.

With regard to customer service and teamwork, the results of several recent meta analyses are particularly relevant. First, Frei and McDaniel's (1998) meta analysis indicated that the average uncorrected correlation between conscientiousness and several commercially available customer service tests was .42 based on a total sample size of over 90,000. These customer service scales, in turn, exhibited a substantial correlation with job performance (ρ=.50) in customer-service-related jobs (N=6,495). Mount, Barrick and Stewart (1998) also conducted a meta analysis that narrowly focused on the validity of FFM constructs measured by the Personal Characteristics Inventory in team environments and jobs requiring interpersonal interactions ($N \approx 1{,}500$, k=11 studies). The results indicated that conscientiousness predicted overall individual job performance in these jobs (ρ=.26), overall job performance in team jobs (ρ =.21) and dyadic service jobs (ρ =.29), and supervisory ratings of interactions with others (ρ =.17, and ρ =.23 in the two different types of jobs, respectively). These results for the prediction of customer service job performance are consistent with another meta analysis that examined scales explicitly measuring FFM constructs (Hurtz & Donovan, 2000; ρ=.29, N = 1,369). While the focus of the present study is on team-based jobs in customer service environments, the research described above makes it clear how conscientiousness relates to individual performance in these jobs specifically, and to job performance in general. This is consistent with link A in the model presented in Figure 2.

Agreeableness. The second personality construct studied here is Agreeableness, typically associated with being trusting, cooperative, good-natured, flexible, caring, cheerful, and tolerant (Mount & Barrick, 1995; Costa & McCrae, 1995; Digman, 1990). Barrick and Mount's (1991) meta-analysis indicated there was little relationship between

this construct and job performance both broadly, and across the different job types and rating criteria they examined (ρ=.07). This, and perhaps the lack of an expected theoretical relationship between agreeableness and job performance, has likely led to it being used less often as a predictor in personnel selection. However, the two meta analyses described above provide some support that Agreeableness should be relevant to team and customer-service based settings. Frei & McDaniel (1998) reported that agreeableness was highly related to valid, commercially available customer service scales ($\bar{r}$ =.61, N=14,786). Costa & McCrae (1995) also reported nontrivial relationships between facets of Agreeableness with J. Hogan and R. Hogan's (1986) Customer Service Orientation scale, with correlations ranging from .21 to .52. Similarly, Agreeableness consistently predicted overall job performance in the Mount et al. (1998) study (ρ =.21), though this effect was substantially higher for team-based jobs (ρ =.33) as opposed to dyadic service jobs (ρ =.13). The Hurtz and Donovan (2000) study mentioned earlier reported an estimated relationship in between these two (ρ=.19) though their sample of studies did not allow them to break this estimate down according to the type of customer service job studied. These results are not inconsistent with the earlier large-scale meta analysis of Barrick and Mount (1991) because these types of jobs were not heavily represented or examined as a moderator in the earlier research. These results are also logical because these specific types of jobs heavily emphasize positive interpersonal interactions. The likelihood of positive interactions, in turn is a core component of Agreeableness' definition as a construct. So much so, in fact, that failure to find a relationship in these contexts would seriously call into question either the theoretical or operational specifications of one or both of the constructs used in these studies. In other

words, people who are agreeable, by definition, should be more likely to respond positively to customer complaints (i.e., to be caring), to be friendly (i.e., to be cheerful), to be open to other perspectives that are inevitable in team settings (i.e., to be flexible and cooperative). Indeed, Pugh (2001) reported that employees' positive emotion evokes similar responses in customers. This suggests that agreeableness fits into the model with regard to job performance as Figure 2 suggests.

Extraversion. The third construct studied here, is very similar to Eysenck's (1947) Extraversion/Introversion and Tellegen's (1982) Positive Affectivity. Extraversion is associated with being sociable, talkative, gregarious, assertive, active, and energetic. Hence, this construct can be well captured by the extent to which someone is likely to be active in social situations. Like Agreeableness, Extraversion showed a very small relationship with overall job performance in the original meta-analysis by Barrick and Mount (1991; ρ =.13) though it showed stronger relationships with individual job performance in team jobs (Mount et al., 1998; ρ=.22). In contrast to Agreeableness, however, Extraversion exhibited a substantially larger relationship with performance in team settings (ρ =.22) as opposed to dyadic customer service settings (ρ=.07). Similarly, Extraversion exhibited a small relationship with the commercially available customer service tests studied by Frei and McDaniel (mean $\bar{r}$ =.07, N=28,502; though these authors also report that excluding a particular test which specifically excludes extraversion substantially increases this relationship to $\bar{r}$ =.22). These authors suggested that talkative customer service employees might unnecessarily extend interactions due to their gregariousness, though they emphasize the need for future research on this issue. Hurtz and Donovan's estimate of the validity for customer service jobs was closer to the

lower bound of these figures (ρ=.11)._Thus, the extant literature suggests that extraversion is important for individual performance in teams though its importance in customer service jobs seems to be an open question.

Summary of Individual-Level Predictor-Job Performance Relationships.

The model depicted in Figure 2 was designed to provide an overall framework that depicts construct relationships underlying the hypotheses examined in this research. The literature described above indicates that the predictor constructs largely relate to performance given the context of this research (i.e., customer service teams). In particular, cognitive ability and conscientiousness were described as having consistent relationships to job performance in all jobs. Agreeableness and Extraversion, on the other hand, do not seem to be universally related to job performance, but nonetheless are related to job performance in team and customer service settings. These findings provide the basis for the following section, which examines the relationships between team composition on these constructs and team performance.

Individual Level KSAOs, Team Composition, and Team Performance

Link B of the models presented in Figures 2 and 3 indicate that individual KSAOs may have an equivalent construct at the team level called composition. Clearly, one can have composition in terms of ability, personality and demographics, for instance, though in this figure the type of composition is not specified because the intent of this figure is to formulate a general framework. To reiterate, the emphasis on the importance of team composition in this research is because this is the aspect of the IPO model over which personnel selection practitioners exert the most control; that is, they determine the particular KSAOs measured in a given selection context (Klimoski & Jones, 1995;

Morgan & Lassiter, 1992). Team composition generally refers to the makeup of a team as a whole as a function of the individual team members' attributes. This statement reveals two critical considerations with regard to team composition: (1) The attributes considered, and (2) the way in which these attributes combine to yield something meaningfully descriptive of the team as a whole. These two issues are discussed in turn below.

First, one must consider team composition with regard to *what*. Researchers have studied the effects of almost every individual difference one can imagine as elements of team and group composition. Team researchers have examined composition with regard to functional background in top management and product development teams (e.g., Ancona & Caldwell, 1992) and cognitive ability and/or personality in a variety of different types of teams (e.g., Barrick, Stewart, Neubert, & Mount, 1998; Barry & Stewart, 1997; Bond & Shiu, 1997; Devine, 1999; Lepine et al., 1997; Neuman, Wagner, and Christiansen, 1999; Neuman & Wright, 1999; Tziner & Eden, 1985). Group researchers, on the other hand, have studied composition variables including group size, group member opinions, ability and personality, and attitudes (e.g., see Bettenhausen, 1991; Cohen & Bailey, 1997). The criteria of interest in these studies are almost invariably aspects of group process or performance. Group and organizational scholars have examined composition as it pertains to a wide array of demographic variables including age, tenure, sex and race, generally studying performance or turnover as the primary outcomes of interest. This high volume of research speaks to the acceptance of team composition as an important determinant of team effectiveness. Because the

emphasis here is on the prediction of team effectiveness from a personnel selection standpoint, KSAO composition is the main focus in this section.

A number of researchers have noted the importance of carefully specifying composition models describing the specific theoretical form of the relationships of supposedly equivalent constructs at multiple levels of analysis (e.g., Chan, 1998; Kozlowski & Klein, 2000; Rousseau, 1985). Failing to do so has been shown to result in a host of problems that cloud valid inference including disconnects between the level of theory and analysis (Chan, 1998), and misleading data analysis results (Glick & Roberts, 1984; Rousseau, 1985). This specification involves articulating the conceptual meaning of the higher level construct which should in turn dictate the way in which the lower level data are aggregated to represent an indicator of the higher level construct (Chan, 1998; Rousseau, 1985). This is not to say that there also could not be meaningful measures of the focal construct at the level at which it is said to exist, but merely that there are oftentimes practical or theoretical constraints that dictate otherwise. Here, aggregation is necessary because KSAOs are measured at the individual level in personnel selection. Thus, we must describe a composition theory that specifies the form of the relationships between individual KSAOs and team composition. Chan presented a typology of composition models that articulates a number of possible ways in which composition might be conceptualized and operationalized. Again, properly conceptualizing composition models allows one to choose an aggregation method that is consistent with the conceptualization.

It appears to be accepted as a fact among team researchers that the nature of the task is critical in understanding how individual performance relates to team performance

(McGrath, 1984). Assuming that KSAOs are related to individual performance, this allows researchers to formulate composition models specifying the way in which aggregated KSAOs might relate to team performance. Of several task typologies (Herold, 1978; McGrath; 1984; Shaw; 1971), Steiner's (1972) has received the most attention. Steiner described three major task types: Additive, Conjunctive, and Disjunctive. Additive tasks are those in which the sum of team members' performances equals the performance of the team; conjunctive tasks are those where the lowest-performing team member defines the overall performance of the team. Conversely, disjunctive tasks are those in which the best team member's performance defines the team's performance. These three task types suggest that the average, minimum, and maximum team member KSAOs would be appropriate representations of team KSAOs, respectively, if one were interested in predicting team performance. The task of fast-service restaurant teams, which are the focus of this study, fits most closely into the additive model, though it is not strictly additive (see the Method section for more details). This suggests that the mean of team member KSAOs represents the most appropriate operationalization of team performance. First, the scope of the team task is relatively large and the task is highly interdependent and thus generally speaking no single team member will have a large overall impact. Second the task does not require specialized skills and individual members exhibit horizontal redundancy in that team members can perform each other's duties (see Lepine et al., 1997). While some research has emphasized the ability of a specific team member, this is only appropriate when a specific team member occupies an especially important role within the team (e.g., team leader; Lepine et al. 1997) or if the ability of one team member can dramatically impact

the overall performance of the team (e.g., Laughlin, 1980). The nature of the task suggests this is not the case here.

The most common ways to index the entire team's standing on a KSAO would be with the mean and the variance. As Chan (1998) and others (e.g., Neuman & Wright, 1999) indicate, multiple operationalizations could be appropriate given theories to support those operationalizations; that is, operationalizations might differ according to the specific KSAO under consideration, or the focus of a particular theory that specifies the process via which the operationalization might relate to another construct of interest (e.g., Kozlowski & Klein, 2000). Hence, the sections below discuss the way in which the mean KSAO level of team members should relate to team performance; the importance of the variance of team member KSAOs is discussed later.

Ability Composition and Performance.

There is research and theory to support the notion that the mean level of cognitive ability within a team represents a meaningful construct. We will call this construct team cognitive resources (TCR). Williams & Sternberg (1988) labeled the same notion as 'group intelligence' described as "the functional intelligence of a group of people working in a unit." (p. 356). There are a number of reasons to think that TCR should relate to team performance (Arthur, Tubre, & Hanson, 2000). First, TCR should be related to the quality of the individual performance exhibited within a team. This is because teams that have a high TCR have members who are likely to exhibit high performance because ability is consistently related to performance. Second, teams high in TCR should also exhibit processes that allow them to better translate individual output into team performance (link F in Figure 2; Lepine et al., 1997). This is because a number

of process variables are thought to have cognitive components, including task monitoring, coordination, and integration of individual performance products. This is consistent with research reported by Stevens and Campion (1994), who developed a test designed to measure teamwork KSAs (i.e., a test to predict effective team process). This test's estimated true correlation with a general cognitive ability measure approached unity, and test scores consistently predicted supervisors' and peers' ratings of individuals' teamwork skills. Further, the cognitive ability subtests also tended to consistently predict supervisory ratings of individuals' teamwork skills, though the correlations with peer ratings were not significant.

Rather than a direct, linear effect from team process to performance, Figure 2 further suggests that team process moderates the relationship between individual and team performance (Weingert, 1997). Recall the earlier discussion in which team process was described as the transformation of team inputs into outputs. Link G implies that teams with better process yield better team outputs than teams with less effective process for a given level of individual performance. In other words, individual performance should make a more direct contribution to team performance when team process is optimal. This is because team process, by definition, ensures that the individuals focus on producing the correct inputs, that the inputs are coordinated, and that they fit together according to the demands of the task. This is especially true if a given task is highly interdependent.

This is consistent with the results of a classic study of team composition effect conducted by Tziner & Eden (1985). These researchers examined the effects of 2-person tank crews' cognitive ability composition on commanders' ratings on tank performance. These authors concluded that the performance of crews composed of two high ability

members was substantially higher than their individual abilities alone would have yielded. This effect is consistent with the logic described above; that is, high ability teams developed effective coordination processes that enabled their overall product to be superior to that of the individual members.

Research by Lepine et al. (1997) is also consistent with the Transformation link in Figure 2. These researchers examined the effects of leader cognitive ability and that of the staff in a cognitively complex laboratory simulation where staff cognitive ability was operationalized as the lowest-ability member (their task was conjunctive; hence, this was an appropriate operationalization). Studying hierarchical decision-making teams, they found that these two ability levels had a multiplicative relationship with team decision accuracy. In particular, staff cognitive ability was positively related to accuracy when leader cognitive ability was high, but there was a slight negative relationship when the opposite was true. This is consistent with the model in Figure 2 because the leader in their task is in charge of integrating and transforming the recommendations of the individual staff members into a team performance product. Hence, in this situation, leaders cognitive ability can be conceptualized as a strong indicator of the quality of team process (as the task of integrating the recommendations is cognitively complex), thus supporting link G in Figure 2. Similarly, several laboratory (Devine, 1999; Williams & Sternberg, 1988) and field studies (Barrick et al., 1998; suggest that mean team level cognitive ability relates to team performance. Thus, the following hypothesis is proposed:

> H1. Mean cognitive ability will be positively related to team performance.

Personality at the Team Level

Numerous authors have indicated that the personality characteristics of team members should relate to team process and effectiveness (e.g., Fleishman & Zacarro, 1992; Hogan, Raza, & Driskell, 1988; Moynihan & Peterson, 2001). Though much laboratory research has been conducted on this issue, this research has used so many different personality traits and tasks that Jackson (1992) concluded that no firm conclusions could be drawn (also see Levine & Moreland, 1990; Morgan & Lassiter, 1992). This research is also hard to characterize because it lacked the organizing framework provided by the FFM (Barrick et al., 1998a). Thus, the interested reader is referred elsewhere for reviews of this research (Jackson 1992).

Examining personality at the team level by aggregating individual personality scores does not necessarily imply that such scores described the team's personality (Lepine et al., 1997). Rather, it merely suggests that aggregating individual personalities tells us something meaningful about the people in the team itself (Lepine et al., 1997). This is likely to be fruitful to the extent that personality is related to either individual performance or team process in that either of these relationships implies that team personality and team performance might be related. The specific ways in which aggregated personality might relate to team performance is discussed below.

Conscientiousness Composition and Team Performance. Earlier in this paper the nature of conscientiousness at the individual level was described. The discussion now turns to the notion of conscientiousness at the group level. One of the steps in adequately specifying a composition model is theory relating the composed construct to other constructs of interest. The personality levels of a group as a whole have been cited by

numerous authors (e.g., Driskell et al., 1987) as likely determinants of team process and outcomes, and several authors have indicated this should be the case for conscientiousness in particular (Barrick et al., 1998; Barry & Stewart, 1997; Bond & Shiu, 1997). The mean level of conscientiousness should be related directly to team performance and process, and one would expect there to be indirect effects as well. In terms of direct effects, Barry and Stewart (1997) argue that a conscientious team should have less process loss (Steiner, 1972) due to free riding and social loafing, two phenomena commonly examined by social psychologists in which group members withhold effort from group tasks (Forsyth, 1999). There is a large body of research indicating that these effects do indeed occur, and due to the task-oriented nature of people who are conscientious, team processes should be less likely to suffer from these effects when a team's members are high in conscientiousness. Similarly, teams with high levels of conscientiousness by their nature should attend more diligently to the importance of team process. This is because teams such as these are composed of individuals who are by their nature interested in efficiency and producing work that is of high quality. Thus, there are several reasons to believe that a team high in conscientiousness should display more effective team process. Given the research described earlier, conscientious teams should also exhibit better performance via the indirect effects on individual performance shown in Figure 2. Similarly, the individual work products should be more effectively compiled and coordinated into effective team performance.

Very few studies have examined the role of conscientiousness at a team level. Barry and Stewart (1997) hypothesized that the proportion of highly conscientious members of a student group would positively relate to task focus, cohesion, group process and

performance; however, none of these hypotheses were supported. This may be because they reduced the variance of their team level conscientiousness variable by examining a proportion of members exceeding a particular threshold. Equally odd, individual level conscientiousness scores were not associated with individual performance. These researchers mentioned several potential problems with the measures and criteria they used as other possible explanations for the unexpected results.

Barrick et al. (1998) examined the relationship between the mean level of conscientiousness in a sample of 51 assembly and maintenance teams in manufacturing plants. In this research, the correlation between mean conscientiousness score and supervisory ratings of team job performance was significant (r=.26), which was consistent with the hypotheses. Further, as would be expected by the typically near-zero correlation between conscientiousness and cognitive ability, the incremental validity of conscientiousness was very similar to its zero-order validity. Neuman et al. (1999) also examined the team mean conscientiousness scores as a predictor of small teams' performance in automotive stores. Their results were consistent with those of Barrick and colleagues, though the correlation obtained in this sample was somewhat higher (r=.40) for performance ratings from two different sources.

Neuman and Wright (1999) reported similar results in their sample of 79 human resources service teams; however, these researchers examined the minimum of conscientiousness scores within the team rather than the mean. This was because the teams' task was described as conjunctive and required distributed expertise (Ilgen et al., 1995), implying that the team's performance should be a function of the lowest-performing member.

In summary, only 3 studies have examined the relationship between mean conscientiousness scores and team performance. All three utilized small samples (though fairly large in the teams research literature), and 2 reported evidence of the hypothesized relationships between conscientiousness and ratings of team performance. A fourth study, which examined the team minimum on conscientiousness due to the nature of the team task, also reported a significant relationship between this index and team performance. One of these studies reported that conscientiousness added incremental validity above that provided by cognitive ability. Thus, while the research in this area certainly suggests that conscientiousness does indeed predict team performance, additional research is needed. Specifically, the research reported here examines team financial performance, as well as other aspects of team performance that should be closely linked to the overall conscientiousness of a team. The value of using criteria specifically targeted to predictor constructs has been widely acknowledged; however, none of the research described above has either used financial performance or specifically targeted performance outcomes as criteria. While some may argue that financial performance is not an outcome that is specifically targeted towards any predictor constructs, given that the end goal of business is to make money, and that the major role of personnel selection is to select employees for businesses, it seems reasonable to say that profitability is a very appropriate (though admittedly stringent) criterion for personnel selection research. Thus, it is hypothesized that:

> H2. The team mean conscientiousness score will be positively related to team performance.

H3. The team mean conscientiousness score will be positively related to team performance after controlling for the effects of team cognitive ability.

Agreeableness Composition and Team Performance. Based on the results of the meta-analysis by Barrick et al. (1998) one would expect that teams with higher levels of agreeableness would perform better. This is because such teams should experience more positive interpersonal interactions, thus allowing their teamwork to result in better translations of individual outputs into team performance. Research conducted by George (1990) provides some indirect support for this notion. In her research, sales teams' negative affective tone was strongly and negatively related to prosocial behavior within the work team ($r = -.57$). While it is easy to see how negative affective tone might be negatively related to high levels of agreeableness in a team, unexpectedly, this construct was not related to team performance. Similarly, positive affective tone was unrelated to team performance as well. Thus, this study provides conflicting messages about the importance of team level agreeableness.

The Barrick et al. (1998) and Neuman et al. (1999) studies discussed earlier, however, provide stronger support for the notion that mean levels of agreeableness should be related to team performance. These researchers found that the mean level of agreeableness was positively related to team performance ($r=.34$, and $r=.41$, respectively). These results are consistent with those of Neuman and Wright (1999) who reported that minimum team agreeableness on a conjunctive task was positively related to team performance and team interpersonal interactions. In both of these studies agreeableness was essentially unrelated to cognitive ability, and thus these constructs provided validity that was incremental to that of ability.

In contrast, Barry and Stewart (1997) found that mean levels of agreeableness in MBA student workgroups were unrelated to their measures of group process and performance. One explanation for this effect may lie in the single rating of overall performance that served as the criteria in this study. Based on the results described above, the following hypotheses were proposed:

> H4. Mean level of team Agreeableness will relate to team performance
>
> H5. Mean level of team agreeableness will relate to team performance after controlling for cognitive ability.

The impact of the minimum level of agreeableness within a team will also be examined. This is because it has been suggested that a single disagreeable team member can seriously disrupt the functioning of a team (Barrick et al., 1998). However, because the teams studied here are fairly large (i.e., generally more than 10 employees), it is not clear whether the impact of a single disagreeable team member will be substantial (i.e., detectable). Thus, these relationships will be examined in an exploratory fashion.

Extraversion Composition and Team Performance. The form of the relationship between extraversion and team performance might be less straightforward than that proposed for agreeableness and conscientiousness. There are several theoretical and empirical findings suggesting that extraversion should impact team performance via an indirect effect through team process. First, recall from the earlier discussion of several meta analyses that extraversion did not consistently relate to overall job performance unless the job was in a team setting. This may be because the criterion space in team settings includes team process relevant contributions whereas this is not the case for individually

or dyadically based jobs. Second, extraverts by their nature prefer interpersonal interactions (Moynihan & Peterson, 2001), and communication has been shown to be an important part of team process; hence, as team process is related to team performance, one would expect that Extraversion should be related to team performance. Third, Jung and Sosik (1999) recently found that a group's overall preference for group work predicted team performance in group decision-making tasks after controlling for several other relevant variables. Campion, Medsker, and Higgs (1993) found that the same preference predicted satisfaction within the group. Preferring group as opposed to individual work settings is emblematic of preferences that extraverts might hold in work settings.

Indeed, there are several studies that have examined empirically the relationship between team extraversion and performance in work settings. Barrick et al. (1998) found that the mean levels of extraversion were unrelated to team performance. In contrast, team extraversion was positively related to team social cohesion yet negatively related to team conflict. Similarly, neither Neuman and Wright (1999) nor Neuman et al. (1999) reported a significant relationship with their operationalization of team extraversion and team performance in human resource teams. These results are not consistent with those of the Barrick et al. (1998) meta analysis, nor are they consistent with the theoretical underpinnings described above that suggest extraverts should perform better in team settings.

In contrast, Barry and Stewart (1997) hypothesized and found nonlinear relationships between extraversion and task focus and performance. In particular, the proportion of high extraverts in a team had a U-shaped relationship to task focus, which was contrary to

their expectations; the prediction of performance showed an opposite pattern, which was expected. These results indicate that teams that are either relatively introverted or extroverted exhibited lower task performance than those with moderate amounts of extraversion. Levels of extraversion that are too low may be detrimental to team performance because members may not interact, while teams very high in extraversion may socialize too much and thus do relatively little work. This is consistent with what one might expect in a large group of extraverts and also might explain the null results obtained by three of the studies described above. (Interestingly, this is one of few cases where increased supplemental fit within a group is posited to have negative consequences for the organization.) Hence, it is hypothesized that:

> H6. Mean levels of team extraversion will exhibit a nonlinear (inverted U-shaped) relationship with team performance.

The hypotheses described above all relate to the team's average ability or personality as a predictor of team performance. The research of Lepine et al. (1997) described earlier suggests that characteristics of team members with strong influences on team performance (e.g., a team leader) interact with team KSAO levels to predict performance. In particular, team ability or personality was more strongly related to team performance with a high-ability leader. While there are no team members in fast-service restaurant teams that have as much structural influence as a team leader, the manager of a restaurant can be thought of as occupying a similar position. This is because the manager has a great deal of control over many aspects of the team's work environment that are likely to shape team effectiveness. Thus, effective managers might develop work processes that allow team members' abilities to be more strongly expressed in terms of performance

(e.g., Schneider et al., 2000). For instance, managers with longer tenure might have stronger HR systems in place (Ostroff & Bowen, 2000) that remove barriers to the expression of abilities and skills. Conversely, less effective managers might create work environments that do not facilitate the translation of team abilities into performance. Because job experience is consistently related to performance, we expected that the managerial tenure would impact team KSAO-performance relationships in the following manner:

> H7. Team manager tenure will moderate the relationship between mean team KSAO levels and team performance such that there is a stronger relationship between KSAO levels and performance for teams with high-tenure managers than with those with low tenure managers.

Turnover at the Individual Level

Turnover is commonly defined as an employee leaving an organization. Researchers usually do not include promotions, demotions, or intra-organizational transfers as turnover, although these job changes may be of importance in their own right. Turnover has numerous implications for individuals and organizations. For instance, a substantial amount of research and theory has addressed the psychological and economic well being associated with employee turnover (Leana & Feldman, 1994), as well as that of those who remain on the job after large waves of turnover (e.g., Brockner, 1986; 1987; 1988; 1994; Gowan, Riordan & Gatewood, 1999). Turnover also impacts organizations in terms of increased training costs, lowered effectiveness, demoralization, and lost knowledge. Given the array of impacts that have been associated with turnover, it is not surprising that there is a large body of literature examining the prediction of turnover or intent to turnover including individual difference, contextual, and attitudinal variables.

There are also numerous theories that seek to explain why individuals leave jobs. A review and integration of these theories is beyond the scope of this paper and is presented by Griffeth & Hom (1995). These authors present an integration of these theories based on available theory and meta-analytic evidence. Job satisfaction, commonly conceptualized as an employee's global attachment to the job, is one of the central antecedents of turnover in this integrated model, and the main pathway is via withdrawal cognitions (i.e., thinking about turning over). Indeed, several recent meta-analyses have indicated that the estimated true correlation between global job satisfaction and turnover is approximately -.20 (Link H in Figure 3; Tett & Meyer, 1993; Griffeth & Hom, 1994; Griffeth, Hom, & Gaertner, 2000). While this correlation might sound trivial, it is important to remember that turnover is a dichotomous variable, which limits the degree to which other variables can correlate with it. Similarly, even relatively weak prediction of turnover can have substantial implications for an organization because of its wide-ranging impact.

Constructs Used in Personnel Selection as Predictors of Turnover.

Despite the large body of literature devoted to examining the impact of individual-level predictors of turnover there has been relatively little emphasis on constructs typically used in personnel selection. This may be the case either because there is little theoretical justification for such research, or because turnover is not typically used as a criterion in personnel selection. In personnel selection, cognitive ability by far has received the most attention; however, it is not obvious why one would expect there to be a relationship between cognitive ability and turnover. Indeed, in a recent meta-analysis, Griffeth et al. (2000) reported the average corrected correlation between cognitive ability

and turnover to be only -.07. Only seven studies in their meta analysis examined this relationship, as opposed to 67 that examined job satisfaction. While job performance has been shown to consistently relate to turnover, and cognitive ability to job performance, the relationship between the former set of variables may be nonlinear (Trevor, Gerhart & Boudreau, 1997), and thus any relationship between cognitive ability and turnover is likely to be complex. Thus, the following sections discuss ways in which the personality constructs studied here might impact turnover. These discussions are preceded by a brief discussion of the distinction between voluntary and involuntary turnover.

Voluntary versus Involuntary Turnover.

Researchers often draw a distinction between voluntary turnover (VT) and involuntary turnover (IVT). VT is widely regarded as the appropriate criterion when examining motivational antecedents of turnover because these approaches are chiefly interested in what leads an employee to choose to leave a job. The rationale for treating these two phenomena differently is highlight in the following:

> An instance of voluntary turnover, or a quit, reflects an employee's decision to leave an organization, whereas an instance of involuntary turnover, or a discharge, reflects an employer's decision to terminate the employment relationship...To treat quits, discharges, and total turnover as synonymous ignores the markedly different etiologies and effects of these phenomena. In an organization with high quit rates, for various reasons employees find it more attractive to leave than to stay. In an organization with high discharge rates, however, presumably incorrect hiring decisions are remedied through termination (Shaw, Delery, Jenkins, & Gupta, 1998, p. 511-512).

Consistent with this argument, in their organizational-level analyses, these researchers found that the two types of turnover were distinct (though related), and had

different precursors. Despite the near consensus of the important distinction between VT and IVT there has been disagreement and difficulty in drawing an operational distinction between the two. For instance, in a review of the turnover literature, Griffeth and Hom (1994) point out that turnover due to pregnancy is classified as VT or IVT by different researchers, as is leaving a job at the insistence of one's spouse. Similarly, in practice it has proven very difficult to disentangle the two types of turnover. This is true when using personnel records' reasons for turnover due to a host of issues relating to litigation, unemployment benefits, the reputation of a terminated employee, for instance. Similarly, employees' own accounts of reasons for turnover may also be inaccurate for a variety of reasons including social desirability, embarrassment, and a lack of clarity, among others.

Nonetheless, IVT still remains a topic worthy of study. This is because organizations incur costs of turnover regardless of which party terminates employment. Though one might think IVT would yield long-term organizational benefits (as terminated employees are supposedly either performing poorly or perhaps are engaging in misconduct at work) it would be optimal if such employees could be identified before they are hired. Choosing an appropriate employee (i.e., one who is less likely to be involuntarily terminated) should reduce costs to the organization associated with hiring and training replacements. Similarly, changing jobs also causes disruption in employees' lives, so achieving a better fit would seem to benefit both parties in the employment relationship. Poor job performance or other misconduct that leads to involuntary termination might be at least in part due to underlying motivational components that can be identified during selection. Thus, it would seem that a motivational approach to turnover might also examine IVT though this has largely been ignored in the turnover literature.

For instance, Barrick and Mount (1994) located only four studies examining antecedents of involuntary turnover (LaRocco, Puch & Gunderson, 1977; Stumpf & Dawley, 1981; Wanous, Stumpf, & Bedrosian, 1979; Wells & Muchinsky, 1985). The authors indicated these studies tended to focus on demographics and job performance as antecedents, reporting consistent demographic differences, as well as a tendency for IVT to be more prevalent among poor performers. None of these studies examined constructs typically used in personnel selection. Barrick and Mount filled this void by examining the role of demographics, conscientiousness and supervisory ratings of job performance, and job involvement as predictors of IVT during downsizing. This research found that the relationship between conscientiousness and IVT was moderate (r=-.33), and the results suggested that this effect was completely mediated by job performance.

The results described above are promising in that they suggest that conscientiousness should be related to IVT in at least certain situations; however, the context of the present research is very different. First, the focus here is teams. A number of researchers have suggested that personality characteristics are more likely to be relevant in team-based work settings as opposed to jobs that are individually based. Second, from a selection standpoint it is important to identify employees likely to turnover involuntarily due to reasons other than downsizing. In downsizing many employees who would otherwise be retained are terminated, generally due to an organization's need to cut costs or adverse market conditions (Cameron, 1998); however, the termination of an employee outside the context of downsizing clearly suggests something problematic about that employee's job performance or work behavior more broadly. Based on the traits associated with conscientiousness, one would expect individuals high in conscientiousness to be less

likely to be terminated outside of a downsizing event (Rosse & Noel, 1996). Hence, in non-downsizing contexts, conscientiousness should be more closely linked to IVT. First, conscientious employees should be less likely to engage in gross misconduct, poor performance, or other unreliability that would lead to IVT. For example, Mount and Barrick (1995) reported a mean corrected correlation between conscientiousness and criteria they termed 'employee reliability' to be .41; the corrected correlation with criteria they termed 'effort' was .51. Similarly, conscientiousness seems to embody traits that would lead one to remain on the job. In other words, conscientious individuals are risk-averse, should be less likely to act impulsively or recklessly (Baron, 1996), and have a desire to work hard. Indeed, these characteristics of conscientiousness are the basis for its inclusion in most personality-based integrity tests (e.g., Schmidt, Ones, & Hunter, 1992). It is somewhat surprising, then, that Barrick and Mount's (1991) meta analysis reported a corrected correlation between conscientiousness and turnover/retention to be only .12. The only recent research I have located that examines the relationship between conscientiousness and turnover reported a correlation of -.23 in a sample of approximately 300 long-haul truckers (Barrick & Mount, 1996); I have found no research examining the role between conscientiousness and IVT. Based on this, the following hypotheses are proposed:

H8. Conscientiousness will be negatively related to VT.

H9. Conscientiousness will be negatively related to IVT.

Though it is predicted that conscientiousness will relate to turnover, there is relatively little reason to believe that the two other personality constructs studied here, extraversion

and agreeableness, relate to turnover at the individual level. Thus, the available evidence suggests that Link I in Figure 3 does not apply to all the constructs studied here.

Turnover at Higher Levels of Analysis

The discussion thus far has examined the notion of turnover at the individual level. Since psychologists tend to study individuals, it is not surprising that psychologists have paid relatively little attention to turnover at the group or team level (i.e., turnover rates). Organizations often examine the performance of their units, and thus it is clear why the turnover rate within a given unit is of interest. Clearly, the turnover rate for an organizational unit is the aggregation of individual turnover occurring within that unit (Figure 3, link J); however, there may be group-level predictors of turnover rates that do not exist at the individual level (George & Bettenhausen 1990; Rousseau, 1985). Understanding such relationships would not only further our ability to predict organizationally relevant phenomena, but also to gain insight into the processes that underlie turnover in teams. The approach taken in the research reported here is that an employee's fit within a given team is a potential precursor to both individual level turnover and turnover rates within teams. Thus, the following section broadly describes the research on fit within various work environments, becoming more detailed at the team or workgroup level of analysis.

Person-Environment Fit in Organizational Research

Individuals are embedded in a number of contexts that shape their behavior. At work, a person can be embedded in all of the following environments: A dyad, team, unit, department, division, organization and vocation. A great deal of research has been conducted examining the influence of the fit between employees and these environments.

Most of this research in the organizational sciences has specifically examined person-organization (PO) fit. PO fit is often described as having two key components: Complementary and supplementary fit (Kristof, 1996). In complementary fit, the organization and/or the employee provides what the other needs. Supplementary fit occurs when there is a similarity between the organization and the person. Fit is thus defined as one or both of these two types of matches, which can occur on a variety of dimensions including values, attitudes, personality, goals, ability and demographics (Kristof, 1996). The impacts of PO fit have been examined throughout the relationship an employee has with an organization (i.e., job choice, socialization, work behaviors and attitudes, and turnover).

There are several frameworks that provide theoretical support for the notion that PO fit should be related to positive outcomes at work. One of the most well known theories is Holland's (1985) theory of vocational choice. This theory states that over time, people choose work environments that are compatible with their orientations. The Theory of Work Adjustment (Lofquist & Dawis 1969) goes one step further by predicting that the fit between the person and the work environment directly impacts job tenure. Two other influential models are more specific in their predictions. Schneider's (1987) attraction-selection-attrition (ASA) framework states that people are attracted to, selected into, and turnover from jobs as a result of the similarity between their own personality and that of the other organizational members. Over time, this should result in organizations that are relatively homogeneous with regard to personality. In social psychology, the similarity-attraction paradigm (Berscheid & Walster, 1978; Bryne, 1971) states that similarity of all types leads to liking and attraction, which in turn leads to a host of positive outcomes.

The cumulative body of empirical research strongly supports these theories as providing powerful rationales underlying the notion that the fit between an individual and his or her work environment predicts important individual and organizational outcomes. Clearly, some of these theories are more important than others with regard to different types of fit. Thus, a brief overview of the PO fit literature is below, while more thorough explications of the relevant theories and research are given when the constructs studied here are discussed.

A consistent finding is that both complementary and supplementary fit are related to individual behavior in organizations. With regard to job search, for instance, attraction to organizations has been correlated with PO fit on preferences for reward structure, individual personality to organizational climate, and value congruence between job seekers and potential organizations. From the organization's perspective, perceived (rather than actual) PO fit between applicants and interviewers has also been shown to be a predictor of interview outcomes; applicants who were perceived as having better PO fit with the organization were more likely to receive higher interview ratings. PO fit has also shown positive individuals' attitudes at work. This has been found in a number of studies examining effects of value congruence on attitudes such as motivation, job satisfaction, work group cohesion, and organizational commitment. Similar findings have been reported in several studies examining the congruence between employees and their supervisors. Dipboye (1994) suggests that person-job fit is so important that it is at least in part responsible for the entrenchment of unstructured interviews in organizational settings despite a large body of literature indicating that structured interviews are far superior (e.g., McDaniel, Whetzel, Schmidt, & Mauer, 1994).

As one of the primary outcomes of interest in this study is turnover, it is helpful to review the effects of fit on turnover in more detail. The reasoning underlying this research is that positive outcomes associated with PO fit such as those described above lead to an increased propensity to remain on the job. Similarly, a lack of PO fit is thought to produce dissatisfaction, strain and other negative outcomes. Hence, turnover should result from a lack of PO fit. Indeed, Chatman (1991) reported that PO fit based on values measured by the Organizational Culture Profile (O'Reilly, Chatman, & Caldwell, 1991) predicted auditors' satisfaction and intentions to remain with their firms. O'Reilly et al. (1991) reported similar findings with actual turnover in the same industry, as did Vendenberghe (1999) in the health care industry. Bretz and Judge (1994) reported that PO fit with regard to several important aspects of fit predicted both satisfaction and job tenure. Kristof-Brown and Kay Stevens (2001) examined the role of goal congruence, a study that was conducted within the context of students completing a group assignment for a class requirement. Their results indicated that the perceived congruence between each student and the team predicted student satisfaction with the team Similarly, Vancouver and Schmitt (1991) found that the fit between teachers' and principals' goal congruence predicted satisfaction, organizational commitment and intention to remain on the job. Chan (1997) examined something he termed cognitive misfit, the lack of fit between the cognitive demands of a job and the preferred cognitive style of employees in that particular job. His results indicated that cognitive misfit between the employee and the job was related to turnover after controlling for the effects of job performance. Other research examined the fit between job creativity demands and the environmental conditions that facilitate creativity (Shalley, Gilson, & Blum, 2000). Job satisfaction and

intent to remain on the job was higher when there was a match between the job demands and the supporting environment. Again, the underlying theme of these studies is that misalignment between the individual and environment (e.g., organizational values or goals, or job demands) results in a variety of negative outcomes such as stress, low levels of job commitment, and satisfaction, and that these outcomes, in turn, produce turnover. Indeed, Intent to remain on the job is an important cognition, in part because it is moderately related to actual turnover (ρ=.38; Griffeth et al. 2000). These results, taken together, suggest that the fit between an employee and the job or the organization exerts consistent effects on turnover and related outcomes.

The gravitational hypothesis (McCormick, DeNisi, & Staw, 1979; McCormick, Jeanneret, & Mecham, 1972) addresses fit more specifically by suggesting that people gravitate towards jobs throughout their life based on a fit between their own abilities and job complexity. This is thought to occur both because people should choose jobs that are commensurate with their abilities, and because employers select people whose abilities can meet the demands of a given job. Given that there are often educational requirements for entry into jobs, and that cognitive ability is frequently used to select employees, it is logical to think that a person's ability would thus be central in determining how people move from job to job over time. There is some research supporting the notion that this indeed occurs.

Wilk, Desmaris, & Sackett (1995) examined job mobility over a five-year period for people who had finished high school. These researchers found that cognitive ability predicted job change at the end of the study after controlling for initial job complexity. Wilk and Sackett (1996) extended this study by examining ability-job complexity

changes in two large national databases spanning ten and twelve years. These researchers found significant positive relationships between ability and change in objective and subjective job complexity over time. These results are consistent with studies reporting that job performance-turnover relationships might be non-linear. For instance, Trevor et al. (1997) reported an inverted parabolic relationship between job performance and turnover, and that salary growth and promotions impacted the strength of this relationship. Because ability is a consistent and strong predictor of job performance, ability should have a similar relationship to turnover; however, the form of this relationship is likely to be different depending on whether one examines VT or IVT. In particular, based on the gravitational hypothesis and the results described above, high ability employees in low complexity jobs should have higher rates of VT but not IVT as compared to low-ability employees. The jobs examined here are very low on job complexity. Hence, the following hypothesis is proposed:

> H10. High ability employees will be more likely to exhibit VT as opposed to IVT after controlling for Conscientiousness.

The research described above indicates it is possible to examine many types of fit rather than fit within an organization per se. Indeed, fit has been studied in a wide variety of contexts including organizations, jobs, teams, groups, leaders, dyads and supervisors. A number of authors have indicated that the boundaries between these different types of fit are often blurred, though they all fall under the more general rubric of person-environment fit. What makes these results especially intriguing for teams is that they study the fit between the person and relatively distal contexts as compared to the team context. That is, the context provided by a team should be much more immediate to the

everyday experiences of an employee than the overall organizational context. This is based on the notion of "bond strength", which states that constructs at the same or adjacent levels of analysis (e.g., a person and team) should be more strongly related than those that are more distal (person and organization; Simon, 1973). There may be several levels of hierarchy separating an individual and an organization (e.g., team, sub-unit, and department) whereas persons and teams are adjacent levels in an organizational hierarchy (dyads are not normally considered to be an organizational level). This is similar to House et al.'s (1995) and Tosi's (1992) notion of 'tight coupling', which states that units across different levels of analysis that are more interdependent will influence each other more strongly. Individuals actually make up a team, suggesting the two are tightly coupled; hence, one would expect to find strong relationships between teams and individual behavior. In fact, where team based work systems predominate, the team context should be the main context experienced by an employee. For instance, Hackman (1992) describes this context as the "social world" of an employee, impacting the ambient stimuli to which a given employee is exposed. He argues that these stimuli, in turn, affect employees' cognitions, affect, and behavior. Bedeian, Kemery and Mossholder (1989) and McGrath (1998) also argue that team composition should be viewed as a context that can shape individual behavior.

This suggests that person-team fit should be strongly related to individual and team outcomes (the 'Fit' and k link in Figure 3), especially given the consistent relationships between a variety of variables and PO fit. Similarly, it is also important to examine how team characteristics may serve as a context that influences not only team-relevant outcomes, but also relationships between individual level construct relationships

(Bedeian et al.1989; Moreland & Levine, 1992; Moreland, Levine, & Wingert, 1992). For instance, Van Der Vegt, Emans, and Van De Vliert (2001) examined the effects of a match between team-level outcome interdependence and the extent to which tasks were interdependent at the individual level. These researchers hypothesized that individual task interdependence would more strongly relate to job and team satisfaction and job and team commitment when there was outcome interdependence at the team level (i.e., that the team level variable moderates the individual level relationships). The results generally supported these hypotheses: A match between the rewards in a team and the extent to which those rewards are obtainable via task interdependence yielded positive affective outcomes. Again, while not examining the composition-based fit, this research nonetheless broadly suggests that fit between individual level and team level constructs leads to positive outcomes and that this can be examined as an effect that moderates relationships at the individual level of analysis.

Teams as a Context

The foregoing discussion shows why the fit of an individual within a team should be related to important outcomes. Indeed, several authors have recently focused on the idea of person-team fit in personnel selection (Werbel & Gilliland, 1999; Werbel & Johnson, 2001) and person-group fit more generally (House et al., 1995; Morgan & Lassiter, 1992). Werbel and Johnson suggest that positive interpersonal interactions are very important in team based work environments because they should positively impact both team process and performance. This suggests one should emphasize supplemental fit as a predictor of effective team process and outcomes. In terms of process, these authors, as well as others (e.g., George, 1990) posit that increased supplementary fit will positively

impact team cohesiveness, or the bonds between team members. Cohesiveness, in turn, has been related to team performance (Gully, Devine, & Whitney, 1995), and is also expected to relate to turnover within teams (Griffeth & Hom, 1994). Supplemental fit is also expected to facilitate cooperation as well, another critical aspect of team process (Werbel & Gilliland, 1999) that has been linked to team performance. When is complementary fit important? Theory suggests that complementary fit within a team is important when the task requires specialized knowledge and skills (i.e., if there are a number of specialized roles within the team). This is likely to be the case when a task is complex (Morgan & Lassiter, 1992). For instance, in cross-functional teams one might seek complementary fit on functional backgrounds to increase the team's breadth of expertise, and thus supposedly the effectiveness of the team as a whole. Indeed, a number of studies examining cross-functional teams indicate that heterogeneity with regard to skills or functional background is beneficial in this regard (e.g., Ancona & Caldwell, 1992; Voiers, 1956, cited in Jackson, 1992). Complementary fit also tends to be emphasized when examining the compatibility between job demands and employees' abilities or environmental facilitators of those demands (e.g., Chan, 1997, Shalley et al., 2000). In contrast, I have located no research suggesting that the personality-based complementary fit within large, undifferentiated teams is critical. Accordingly, the present research largely emphasizes supplemental fit.

Several authors argue that fit is inherently multilevel (Schmitt & Chan, 1998; Werbel & Gilliland, 2000) and that selection should be evaluated from a multi-level perspective (Ployhart, 1998; Schneider, 2000), an idea that is central to the research reported here. For instance, House et al. (1995) suggest that a lack of fit will result in decreased

influence across levels of analysis (e.g., from the team to the individual). This supports the notion that person-team fit might be related to turnover because there will be weaker bonds holding the individual to the team. At the team level, one might think about misfit as team-level diversity on important individual attributes of interest. This is appropriate if there is good reason to believe that the variability within a team on a given attribute tells us something meaningful about the team itself that can be related to other key team level constructs. This is likely to be true when the focus is on supplemental fit because variability within a team implies supplemental misfit. On the other hand, one can construe person-team fit as a cross-level effect where the team level mean (or other appropriately conceptualized operationalization of team standing) moderates relationships between individual level variables of interest (Rousseau, 1985). This approach would be helpful if one were interested in treating teams as a context shaping relationships among individual level variables (Bedeian et al., 1989; Markham, 1988). Very few studies have empirically studied teams in such a way. One exception is research reported by Kidwell, Mossholder, and Bennett (1997) who found that increased work group cohesiveness strengthened the relationship between job satisfaction and courtesy. Underscoring how little work has been done in this area is that Kidwell et al.'s study was more a demonstration of statistical techniques rather than a substantive focus on teams as a context. Treating person-team supplemental fit as a cross-level effect is conceptually closely tied to examining the effects of team level diversity; in fact, the two can be conceptually equivalent given correctly specified composition models. Accordingly, the following sections discuss person-team fit as both team-level diversity and a cross-level effect.

Personality-Based Team Context

There are several reasons to think that the fit of personalities within a team should impact team process and team effectiveness. First, a particular mixture of personalities may lead to competition or conflict in a team. For instance, it is easy to imagine how a team with little or no variability on the personality trait of dominance might have difficulties. If all the team members are very similar on dominance, leaders might not emerge in the team, or everyone might be competing for a leadership position. In either case, one would expect there to be negative effects on team process and performance; however, more commonly it is thought that similarity in personalities should lead to better interpersonal relationships. For instance, there is a substantial amount of research supporting the notion that personality similarity is related to more positive exchanges between leaders and team members (e.g., Askanasay & O'Connor, 1997; Engle & Lord, 1997; Phillips & Bedeian, 1994), job satisfaction (Meglino, Ravlin, & Adkins, 1989; 1991) and performance ratings (Antonioni & Park, 2001). Examining job search preferences, Judge and Cable (1997) linked the congruence between job seeking students' personalities to attraction to organizations. The theory most often cited to support this notion is the similarity-attraction paradigm (Berscheid & Walster, 1978; Bryne, 1971) which states that similarity (e.g., racial, attitudinal, etc.) leads to liking and attraction. This is because similarity is thought to lead to positive reinforcement of one's attitudes and beliefs. This reinforcement, in turn, leads to smoother and desirable social interactions. There is an overwhelming amount of research conducted over several decades in a variety of contexts supporting this perspective. This suggests that people should find it less rewarding and more dissatisfying to work with people that are

dissimilar to them. This is also consistent with Schneider's (1987) ASA model discussed earlier.

Research conducted by Ferris, Youngblood, and Yates (1985) also provides some broad support for the notion that the similarities within a team (i.e., supplementary fit) predict organizationally relevant outcomes. Like Kidwell et al. (1997), these researchers examined the group as a context that shapes individual level outcomes. Their results indicated that the overall FFM-based personality fit between a sample of newcomer flight attendants and a sample of incumbents moderated the relationships between performance and attendance and turnover. In particular, performance was unrelated or negatively related to these outcomes where person-group fit was low, whereas the opposite was found where fit was high. Several caveats are worth mentioning, however; first, person-group fit did not predict turnover or absenteeism. This is inconsistent with the broader notion that person-group fit is associated with positive outcomes (i.e., attendance and remaining on the job). Second, the referent group was a small sample of flight attendants taken from the larger pool of flight attendants as a whole. Since there was little data or discussion of the representativeness of this sample of flight attendants, it is unclear exactly what the referent group represented in this research. Third, these researchers calculated a distance score for each newcomer flight attendant across all the FFM personality constructs. Thus, the nature of this measure was unclear because it consisted of five different distance scores. Nonetheless, this study is still somewhat informative because it broadly suggests that groups' personalities might provide a context that shapes relationships among variables or constructs at the individual level.

There is theoretical support for the notion that team process and thus performance might decline as a function of personality misfit. Shaw and Barret-Power (1998) argue that 'cognitive diversity' impedes a group's ability to integrate their behaviors, and that this is increasingly likely to occur on tasks where multiple perspectives are not critical (like the task of the teams studied here). This should also be true when there are multiple preferences for how work is conducted (e.g., levels of conscientiousness). Overcoming the effects of this diversity is said to have a "cognitive cost" which makes this difficult (p. 1315). The literature on shared mental models is also consistent with this notion. Shared mental models are schemas that are shared within the team that are in some way relevant to the team task at hand. These schemas can relate to either how the task should be performed (i.e., taskwork) or how team members should interact to achieve the task (i.e., teamwork). Accordingly, Rentsch and Hall (1994) state that shared mental models can manifest themselves as expectations and guidelines that structure team interaction and the compilation of individual inputs into outputs.

These authors present a theoretical model that suggests several different ways in which schema dissimilarity can disrupt teamwork and taskwork and thus impact team performance and the desire of the team to work together in the future. More concretely, however, it is easy to give examples of how schema dissimilarity might lead to less efficient team processes. For instance, if two individuals have different schemas regarding the notion of cooperation they might encounter serious communication and coordination problems if they are asked to cooperate. The effects of this similarity might be expected to cascade if it happens at the team level. These authors also assert that schema similarity is more likely to occur when the team members show supplemental fit

in terms of their values, personalities, and preferences. This is important because the extent to which effective mental models are shared have been theoretically (Salas et al., 1992) and empirically (e.g., Mathieu et al., 2000) linked to impact performance and are considered by some to be critical determinants of coordinated team performance (e.g., Rouse, Canon-Bowers & Salas, 1992). To the extent that Conscientiousness, Agreeableness, or Extraversion represents a preferred way of behaving at work in team settings, one might expect that the sharedness of these preferences should relate to important outcomes. Jehn, Northcraft and Neale (1999) reported some results consistent with this notion in that values for the way in which work was conducted in a sample of moving employees predicted intent to remain on the job and actual group performance; however, these researchers used D-scores and analyzed their data at the individual level, so these results must be interpreted cautiously.

Conscientiousness. Of the three personality constructs examined here, conscientiousness is clearly the most directly related to tasks at work; it is easy to translate the traits associated with conscientiousness to specific behaviors one might exhibit at work. For example, someone who is detail oriented would tend to complete tasks at work to higher standards than someone who is low on conscientiousness (i.e., is not detail oriented). These behaviors, in turn, are likely to be at least in part due to an underlying cognitive preference. For example, one might reasonably infer that someone who generally describes herself as detail oriented prefers to ensure that the details of tasks are completed. To the extent that this is true, then, one might expect that the fit between an individual's standing on conscientiousness and that of her team might lead to important outcomes. This is because a lack of fit would imply that work tasks are being done in a

manner that is discordant with the preferences of the individual: In our example of detail orientation either too much or too little attention is paid to the details of work tasks. Either situation can cause dissatisfaction in a worker that does not fit in this way. For instance, an employee that is substantially lower than the team-level mean on conscientiousness might be pressured by coworkers to put more effort into the details or work tasks. On the other hand, the work products of coworkers he or she considers sloppy might frustrate someone markedly higher on conscientiousness. At an individual level, dissatisfaction may result. If there is large variance in conscientiousness at the team level (i.e., an overall lack of supplemental fit within the team), overall process might suffer as a result because high conscientious individuals might have to redo the work of low conscientiousness individuals (Barrick et al., 1998a).

Several studies have examined or reported the impact on conscientiousness similarity in team and work units. Barrick et al. (1998), in the study of manufacturing and maintenance teams described earlier, reported that the variance of conscientiousness scores within a team was negatively related to team performance (r=-.33); however, this effect was not hypothesized. Neuman et al. (1999) also examined the same index in an exploratory manner and found no relationship between conscientiousness diversity and team performance. These researchers did not provide a detailed description of the extent to which the team members were interdependent; hence, it is difficult to develop possible explanations for these null results.

Day and Bedeian (1995) examined the effects of conscientiousness similarity on the job satisfaction, job performance and tenure in a sample of Black nurses. Their sample represented approximately one-third the entire sample of nurses in a medical center and

similarity scores were calculated for each nurse. These similarity scores were designed to index the extent to which each nurse was similar to the larger sample of nurses within the entire medical center. These researchers found that higher similarity was associated with lower organizational tenure and higher job satisfaction. The former finding was contrary to what one would expect based on the research and theory described above. This null result may have occurred because in their particular context nurses may not have necessarily worked together. Perhaps conscientiousness similarity is not consistently related to outcomes in predictable ways when employees do not work together on team tasks.

Thus, to summarize, the existing literature paints an unclear picture of the relationship between person-team conscientiousness fit on individual and group outcomes; however, based on the reasoning put forth above, it is expected that conscientiousness fit at the team level would indeed predict team performance. In particular:

> H11. Increased team level diversity in conscientiousness should be associated with lower team performance.

Recall the earlier discussion describing how person-team fit can also be examined as a cross-level effect as well as a team-level diversity effect. The individual level outcome examined in the present study is turnover. Based on the general findings in the PO fit literature, as well as the reasoning described above, it was expected that conscientiousness fit would relate to increased VT but not IVT. Hence, it is hypothesized that:

> H12. Increased diversity in conscientiousness should be associated with higher team voluntary turnover (but not involuntary turnover).

H13. Increased conscientiousness-based person-team fit should be associated with a lower likelihood of individual level voluntary turnover (but not involuntary turnover).

Agreeableness. Two of the three studies described above (Barrick et al., 1998a; Neuman et al., 1999) found no relationship between Agreeableness diversity and team performance. Barrick et al. did report, however, that increased Agreeableness diversity was associated with lower levels of cohesion and communication, and higher levels of conflict. Day and Bedeian (1995) found that increased Agreeableness similarity was associated with higher job performance (r=.17). Like Conscientiousness, these results are mixed; however, it is less clear why one might expect the diversity or fit with regard to agreeableness to impact team process or performance. This is because people that are agreeable by their nature are supposed to be understanding and tolerant. One would thus expect people that are high in agreeableness not to react negatively to having to work with people that are low in agreeableness. Hence, given the lack of theoretical basis and conflicting empirical results, no hypotheses regarding agreeableness fit are proposed and exploratory analyses will be conducted.

Extraversion. The existing research on extraversion diversity in teams is somewhat clearer than that on Agreeableness. Barrick et al. (1998) hypothesized and found that it would positively relate to team social cohesion; however, there was no explicit discussion of why they expected this to occur. Neuman et al. (1999) also reported that the variance in team extraversion was positively related to team performance ratings (r=.26), though again, their research was exploratory. In contrast, Day and Bedeian (1995) found that extraversion similarity was related to job satisfaction, though not to performance. However, these researchers used difference scores to index similarity, an approach that

has been widely criticized (see below). Perhaps a mix of extraversion is necessary within a team because otherwise interpersonal conflict will arise as the extraverts 'jockey' for social interactions. In other words, at the team level, supplemental fit might actually lead to negative effects on team outcomes because of the nature of Extraversion as a construct. On the other hand, it is somewhat difficult to see how an extravert might prefer to work with introverts. Hence, the relationship between extraversion fit will be examined in an exploratory fashion.

Demographics. Jackson et al. (1995) differentiate between readily detectable ('surface level') and underlying ('deep level') qualities that can be used to examine the effects of diversity. The discussion thus far has focused on deep level diversity and composition effects. The following sections discuss the literature surrounding the effects of surface level diversity, laying the groundwork for hypotheses that focus on this aspect of person-team fit.

It is widely acknowledged that the demographics of the American workforce are rapidly changing. Over the last few decades there have been marked changes with regard to the racial, gender, and age composition of the workforce. In particular, there is an increasing diversity in these three demographic variables in the workforce (National Research Council, 1999). Not surprisingly, there is a substantial amount of psychological literature devoted to the effects of this increased diversity (i.e., the demographic mix) within teams and other organizational units (Williams & O'Reilly, 1998).

This literature was sparked by Pfeffer's (1983) seminal article that coined the term organizational demography. Pfeffer argued that the distribution of demographics within a given organizational unit (e.g., organization, department, workgroup, etc.) impacts the

quality of social integration within that unit. Focusing on tenure, Pfeffer argued that the distribution of demographics within an organizational unit (e.g., organization, department, workgroup, etc.) impacts the amount of conflict within that unit. This disruption, in turn, was theorized to impact performance, innovation, turnover, and power distributions. In the years following Pfeffer's article researchers have examined the effects of a wide array of demographic variables including race, age, sex, job tenure and functional background on outcome variables such as unit performance, turnover, and unit-processes such as social integration, social network strength, and conflict (e.g., O'Reilly, Williams, & Barsade, 1998). An offshoot of this thinking was research on the effects of demographic similarity at the individual level, referred to as relational demography (Tsui, Egan, & O'Reilly, 1992; Tsui & O'Reilly, 1989). The idea underlying this research is that demographic similarity positively impacts the social relationships between individuals. Conversely, demographic dissimilarity is said to disrupt social relationships between individuals. These putative effects of dissimilarity are thought to negatively impact affect and attitudes, which in turn are hypothesized to impact performance and turnover. Thus, despite the focus on work units and individuals, respectively, the underlying processes and theoretical rationales for both relational and organizational demography are very similar.

There are several major theories in social psychology that support the notion that the mix of demographic characteristics within a group or team impact the outcomes described above. First, there is the similarity-attraction paradigm, described earlier. One would expect this theory to apply to demographics to the extent that demographics are salient indicators of similarity in a particular environment. Two theories that speak to

this point are categorization (Tajfel, 1981; Turner, 1987) and social identity theory (Turner, 1982). Social categorization theory states that individuals have a strong need to maintain high self-esteem. In doing so, they categorize others according to salient characteristics including race, sex, and membership in groups. The need to maintain positive self regard leads people to strongly identify with their own group and to maximize perceived differences between groups. The end result, which has received an overwhelming amount of empirical support, is that individuals tend to react more positively in interactions with people in the same group, even when group distinctions are arbitrary (Sherif, 1961). Thus, even if demographics are not related to underlying attributes on which people might differ, it is still possible that people assume that they are (e.g., McGrath, 1998; Milliken & Martins, 1996). Before this large body of research is reviewed, however, it is important to first discuss several important theoretical and methodological limitations that pervade this research. This is because these limitations make it difficult to draw unambiguous conclusions from these studies, and so they should be considered in any review of this research.

Methodological and Conceptual Issues in Relational Demography Research. Failure to pay careful attention to issues relating to levels of analysis can lead to a host of conceptual and data analytic problems (e.g., Klein & Kozlowski, 2000; Ostroff, 1993; Roberts & Burstein, 1980; Roberts, Hulin, & Rousseau, 1978; Rousseau, 1985). Relational demography research largely suffers from the several problematic ways in which researchers conceptualize and thus operationalize demographic diversity. Like the other composition variables discussed earlier, demographics are measured at the individual level yet the key predictor of interest is demographic dissimilarity between an

individual and a referent organizational unit. In reviewing the relational demography literature, Riordan (2000) outlines the three major approaches to operationalizing this dissimilarity.

The most common approach calculates a difference (D) score for each individual that is said to represent the average demographic distance between an individual and the members of the focal organizational unit of interest. D-scores have been widely criticized on a host of conceptual and methodological grounds (e.g., Edwards, 1994; 2002; Johns, 1981; Riordan & Shore, 1997; Riordan, 2000) yet they are still commonly used. The second approach which Riordan termed the 'interaction approach', involves disaggregating an index of group composition to the individual level and performing moderated regression analyses to test the significance of the interaction between group composition and individual level demographic characteristics (e.g., Chattophaday, 1999; Ferris, Judge, Chachere & Liden, 1991; Flynn & Shore, 1994, cited in Riordan, 2000; Mellor, 1995; Riordan & Holliday-Wayne, 1998, cited in Riordan, 2000; Riordan & Shore, 1997; Riordan & Weatherly, 1999, cited in Riordan, 2000). Although Riordan (2000) criticizes this approach as lacking in power, in fact, tests of these interaction terms are biased upwards because the individual level sample size severely overestimates the number of groups that exist (i.e., by several fold; Bryk & Raudenbush, 1992). The interaction approach, as well as the D-score approach, introduces correlated errors of prediction that are not accounted for in ordinary least squares regression which can yield unpredictable results (Bryk & Raudenbush, 1992; Hannan, 1990; Hofmann, 1997; Kennedy, 1998; Kenny & Judd, 1986). For instance, when nonindependence due to groups is ignored decisions about higher-level variables (e.g., group racial composition)

will be more prone to Type I errors, whereas those concerning lower-level variables (e.g., individual race) will exhibit more Type II errors (Bliese, 2000; 2002; Heck & Thomas, 2000). Bryk and Raudenbush (1992) also note that ignoring independence likely violates the constant variance assumption. Clearly, when forming an interaction term the true results are muddled to say the least. Indeed, it has been well documented that disaggregation is associated with a host of problems, both conceptual and analytical (e.g., Glick & Roberts, 1984; Hannan, 1971; Langbein & Lichtman, 1978; Nezleck & Zyzniewski, 1998; Rousseau, 1985; cf. Roberts et al., 1978). Thus, in research using the first two approaches (e.g., Barsade et al., 2000; Chatman, et al., 1998; Chatman & Flynn, 2001; Chattopadhyay, 1999; Day & Bedeian, 1995; Ferris et al., 1991; Flynn & Shore, 1994; Green, Anderson, & Shivers, 1996; Iverson & Buttigieg, 1997; Jackson et al., 1991; Mellor, 1995; Mueller, Finley, Iverson, & Price, 1999; O'Reilly et al., 1984; O'Reilly et al., 1989; Pelled, 1996; Riordan & Holliday-Wayne, 1998; Riordan & Shore, 1997; Riordan & Weatherly, 1999; Tsui, Egan, & O'Reilly, 1992; Tsui & O'Reilly, 1989; Turban & Jones, 1988; Wagner et al., 1984; Westphal & Zajac, 1995; Wiersema & Bird, 1993; Zenger & Lawrence, 1989) it is difficult to determine the extent to which the reported results accurately describe the true nature of the relationships of interest. Yet, these approaches continue to be used despite the acknowledgment by influential authors in the area that demographic effects can occur at both the individual and group level (Jackson et al., 1995) and that attribute dissimilarity is a multi-level construct (Jackson, May, & Whitney, 1995; Jackson, Stone, & Alvarez, 1992, Figure 1). Theoretically, the first two approaches misspecify the level at which the hypothesized processes occur (Rousseau, 1985). This is because these approaches examine effects at the individual

level despite theory focusing on the dissimilarity between an individual and a group. Hence, applying the results to any level of analysis is difficult because model misspecification effects are pervasive (e.g., Klein, Dansereau, & Hall, 1994; Roberts et al., 1985; Rousseau, 1985).

A third flawed approach is the use of perceived demographic similarity which asks individuals to indicate how demographically similar they perceive themselves in relation to a work unit (e.g., Kirchmeyer, 1995; Ragins & Cornwell, 2001; Riordan & Weatherly, 1999, cited in Riordan, 2000). While perceived demographic similarity may have some limited interest in its own right, it should not be treated as the same as actual demographic similarity for several reasons. First, demographic similarity is a broad term and it is likely that individuals have very different internal conceptualizations of what that might mean (Riordan, 2000). Second, because it is so broad it cannot easily be meaningfully compared to other indices of demographic similarity. Third, because demographic information is often readily available and clearly observable (McGrath, Berdahl, & Arrow, 1996) it seems that this information should be gathered instead or that actual and perceived similarity should be compared, at the very least. Perceptual measures such as this are generally only used when the characteristics of interest are not outwardly observable or when there is clearly articulated reasoning explaining the interest in the perceptual approach to similarity. Not only are all three approaches described above seriously flawed, they also yield results that do not exhibit convergent validity (Riordan, 1997) which makes drawing conclusions across studies difficult.

These limitations reflect a lack of conceptual clarity and methodological sophistication. This is because dissimilarity between an individual and a group is not

solely an individual or group-level phenomenon; it is a cross-level effect, referring to the demographic fit between an individual and a group. That is, the composition of the group is thought to moderate the relationship between the individual level demographic variable and the individual-level outcome of interest. Given a significant interaction effect, its form, in turn, will reveal whether the results are consistent with theory. This approach is consistent with the fundamental assumptions of relational demography stating that demographic dissimilarity is the key predictor of interest. Data analytic tools that allow for modeling both individual and group level variables do exist (Bryk & Raudenbush, 1992), though they are not as widely understood and used as ordinary least squares regression, the analytic technique of choice in much industrial/organizational psychology and management research. Alternatively, if the outcome of interest is at the group level it is reasonable to examine the relationships between demographic diversity and the outcome variable of interest as long as the outcome variable is either (1) meaningfully measured at the group level, or (2) if aggregation from the individual level is theoretically justified given the data at hand. Modeling relational demography as both a cross-level and a group level diversity effect is consistent with the notion that different processes, outcomes, and relationships among variables can occur at different levels of analysis (Ostroff, 1993; Rousseau, 1985). Note, however, that it is also possible to operationalize team level diversity in different ways, and that the theory underlying a given study should guide the approach.

The literature generally reports that increased diversity or dissimilarity leads to a host of negative outcomes in terms of communication, satisfaction, cohesion, and social integration in the workplace, and that these in turn impact turnover and the performance

of organizational units. Despite the criticism leveled above, it is still instructive to review this research to understand the overall tone of the findings and the types of questions that are asked. Further, some (albeit very few) studies do not use the flawed data analysis methods described above. Thus, the following sections discuss this literature as it relates to team and work-unit composition.

Race. Race plays an especially important role in personnel selection because there is often pressure not only to remain within legal guidelines but to also racially diversify workforces beyond legal requirements (Arvey & Faley, 1988; Hartigan & Wigdor, 1989; Schmidt, 1993; Schmitt, 1989). This issue has generated an extensive body of literature (e.g., de Corte, 1999; Hough, Oswald, & Ployhart, 2001; Roth, Bobko, Switzer & Dean, 2001; Sackett & Ellingson, 1997; Schmitt et al., 1997). Thus, from a practical perspective it is especially important to examine the effects of team racial composition.

Several studies have examined the effects of unit racial composition on aspects of unit effectiveness. Two of the very few studies in this area that have appropriately modeled relational demographic variables examined the effects of unit racial diversity and supervisor-subordinate racial dissimilarity on job satisfaction. Wharton, Rotolo, and Bird (2000) found that race heterogeneity affected mean levels of job satisfaction within university departments even though race itself was not significantly related to job satisfaction at the individual level. Similarly, Wesolowski and Mossholder (1997) found that the interaction between supervisor and subordinate race predicted subordinates' job satisfaction such that subordinates with supervisors of different races exhibited less job satisfaction. Mueller et al. (1999) reported that increased teachers' racial diversity in schools was negatively associated with job satisfaction; however, in school settings it

seems likely that there are so many other possible influences such as resource availability, which is known to be strongly correlated with race in educational settings, that it may be difficult to draw firm conclusions from this study. Pelled, Eisenhardt, and Xin (1999) reported that racial diversity in process improvement teams was positively and significantly related to mean group-level perceptions of conflict. In contrast, Pelled (1996) conducted another study in which racial dissimilarity was unrelated to several different types of perceived conflict. Pelled speculated that she might not have obtained the expected results because her small teams were so highly integrated (she gives a hypothetical example of a team consisting of one Black, one Asian, one Hispanic, and one White) that employees of different races were less likely to feel like outsiders. Tsui et al. (1992) reported that higher racial distance was related to lower individual psychological commitment and intent to remain with the organization and absence. Chatman and Flynn (2001) also reported that demographic similarity (calculated as an aggregate of race, sex, and citizenship) was negatively associated with group process. In particular, these researchers found that demographic diversity was negatively associated with individuals' perceptions of cooperative group norms and satisfaction; however, the results of these last three studies are ambiguous because all used the D-score approach.

Riordan and Shore (1997) studied the attitudinal and perceptual outcomes associated with racial diversity. These researchers reported that racial diversity of work units was related to lower perceived work group productivity and advancement opportunities, commitment and cohesiveness; however, racial diversity was assessed for three different races (Whites, Blacks, and Hispanics) coded as a single variable and group racial composition was split into 3 categories, disaggregated, and used in calculating an

interaction term for moderated regression. Hence, although this is one of the most widely cited studies in this area it is unclear what conclusions can be properly drawn.

Several authors have suggested that demographic diversity might relate to team performance (e.g., Shaw & Bennet-Power, 1998) though only a few studies have specifically examined racial diversity. Timmerman (2000) examined the effects of racial diversity in two different types of teams that exhibit varying levels of task interdependence: Professional baseball and basketball teams. The theory underlying this approach is that more coordination is required for a sport like basketball, which is more interdependent than baseball. Hence, diversity effects were expected for basketball as opposed to baseball because any negative effects of diversity should be more disruptive (i.e., should lead to more process loss) in a sport that requires more coordination. Similarly, the higher level of coordination should be more difficult to achieve in the face of any negative effects that might result from racial diversity. Data from 38 years of baseball and basketball teams were used, and racial diversity predicted team basketball performance (i.e., winning percentage) after controlling for aggregated individual level indices of performance (e.g., field goal percentage). Consistent with expectations, the same effect was not found in baseball. This suggests that racial diversity disrupts team cooperation or coordination, which in turn impacts performance. This finding is consistent with those of Hoffman (1985) who reported that the proportion of Black supervisors was positively associated with communication via formal meetings. As formal communication may be a substitute for more informal communication, these results suggest that increased racial diversity leads to weaker social network ties related to social communication (Shaw, 1981). A number of studies suggest that weaker social

network ties lead to turnover in organizations and groups (e.g., Feeley & Barnett, 1997; McPherson, Popielarz, & Drobnic, 1992) implying that racial diversity might lead to turnover.

Several authors have compared racially homogeneous and non-homogeneous teams. Watson, Kumar, and Michaelson (1993) compared such teams of college students who had group assignments for a management course. These researchers found that homogenous groups had higher ratings of group performance and several process variables, though these effects faded at the end of the 17-week assignment. Baugh and Graen (1997) also compared the same types of teams in a state government agency. Racially non-homogenous teams perceived themselves as less effective than racially homogenous teams, though this result did not hold for external evaluations of team performance. These researchers noted their small sample size (N=31 teams) might have accounted for the lack of significance. At the least, this research suggests that racial diversity impacts team members' attitudes.

Another aspect of effectiveness, turnover, has also been examined as a potential outcome of racial dissimilarity. Chattopadhyay (1999) used D-scores to examine the effects of racial dissimilarity on organizational citizenship behaviors (OCBs). Racial dissimilarity was negatively related to three self-report facets of OCBs. Further, race interacted with racial dissimilarity in this regard; that is, this effect occurred for Whites working in minority-dominated groups but not when the opposite was true, a result also reported by Tsui et al. (1992). This effect was found to be mediated by organization-based self-esteem rather than peer relationships, suggesting that the Whites who worked in these settings felt undervalued. This is also consistent with research reported by

Hinds, Carley, Crackhart, and Wholey (2000) who reported that students choosing work group members on long-term projects preferred members of the same race. Turnover from a group should result if dissatisfaction with the group members surpasses a certain threshold.

Although I have found no studies of the effects of racial diversity on team performance in organization settings, Richard (2000) studied this issue for the entire set of employees working in banks. Using mail surveys, this author obtained racial composition information from the banks' human resource managers and compared it to revenue data and two other indicators of bank performance. The zero-order correlation between racial diversity and market performance was significant and negative (r=-.32), though after entering relevant control variables into the regression equation this effect was largely eliminated. It is important to note that this study had a fairly small sample size (N=63 banks), 75% of which was drawn from California. Not only is California widely considered to be among the most liberal states in the country ("State's Diversity Becoming Its Strength", 2001), it is also the most racially diverse. Hence, these null results are not necessarily very surprising. One might expect to find more positive results in a more representative sample. It is also possible that the results might have been different had the unit of analysis been work teams rather than the entire bank as a whole.

In sum, the results reported above suggest that racial person-team fit has several important behavioral and attitudinal outcomes; however, due to serious methodological and conceptual problems that underlie almost all of the studies conducted to date further research is needed to more accurately assess these effects. Nonetheless, these results taken together, and the substantial theory underlying this research, suggest that racial

misfit within organizational units leads to a number of important outcomes. Similarly, because a number of researchers suggest that demographics might be proxies for other deeper-level variables, this research will control for ability or any of the personality constructs that are related to demographics and turnover. Hence, the following hypotheses are proposed:

> H14. The racial fit of the individual within the team will predict individual turnover such that misfit is associated with a higher likelihood of turnover. Based on work by Tsui et al (1992), this relationship will be more evident in Whites working in teams with higher minority representations rather than vice-versa. This effect will be evident after controlling for the relevant personality constructs and ability.
>
> H15. Increased team racial diversity will be related to increased team turnover rates. This effect will be evident after controlling for the relevant personality constructs and ability.
>
> H16. Increased racial diversity will be negatively related to team performance. This effect will be evident after controlling for the relevant personality constructs and ability.

Age. A number of studies have examined the effects of the mix of ages or age dissimilarity in work units on a variety of psychological and group process variables, and turnover. Most of these studies used the D-score approach. Tsui et al. (1992) studied entire organizational units (e.g., an entire hospital) and reported that age dissimilarity predicted intent to remain with the organization but not psychological commitment. This odd result suggests that perhaps one or both measures lacked construct validity or that the entire organizational unit represented a focus that was too broad. In contrast, Zenger & Lawrence (1989) found that age dissimilarity was negatively related to communication

outside of the project team, and that dissimilarity was similarly related to intra-team communication. Two other studies in top-management (Wagner et al., 1984) and work teams (O'Reilly et al., 1989) found that age similarity was negatively related to turnover; however, the latter study reported that age similarity was not related to social integration, a finding that is conceptually at odds with the reported turnover relationship.

Other researchers have used different approaches to study age dissimilarity. Kirchmeyer (1995) used the perceptual approach to examine what he termed 'generational similarity', which was a combination of perceived age, education and lifestyle. This author reported that it was not significantly related to either organizational commitment or turnover. Jackson et al. (1991) reported that group age variability predicted group turnover rates. At the individual level using the D-score approach, however, these researchers reported that age dissimilarity was unrelated to turnover. Riordan and Weatherly (1999; cited in Riordan, 2000) reported that perceived age similarity was related to group performance, and that age similarity at the individual level was related to OCBs. The Timmerman (2000) study reported earlier on professional basketball and baseball teams reported results that were similar to those with regard to race; that is, age variability impacted basketball but not baseball team performance after controlling for average individual level performance. In this context, however, age effects might be overstated in relation to organizational contexts because age is such a salient issue in sports.

These studies, taken together, suggest that an employee's age-based fit within a team might impact team process and team performance. However, these studies also illustrate the fragmented nature of the approaches (and thus, results) in this literature. For instance,

using the D-score, some studies indicated dissimilarity was related to turnover but not to hypothesized intervening variables. This is especially surprising because turnover is dichotomous whereas the intervening variables are continuous; thus, correlations with the latter should be substantially easier to obtain. Similarly, Jackson et al. (1991) modeled relational age at both the individual and group levels but reported peculiar results: Relationships between age and turnover at the group level but not at the individual level. These seemingly contradictory results suggest these researchers might have omitted relevant variables that were perhaps 'deeper' psychological or motivational correlates of race and sex. One cannot tell, however, because the D-score approach is seriously flawed. Overall, it is clear that additional research with a clearer conceptualization of age dissimilarity would be informative. Thus, the following hypotheses are proposed:

> H17. The age-based fit of the individual within the team will predict individual turnover such that misfit is associated with a higher likelihood of turnover. This effect will be evident after controlling for the relevant personality constructs and ability.
>
> H18. Increased team age diversity will be related to increased team turnover rates. This effect will be evident after controlling for the relevant personality constructs and ability.
>
> H19. Increased age diversity will be negatively related to team performance. This effect will be evident after controlling for the relevant personality constructs and ability.

Sex. A number of authors have also examined the effects of sex diversity on individual and group outcomes in work settings. As a whole, these results are mixed. Several authors (Pelled et al., 1999; Riordan & Shore, 1997) examined group level sex heterogeneity and found no relationships to a host of outcomes including task and

emotional conflict, work group productivity, commitment, and cohesiveness. On the other hand, Harrison et al. (1998) reported that heterogeneity was negatively related to group cohesion. Similarly, Baugh & Graen (1997) found that gender homogeneous groups perceived their performance to be higher than non-homogeneous groups; however, there was no difference between these groups on perceptions of the quality of intra-workgroup working relationships or on external evaluations of workgroup performance.

At the individual level, most researchers have used the D-score approach. Chatappadaphay (1999) found dissimilarity as unrelated to peer relations and altruism, and Tsui et al. (1992) reported that it was unrelated to intent to stay with the organization. On the other hand, Riordan and Weatherly (1999; cited in Riordan, 2000) reported that actual gender dissimilarity using disaggregated group-level indices was unrelated to liking of peers or job satisfaction, but that perceived gender similarity was positively related to both outcomes. Other researchers (Flynn & Shore 1994, cited in Riordan, 2000; Riordan & Shore, 1997) using disaggregated group-level data to form interaction terms, reported finding no relationships with perceived group productivity, perceived advancement opportunities, perceived group communication, or satisfaction with several facets of the work environment. Again, like the studies examining age heterogeneity, this research reports mixed results using a variety of approaches, some of which are seriously flawed. Hence, the present study aims to examine these same issues using a novel, yet theoretically sound way of examining these issues. Thus, the following hypotheses will be tested:

H20. The sex-based fit of the individual within the team will predict individual turnover such that misfit is associated with a higher likelihood of turnover. This effect will be evident after controlling for the relevant personality constructs and ability.

H21. Increased team sex diversity will be related to increased team turnover rates. This effect will be evident after controlling for the relevant personality constructs and ability.

H22. Increased sex diversity will be negatively related to team performance. This effect will be evident after controlling for the relevant personality constructs and ability.

The Potential Moderating Effects of Time and Extraversion on Demographic- and Personality-Based Fit Effects

The sections above provide the rationale underlying the expectation that in general, demographic fit will be related to increased turnover at the team and individual level. Riordan (2000) suggested the impact of demographic diversity will be initially important but will fade over time as people have an opportunity to 'see past' surface-level features. Shaw and Barret-Power (1998) also allude to this notion, in that their model of demographic and cognitive diversity effects suggests that the former are more important at the outset of group development whereas the latter are linked to later stages. Indeed, this is consistent with Watson et al.'s (1993) research discussed earlier, which showed that racial homogeneity was related to positive outcomes early in team development but that these effects faded over time. Ancona and Caldwell (1998) reported similar results in that the effects on several demographic diversity variables were weaker in groups that had been established longer. The Chatman and Flynn (2001) study mentioned earlier directly assessed whether demographic diversity effects varied as a function of time. In

their first study, the results indicated that the negative effects of diversity on perceptions of cooperative group norms were mainly present early on in the group's lifespan. In another sample, negative effects of demographic diversity were only present among newcomers. Although these results should be interpreted with caution because D-scores were used, they still suggest that the negative effects of demographic differences might be ameliorated by time. Similarly, Harrison et al. (1998) found that gender (though neither racial nor age) diversity and affective diversity interacted with time such that the former was more strongly related to group cohesion in groups that had not been together very long. Conversely, underlying attitudes became a stronger predictor of group cohesion as time together increased. This suggests that the interpersonal interactions that naturally occur over time mitigate the effects of demographic diversity but that other types of diversity (e.g., attitudinal) that are only apparent after extended interactions have subsequent impacts. This is consistent with the underlying notion that individuals form similarity judgments based on the available information at a given time; it is logical to think that readily observable information such as demographics are used in people's early interactions because other information will be less available.

Chatman and Flynn (2001) indicate that time effects in relational demography research are inconsistent with the similarity-attraction paradigm because similarity does not change over time. In contrast, they argue that social categorization theory allows for changes in what is salient in a given situation. As time goes on and group members know each other more closely, other, 'deeper-level' similarity may become more important. Indeed, Shaw and Barret-Power (1998) discuss the notion of 'willingness to communicate' as a potential diversity management skill that might mitigate the negative

effects of diversity. Though these authors suggest this is likely to be most important with regard to deep-level diversity it seems likely that this would be true of demographic diversity as well.

Work by Jackson et al. (1992) and Arrow (1998) also suggests that demographic effects should change over time. These researchers see socialization, the process via which new employees become integrated into a group, as a critical mediator of the effects of demographic dissimilarity. In particular, these authors present a number of specific theoretical mechanisms via which dissimilarity should impact socialization, which in turn suggests that demographic effects should vary over time because socialization is a longitudinal process. While little empirical research to date has focused on the role demographic dissimilarity may play in socialization, this work nonetheless provides a framework that explains how well known organizational processes associated with socialization might be impacted by demographic dissimilarity.

Consistent with these notions is research reported by Flynn, Chatman and Sparato (2001) that studied Extraversion as a personality construct that may facilitate the extent to which these putative effects occur. Using MBA student teams working on a semester project, these researchers hypothesized that extraversion would mitigate negative impressions formed of other students that were demographically dissimilar. The rationale for this expectation was that individuals that are extraverted are more likely to get to know their team members more quickly than those who are introverted; hence, as interpersonal familiarity increases, similarity on demographic characteristics should become less important. Demographic dissimilarity was calculated by averaging dissimilarity of race, citizenship and sex at the individual level using the D-score

approach I critiqued earlier. The results were consistent with the hypotheses in that increased extraversion slightly attenuated the negative effects of demographic dissimilarity on impressions of their team members. However, a serious methodological caveat with this research includes the D-score approach that averaged three different demographic indices, further clouding inference. Nonetheless, these results and approaches, taken together, suggest the following hypotheses:

> H23. The fit between an individual's demographics and that of the team will be more strongly related to turnover at lower levels of tenure.
>
> H24. The effects of demographic fit within a team will depend on an individual's extraversion such that high extraversion individuals that do not demographically fit will be less likely to turnover than low extraversion individuals who do not fit.

The hypotheses formulated above all pertain to the effects of time and Extraversion on individual behavior. Team-level effects are not studied here because unlike laboratory studies the teams studied here are examined well after they have initially formed.

The Present Study

The present study broadly examines team composition as a predictor of team effectiveness in a large sample of fast-service restaurant teams. The goal is to contribute to the existing literature in a number of ways. First, there are virtually no large-scale studies of teams in the psychological literature. The benefits of large-sample studies have been widely acknowledged because they yield more power and more accurate parameter estimates. Second, this research examines financial performance as an indicator of team effectiveness. Though teams are more likely to yield such organizationally relevant

output, I have located no research on team composition and selection that has used such outcomes. Because each of the teams in this study can also be seen as an independent organization, this research answers calls to examine whether personnel selection has implications for organization-level effectiveness (Schneider, 1996; Schneider et al., 2000). In other words, this research examines mesotheoretical models linking individual, team and organizational effectiveness (House, Rousseau, & Thomas-Hunt, 1995; Rousseau & House, 1994). In addition, examining the financial performance of teams allows for extensions of existing theory and application of utility modeling (Zeidner & Johnson, 1991) because the financial value of performance is directly predicted at a higher level of analysis (Schneider et al., 2000). Third, this study examines multiple predictor constructs and multiple conceptualizations of team diversity. Very little research has formulated hypotheses incorporating both multiple predictors and multiple conceptualizations of team composition (McGrath, 1998). The addition of racial composition as a critical concern in personnel selection also adds another dimension not seen in existing research.

This research also seeks to make several contributions to the relational demography literature. First, there have been very few large sample studies in this area and there is a need for relational demographic studies in service teams (Tsui, Egan, & Xin, 1995). Second, this research seeks to clarify a number of theoretical and methodological limitations that are pervasive in the existing relational demography literature. Theoretically, the contribution is the conceptualization of demographic dissimilarity as either a cross-level effect or one that can be examined as a team-level diversity effect. Methodologically, this research relies on hierarchical linear modeling (Bryk &

Raudenbush, 1992) as a data analytic system that is both powerful and tied to an appropriate conceptualization of demographic similarity. This allows the study of individual and group effects using a unified analytic framework. Conceptualizing and analyzing the data in such a way is consistent with Riordan's (2000) call for relational demography research that is conceptually, methodologically, and analytically integrated. This research takes Riordan's notion even further by examining these effects at multiple levels of analysis. Practically, this should enhance our understanding of relational demographic effects by clearly specifying the level that these supposed effects might occur. Last, this research examines the impact of time and extraversion as potential moderators of demographic similarity effects. These represent potentially moderating effects that have only recently been incorporated into relational demography research, although numerous authors have suggested that composition effects should be studied over time (e.g., McGrath, 1998).

The research reported here is also one of the only studies to conceptualize and examine team effectiveness within the context of person-team fit, a notion that should relate to multiple aspects of team effectiveness including team performance and turnover. This is consistent with Moynihan and Peterson's (2001) 'contingent configuration' approach to the study of personality in teams. In other words, the nature of the relationship between personality and team outcomes is thought to depend on the personality dimension studied, the task-driven conceptualization of team personality level, and the outcome of interest. Second, by examining person-team fit as both a team-level diversity and cross-level effect, this research seeks to demonstrate how fit researchers can expand the domain of questions they ask. In other words, fit can be

related to both team and individual level outcomes. Similarly, examining variability at the team level is something that a number of authors have emphasized with regard to team composition research (e.g., O'Connor, 1998). Third, the multilevel nature of this research acknowledges arguments that it is unlikely that individual level phenomena will be unaffected by higher levels (Klein et al., 1994; Klein & Kozlowski, 2000, Rousseau & House, 1994; Schneider 1985). In particular, these authors suggest that contextual factors should be modeled when the existing knowledge base permits. Given the large knowledge base in personnel selection and team composition, it appears that the time is ripe for such research.

Finally, this research is conducted in jobs that are characterized by very low pay and high turnover. While it might be intuitively appealing to dismiss research using such samples as lacking any practical import, a large number of jobs like this are inevitably needed in our economy. Further, in light of the September 11th attacks, airport security screeners have come under increasing scrutiny as a possible weak link in our aviation security system. These screening jobs have been widely characterized as having low pay and very high levels of turnover. Thus, there is a clear practical need to further our understanding of team performance and turnover in jobs such as these.

Table 1 lists the hypotheses formulated in the research reported here. These hypotheses are divided into individual, team, and cross-level effects based on how they were described above, and based on the criteria of interest. The analysis method for each is also provided. Because of the complexity of the dataset used here, however, to fully understand the methods used to test these hypotheses, it may be helpful to refer to the Method section, presented below.

METHOD

Sample and Procedures

The sample for this study consisted of a total of 336,085 applicants and employees from 3,454 fast-service restaurant franchises across the country; however, the sample size varies widely depending on the analysis because there was a large amount of individual and team level missing data. The sample sizes for each of the predictors are given in the sections below. Each franchise is independently operated though there are standardized operating guidelines across the organization.

Hiring decisions for the organization were made based on the test used here; hence, this research uses a predictive validation design. Based on test scores the organization assigned candidates to one of six bands: Double Green, Green, Double Yellow, Yellow, Double Red, and Red. The three colors were based on norms that were designed to split the applicant sample into thirds. Restaurants were advised that candidates in the green band could safely be offered jobs but that they should use additional information sources to evaluate candidates in the yellow bands. Applicants in the red band were considered 'not recommended'. The actual cut points, and thus selection ratio, however, were determined by each individual restaurant.

Nature of the Team Task.

Crew members are entry level employees in these restaurants and are responsible for the following tasks: Taking orders, preparing food, delivering food to the customer, cleaning and maintaining the premises, and setting up and preparing equipment. In other words, the crew is entirely responsible for the day-to-day tasks associated with operating the restaurant. The objective of the team is to quickly provide customers with food of a

consistent quality, and in a pleasant manner that encourages them to return. The restaurants also must be kept physically clean for health reasons as well as to facilitate return business. Employees that directly interface with customers also have responsibilities that are directly related to the financial performance of the store in that they are required to ask customers if they would like to order certain accompaniment items such as soda. This task is very highly interdependent in that no one or two employees can complete any significant portion of the team goal on their own. Similarly, teamwork is involved because crew members must assist other team members who require work products or services (e.g., a cashier might have to ask another crew member to prepare a certain item). Crew responsibilities can also overlap. For instance, if a customer returns to the counter to ask for something additional or complain about an item, an employee besides the original cashier might be able to remedy what might otherwise be a less than satisfying dining experience. Similarly, because customers might emphasize any number of different aspects of the dining experience (e.g., speed of service, friendliness of service, quality of food), it is difficult to define this task as conjunctive. For instance, if a majority of customers emphasize fast service an employee who provides unfriendly service might not seriously impact team performance. Thus, to reiterate, the teams studied here do not neatly fall into any of the major task typology categories, though the situation is most closely related to the additive model; accordingly, the mean is used to index the team level of personality and ability standing.

Measures

Overview

Personnel records from 1995-1996 were used to obtain data for individual employees (N = 190,156). The dataset contains race, sex, hire date, termination date and termination reason for all selected employees. In addition, the last five jobs held within the organization are listed, as is the pay for each job, and the date at which the pay or job codes changed. The test used to hire employees provides the indicators of personality and ability studied here, and was first implemented in 1994, though it was not widely used until 1995. Store level financial performance and the number of transactions were available for 13 four-week periods beginning in mid 1995. In 1995, the first time period is Period 8 (P8), followed by P9, and so on, until P13. In 1996, the periods begin with P1 and go up to P7. So, in order, the 13 time periods are as follows: P8, P9, P10, P11, P12, P13, P1, P2, P3, P4, P5, P6, and P7. Individual level data were available from before P8 in 1995 and through P5 in 1996. Thus, the last two time periods, P6 and P7, are not used in any of the analyses reported here.

Independent Variables.

Selection Test. The test used in the current study was an 89-item paper and pencil measure (hereinafter called the CRI) designed for use in the fast service restaurant industry. Test booklets contained pages in both English and Spanish. Applicants were instructed to use the language in which they were most comfortable. Items were translated and retranslated and examined by several language and psychometric experts to ensure that they were parallel across the two languages. Test scores were available for 123,656 individuals in 2,991 restaurants. Of these, 53,661 were hired, representing an

overall selection ratio of 43.4%. Implementation of the test occurred unevenly throughout the organization. Prior to using the test this organization used a test that measured similar constructs. Discussions with informed personnel within the organization indicate that implementation varied considerably because of the decentralized nature of the individual restaurants. The test was based on a job analysis of jobs in different types of fast service restaurants (described below in the section titled 'validation study'). Test items were in either multiple choice or Likert format with response options anchored as follows: (1=Definitely True; 2=Somewhat True; 3=Don't Know; 4=Somewhat False; 5=Definitely False). A job analysis identified the following broad requirements for successful crew members: Work orientation, Fundamental Skills, Teamwork Orientation and Service Orientation. Although there was not a direct mapping of constructs specified by the test developers onto the FFM personality constructs used here, there was some consistency between the original construct labels and these constructs. In particular, Work Orientation items tended to resemble Conscientiousness, Service Orientation items Agreeableness, Team Orientation items Extraversion. Original construct linkages were evaluated and changed when necessary based on consultation with several influential papers on the FFM constructs (e.g., Barrick & Mount, 1991; Digman, 1990; Mount & Barrick, 1995). Examples of items on the Conscientiousness scale are "I consider myself a very dependable person" and "People say I am a very reliable person." Examples of Agreeableness items are "I often lose my patience with others" and "I have a pretty quick temper" (both reverse coded). Examples of Extraversion items are "I would prefer a job where I work by myself" and "I'm generally not interested in joining group activities" (both reverse coded).

The level of the fundamental skills items was aimed at the low cognitive complexity of these jobs. Hence, as the name of this scale implies, it was designed to assess whether candidates possessed minimum cognitive skills that were deemed to be required based on the job analysis. There were a total of 12 multiple-choice items on the fundamental skills scale, with 3 items pertaining to each of 4 problems. However, only 11 of the items were used because one of the items (#81) was not captured in the data archive file. Each problem presented a table with numeric and/or text information. Questions relating to the numeric tables required candidates to look up answers to the questions in the table. Those relating to tables with text were more complex; these items asked candidates to make logical inferences and mathematical computations based on the information provided in the question and the table. Thus, these items seem to require mathematical, verbal, and logical reasoning. Items were ordered so that they increased in difficulty. Cronbach's (1971) index of internal consistency reliability (α), for this scale was .67 It is not uncommon to see α's like this on relatively short scales with dichotomously scored items.

Construct Validity Study of Personality Items. A construct validity study of the personality scales described above was conducted to evaluate the extent to which they overlap with known construct-valid measures of the Big Five. This was done because there exists relatively little independent evidence of the construct validity of these items; however, based on item content it was expected that these scales would yield substantial corrected correlations with the known construct-valid scales. The items used as the referent, construct-valid measures of the Big Five were those on Goldberg's International Personality Item Pool (IPIP) website (http://ipip.ori.org/ipip/). The website presents a

large body of construct validity evidence for these items including correlations with a range of other commonly used and construct valid Big Five measures. 10 IPIP items were used to assess each of the Big Five constructs (see Appendix A).

The IPIP and CRI items were administered to 135 undergraduate psychology students who participated in exchange for extra course credit. The CRI scales were refined based on item statistics (means, standard deviations, and inter-item correlations and corrected item-total correlations), item content and their correlations with the IPI scales. As a result, one item was dropped from the original conscientiousness scale based on content ("Many supervisors expect employees to work too hard") and its corrected item-total correlation. The entire extraversion scale remained unchanged. The original agreeableness scale, however, posed more challenges in that it exhibited statistically significant, yet relatively low uncorrected correlations with the IPI agreeableness items ($r = .23$, $p < .05$). A shorter (10-item), though substantially less internally consistent scale, was thus formed that exhibited a much higher correlation with the IPI scale ($r = .47$, $p<.05$, new $\alpha = .66$, original $\alpha = .82$). This tradeoff was considered reasonable because the only hypotheses concerning agreeableness were at the store level, which yield scale scores that are considerably more reliable than what is indicated by internal consistency estimates because they are averaged across multiple individuals. Further, the average inter-item correlation was the same as on the original scale, indicating that the low internal consistency estimate was primarily due to a relatively low number of items and not their relationships with each other per se. Based on the applicant sample α's were as follows: Conscientiousness (.74), Agreeableness (.62), and Extraversion (.57). Some of these internal consistency estimates of reliability are less than optimal; indeed, they are

close to the lower bound specified by Nunnally (1978) for use in research settings (as opposed to use in making individual selection decisions). This is likely due to the short length of the agreeableness and extraversion scales, as the average inter-item correlations are similar to those on the conscientiousness scale. Given the large sample size used in this research this should not impact statistical power as adversely as otherwise might be the case. These effects will also be mitigated because aggregated data are used in many of the analyses (Rousseau, 1985).

The results of this construct validity study are summarized in Table 2, which presents the means, standard deviations, uncorrected correlations (upper diagonal), internal consistency estimates (diagonal), and correlations corrected using the internal consistency estimates (lower diagonal). As can be seen in the table, the corrected correlations between the CRI and IPI scales are .81, .82, and .65 for Conscientiousness, Extraversion, and the revised Agreeableness scale, respectively. This indicates that there is a substantial amount of overlap between the scales for all three constructs, although the strength of these results is somewhat less compelling for the Agreeableness scale. Accordingly, the results for this scale may have to be interpreted with some caution.

To assess the test-retest reliability of these scales, 29 of these 135 subjects took both the IPI and CRI items twice separated by a one-week interval. The results of the test-retest reliability analyses are presented in Table 3. As can be seen in the table, the test-retest reliabilities, on the diagonal, range from .82 - .83 for the CRI scales, and .83 - .91 for the IPI scales. This indicates that these items exhibit substantial stability over time, although one would certainly expect these correlations to decrease as the time interval between administrations increases.

Restaurant Level Test Saturation. Because the test was implemented unevenly across the organization it was necessary to examine the proportion of employees who were selected with the test used here. This proportion is hereinafter referred to as CRI Saturation. After examining these data it was decided that the last time period (P5) would be used for all the analyses including ability or one of the personality variables. This was done because this time period had the highest CRI saturation, and because restaurants that are among the first to implement a new test might differ on several important dimensions that might impact the results. By studying change over the last time period these effects should be mitigated, as more time has passed allowing the test to be more widely implemented. It should be noted that examining only the last time period as the outcome for the analyses using test scores would likely yield very similar results to the analyses reported below if other time periods were used because the average correlation between controllable profit at adjacent time periods was .87. For this time period, the CRI saturation was 39% (SD = .20). Although this is clearly less than ideal, the effects of test saturation will be modeled in this research. Specifically, saturation will be used as a moderator variable because stronger relationships should be evident in stores with higher CRI saturations. This is because there should be a more accurate index of the store-level standing on the constructs studied here as test saturation increases.

Validation Study. A concurrent validation study of this test was conducted in a number of franchises in this organization using both individual and store-level outcomes as criteria (N=883 employees in 101 restaurants). This serves as a kind of pilot study for some of the relationships examined in this research. Performance data consisted of supervisory ratings of individual crew members' job performance. These performance

ratings were gathered for research purposes only, and confidentiality of responses was assured and maintained throughout the process. Raters were trained on common rating errors and strategies for the avoidance of these errors. Because there are several supervisors with information on each crew member's job performance, care was taken to ensure the supervisor with the most knowledge of a given employee's performance was used to make the ratings for that employee. Training and ratings were made on company time at a regional meeting.

The performance instrument was developed specifically for the purposes of this validation study in this organization. The instrument asked for a total of 10 job skill ratings (e.g., dependability) and 12 work habit ratings (e.g., goes beyond required duties), half of which were positive and half of which were negative. There was also a single item asking for an overall rating of job performance. Each of the 22 dimensions as well as the overall-rating item was accompanied by a definition and behaviorally anchored rating scale. The results of a principal components analysis of these ratings yielded 4 interpretable factors, which correspond with some of the dimensions of the test used in this study (see Appendix B for performance dimensions and factor loadings, Appendix C for performance dimension intercorrelations and internal consistency estimates). These results indicate the performance ratings gathered were of relatively high quality; many researchers have noted that performance ratings often only yield a single interpretable factor (Campbell et al., 1993).

Several store level criteria were also collected in this study, namely, pre-advertising profit, cost of labor, and cost of sales across five periods in time. Team and Interpersonal Skills, Overall Performance, and a composite of the performance dimensions averaged

across employees within a store exhibited significant negative correlations with these store-level criteria (rs = -.22, -.21, and -.18, respectively). This is evidence of the construct validity of these performance ratings in that higher mean performance ratings were associated with lower costs of labor at the store level. These results are even more impressive because not every employee was rated within each store. Of course, it is also possible that store performance impacts mean job performance ratings at the individual level so the direction of the relationships here cannot be conclusively known.

As in a traditional validation study, correlations were calculated between the dimensions of the test and the job performance ratings. The results indicated a statistically significant and practically meaningful relationship between a composite total test score and a performance rating composite (r=.28; Appendix D presents scale-level validities for the four performance factors). Average test scores within each store were also examined in relation to the store-level criteria described above. The only relationship that was significant was the correlation between Fundamental Skills and the cost of labor (r = -.27). This provides some evidence that the test used here predicts store-level outcomes of interest.

Demographics. The following demographic variables were examined in this study: Race, sex, and age. Race was coded according to racial groups identified by the Equal Employment Opportunity Commission guidelines. Races that accounted for more than 1% of the total sample are: White (46.6%), Black (30.5%), Asian/Pacific Islander (2.8%), Hispanic (18.2%), and Native American (2.0%). The sample was 51.3% female. Table 4 provides a breakdown of these data. These data indicate the sample was diverse with regard to race and sex. Table 5 describes the sample's age.

There are two common formulas used to index the composition of a group on categorical variables such as race. A number of studies (e.g., Harrison et al., 1998) have used Blau's (1977) index of heterogeneity:

$$1 - \sum P_i^2$$

Where:

P = proportion of group members in a category

i = number of different categories in the team

The equation above indexes the heterogeneity in a given group for categorical data, and it varies from 0 to 1 asymptotically, with higher values indicating higher heterogeneity.

Other studies (e.g., Ancona & Caldwell, 1992) use Teachman's (1980) index of heterogeneity:

$$-\sum (\ln P_i)$$

Where the elements have the same meaning as those in Blau's (1977) index. A comparison of these indices under a number of conditions for 5 or fewer groups indicated that the correlation between them was above .95. This suggests that respective relationships with external variables will be nearly identical. Thus, the Blau index was used in analyses examining the team-level effects of racial diversity as a whole because it is more commonly used. In analyses where the dissimilarity between an employee and the racial composition as a whole is the focus, two variables are used to indicate the proportions of three racial groups in a given team. (Asians and Native Americans will not be examined in analyses using dummy coded variables to reduce the number of highly

correlated variables in the models. Because these groups represent such a tiny fraction of the sample, this should have a trivial effect, if any). Store-level composition with regard to sex was captured as the mean of a dummy coded variable where male=0 and female=1. Age composition was also calculated directly, as the mean or standard deviation of the ages of the crew members within a given restaurant, as appropriate.

Store Manager Tenure. The tenure of the restaurant general manager (RGM) at the end of data collection was provided.

Dependent Variables.

Dependent variables at the individual and team level were obtained from organizational records. Extensive discussions were conducted with representatives of the organization to ensure that criteria were both organizationally relevant and believed to be accurate/reliable. All the variables described below met these two criteria.

Turnover and Tenure.

The available data contains each employee's hire and termination date, and a reason for termination. Upon the termination of employment (by either party) the restaurant manager completed required paperwork giving a reason for termination. Discussions with informed organizational personnel indicated that these reasons for turnover are likely to describe actual turnover reasons. The organization separates these reasons into two categories: Voluntary and involuntary. Following the recommendation of Hom & Griffeth (1995), I further separated these into Avoidable and Unavoidable categories (only the former will be used in testing the IVT hypotheses). These reasons and their categorizations, as well as their incidence, are given in Table 6. Though it is difficult to determine whether there might have been some pressure exerted on managers to report

turnover as voluntary when an employee was actually fired (i.e., to perhaps avoid legal action), the fact that Avoidable Involuntary Turnover represents a sizeable portion of overall turnover indicates that this type of voluntary turnover still occurred at a relatively high rate.

Team Performance.

Controllable profit was used as the indicator of team performance in this study and was available from 2,502 stores for the 13 periods. Again, for the analyses involving test scores, the controllable profit at P5 was used, and the value at the previous time point (P4) was used as a control variable. Thus, the change in profitability was examined. In addition, the number of sales transactions for the period was also included as a control to help mitigate the effects of large-between store differences in the volume of transactions processed. These two control variables control for a wide range of between-store differences that might be related to the variables studied here. The client organization indicated that controllable profit was of obvious interest to them and it would be an appropriate, though stringent criterion to use in this research.

Control Variables.

Individual Pay. Because a number of turnover theories suggest that pay impacts turnover intentions, analyses were conducted to determine whether pay was related to any of the predictors used here. The data revealed that pay had near-zero relationships with the variables studied here, and thus it was not considered in any of the analyses.

RESULTS

Table 7 provides descriptive statistics and intercorrelations for the variables used in the team-level hypotheses reported here. As can be seen in the table, there was a

reasonable amount of variability in all the measures except for the aggregated personality scales, which are highly elevated and have very little variance, both as compared to the individual level student sample (see Table 2) and based on what one would expect given that the scale ranges from one to five. This is typical of personality items used in selection settings. Also of note is that of the average test scores, only ability exhibits a significant positive (though small) association with controllable profit for Period 5, which is the last period for which both individual and store level data were available. The results of more detailed analyses, which take into account the CRI saturation and other control variables, are described below.

The Relationship Between Mean Levels of Ability and Personality on Changes in Controllable Profit

Hypothesis 1 stated that mean levels of ability would predict the change in controllable profit after taking into account the number of sales transactions and CRI test saturation. It was also thought that the interaction between test saturation and ability would add to the prediction of controllable profit because one obtains a better indicator of ability in stores with higher test saturation. This hypothesis was tested using moderated regression where controllable profit for the prior period, sales transactions and test saturation were entered first, followed by ability scores, and then by the ability X test saturation interaction term at the last step. Table 8, which summarizes the results of these analyses, shows that the three predictors entered at the first step accounted for a very large proportion of the variability in profit, especially profit at the prior time period (β = .64; overall step R^2 = .854, $p < .05$). The incremental contribution of ability was significant (β = .03; ΔR^2 = .001, $p < .05$), though the ability X CRI saturation interaction

term was not significant (β = .03; ΔR^2 = .000, p > .05). To put the effect for ability in more easily understandable terms, a store that has an ability level .5 SDs above the mean generated an additional change of $128.57 in profitability per period as compared to a store whose employees scored at the mean on the ability test (this was calculated by multiplying the unstandardized regression coefficient by ½ the standard deviation for P5 controllable profit). Increasing ability scores by .5 SDs in every store throughout this sample would thus yield an additional $4.18 million in controllable profit over the course of a year ($128.57*2,502 restaurants*(52 weeks/4 weeks)). Thus, hypothesis 1 was supported.

A similar analysis was conducted for Hypothesis 2, which substituted conscientiousness for ability. The results in Table 9 indicate that this hypothesis was not supported, as the increment of conscientiousness was not significant after accounting for the control variables (β = -.01; ΔR^2 = .000, p > .05). Similarly, the conscientiousness X CRI saturation interaction was non-significant as well (β = .04; ΔR^2 = .000, p > .05). Hypothesis 3, which stated that conscientiousness was expected to contribute incrementally after controlling for the effects of ability, was not tested because the results of Hypothesis 2 indicate that this effect will be non-significant.

Table 10 summarizes the results relating to Hypothesis 4, which was parallel in structure to Hypothesis 2 except that agreeableness is the focal construct. As can be seen in the table, agreeableness did not significantly predict a change in controllable profit after the first block of variables were entered (β = -.02; ΔR^2 = .000, p > .05); Consistent with the results reported earlier, the interaction term between test scores and agreeableness did not add to the prediction of change in controllable profit. Thus, neither

Hypothesis 4, nor Hypothesis 5, which pertained to the incremental contribution of agreeableness, was supported

Table 11 summarizes the results relating to Hypothesis 6, which focused on extraversion exhibiting an inverted-U relationship with change in controllable profit. Because the interaction between CRI saturation and the other personality and ability variables were non-significant in the analyses reported above, the product term was dropped from this analysis. At Step 2, extraversion added to the prediction of controllable profit but this effect was negative (β = -.02; ΔR^2 = .0004, $p < .05$) such that higher mean levels of extraversion were associated with an decrease in profitability. The polynomial term, entered at Step 3, was not significant (β = .08; ΔR^2 = .000, $p > .05$); hence, Hypothesis 6 was not supported.

Hypothesis 7 stated that manager tenure would interact with the mean levels on the ability and personality variables within a store to predict change in profits. The results of these analyses are summarized in Tables 12, 13, and 14. In each of the analyses, the control variables were entered in the first step, followed by the ability or personality predictor, manager tenure, and then the interaction term in the third step. There was no evidence for a significant effect for interactions associated with conscientiousness (β = .14; ΔR^2 = .000, $p > .05$), agreeableness (β = .11; ΔR^2 = .000, $p > .05$), or ability (β = -.02; ΔR^2 = .000, $p > .05$). Thus, these hypotheses were not supported.

The Effects of Individual Standing on Ability and Personality Variables on Turnover

These hypotheses were tested with Cox regression (Cox, 1972), which evaluates the effects of a set of covariates on the instantaneous risk of encountering a particular event. Numerous authors (e.g., Harrison, 2002) have suggested this approach because it uses

data from cases that are censored (i.e., those individuals that have not turned over by the end of data collection). In addition, because time is treated continuously and is evaluated with respect to risk for turnover, more information is gleaned as compared to approaches that treat turnover as a dichotomous variable. Much like logistic regression, testing series of Cox regression models yields differences in model fit distributed as χ^2 which can then be tested for statistical significance. Parameter estimates can be anti-logged to yield estimates of the changes in risk due to a given change in the covariate or predictor.

Hypothesis 8 suggested that conscientiousness would be negatively related to the risk of voluntary turnover. Of those that had conscientiousness scores, 22,606 experienced voluntary turnover during data collection, while 15,959 were censored. Conscientiousness was a significant predictor of turnover risk ($\Delta\chi^2 = 37.13$, B = .11, $p < .05$.). Anti-logging the B parameter (hereinafter referred to as Exp(B)) yields a value of 1.11, indicating that a one-unit increase in conscientiousness results in a risk of turnover that is 111% of the original risk. This does not support Hypothesis 8 because higher conscientiousness scores were predicted to lower one's likelihood of voluntary turnover.

Hypothesis 9 was identical lto the previous hypothesis except that it specifed involuntary rather than voluntary turnover. This analysis was based on 10,518 employees with conscientiousness scores who experienced involuntary turnover, and 15,959 who were censored. The results of this analysis did not support the hypothesis because conscientiousness was unrelated to the risk of involuntary turnover ($\Delta\chi^2 = .47$, B = .02, $p > .05$.). (Because conscientiousness was not related to turnover it was not used as a control in any subsequent hypotheses concerning demographic diversity and turnover).

Hypothesis 10 stated that high ability employees would be more likely to exhibit voluntary as opposed to involuntary turnover. This hypothesis was tested using two separate survival analyses – one for each type of turnover. For voluntary turnover (N = 22,596; censored N = 15,948), higher ability was associated with a slightly lower risk of turnover ($\Delta\chi^2 = 23.28$, B = -.013, SE = .003, $p < .05$., Exp(B) = .987). For involuntary turnover (N = 10,511, censored N = 15,948), higher ability was also associated with a slightly lower risk of turnover ($\Delta\chi^2 = 201.28$, B = -.06, SE = .004, $p < .05$., Exp(B) = .944). These results are not consistent with the hypothesis.

Hypothesis 11 stated that increased diversity in conscientiousness would be associated with decreases in store performance. In this analysis, after the control variables were entered, the coefficient of variation for conscientiousness was entered. This coefficient is calculated by dividing the standard deviation by the mean as a way of expressing the former controlling for the latter. As can be seen in Table 15, the increment in this step was not significant ($\beta = .01$; $\Delta R^2 = .000$, $p > .05$). A similar exploratory analysis was conducted with agreeableness, yielding similar results ($\beta = .01$; $\Delta R^2 = .000$, $p > .05$).

Hypotheses 12 and 13 made predictions about the relationship between conscientiousness diversity and voluntary and involuntary turnover rates. This relationship was examined for Period 4 because there were no termination reasons recorded for terminations that occurred during Period 5. These hypotheses were not supported, as the correlation between the conscientiousness coefficient of variation and voluntary and involuntary turnover rates was not significant ($r = .02$ and $r = .01$, respectively, $p > .05$).

The Relationships Between Demographic Diversity and Profitability and Turnover Rates

These hypotheses – 15, 16, 18, 19, 21, and 22 – were tested using latent growth curve (LGC) methodology in structural equation modelling (Meredith & Tisak, 1990; Stoolmiller, 1995; Willett & Sayer, 1994). LGC modelling uses a latent intercept and slope representing the starting point and change over time, respectively, of a given variable with repeated measurements. The slope can be linear or a higher order polynomial (e.g., quadratic), depending on the hypothesized or observed pattern of change over time. The mean of the slope and intercept parameters define the overall trajectory for the entire sample on which they are calculated. The statistical significance of their variances indicates whether there is a significant amount of inter-individual variability in these trajectories. If one can identify a model that adequately describes the latent form of the change pattern underlying the observed data, and if the slope and intercept variances are significant, one can then seek substantive predictors of these two parameters. In this way, LGCs make it possible to identify correlates of the starting point (intercept) or change over time (slope) in a given set of individuals or to relate intercept and slope parameters across different variables (e.g., changes in diversity and changes in profitability). Analyses of the latter type, which are used to test the hypotheses in the research reported here, have been termed cross-domain LGCs (Chan, 1998).

Although LGCs are traditionally used to model individual change over time, in the present study a number of store-level changes were of substantive interest. In particular, the hypotheses mentioned above identify demographics as potential correlates of profit and turnover rates. The dynamic nature of these criteria make them well suited to the

LGC approach described above; however, rather than identifying a single static predictor of changes in profit and turnover rates, the analyses conducted here also take into account the dynamic nature of the predictors. Thus, for this study the intent was to estimate separate LGC models for each of the five variables in these hypotheses: Age, sex, and racial diversity, controllable profit and turnover rates. Once good-fitting models were identified, the respective models were combined as appropriate to see if their slopes and intercepts covary. This allows one to examine whether the starting point or rate of change for the diversity variables is related to the starting point or rate of change in profitability or turnover rates over time.

A graphical representation of a simple LGC model is depicted in Figure 4 (Chan, 2002). The variables Y1 – Y4 represent a variable measured at four different points in time. Associated with each of these variables is a disturbance or error term (ε_1 - ε_4), which includes both the prediction and measurement error associated with each. In addition, each has a factor loading on the intercept (λ_{1I} - λ_{4I}) and slope (λ_{1S} – λ_{4S}) latent factors. In order to define the intercept as such, the factor loadings are constrained to one. For the slope, two loadings have to be fixed at zero and one to facilitate model identification, while the remaining loadings can be fixed to represent other forms of growth (Meredith & Tisak, 1990). For example, a linear growth pattern would fix the remaining two loadings at two and three, whereas a quadratic pattern would specify an additional latent slope factor where the the third and fourth factor loadings are fixed at 4 and 9 (2^2 and 3^2, respectively). Alternatively, if no particular form of slope is expected, or if no easily recognizable form is seen in the data, one could freely estimate the final two slope parameters. As mentioned earlier, the intercepts and slopes also have a mean

(μ_I and μ_S) and variance (σ^2_I and σ^2_S), and a covariance among each other (σ_{SI}). This general model, discussed below, was implemented in the LGCs described here.

LISREL 8 (Jöreskog & Sörbom, 1993) was used to estimate the LGCs reported here. To assess the adequacy of the various models tested, five fit indices were used to obtain convergent evidence of model fit. This has become common practice in studies using structural equation modeling because the original χ^2 fit index is highly sensitive to sample size in that models with excellent fit will often yield significant χ^2 values falsely indicating the hypothesized model does not fit if the sample is large enough. The indices used were: The root mean square error of approximation (RMSEA; Steiger, 1990); normed and nonnormed fit index (NFI and NNFI, respectively; Bentler & Bonnett, 1980); adjusted goodness of fit index (AGFI; Jöreskog & Sörbom, 1989); and standardized root mean residual (SRMR; Jöreskog & Sörbom, 1986). All of these fit indices range from 0 to 1, with 1 indicating better fit except for SRMR and RMSEA. For the former category of fit indices, values above .90 are considered good fit, whereas for the latter, values of .10 or less are considered to be good fit and values of .05 or less are considered close fit. In addition, modification indices (MIs), standardized residuals and standardized expected change parameters for particular models were examined to determine the extent to which fit might be improved by estimating or constraining specific parameters. Also considered in these decisions were substantive concerns (e.g., whether fixing a freely estimated parameter makes sense given the research design and hypotheses of interest).

Before any of the LGC models were tested, the bivariate correlations between the three diversity variables and controllable profit and turnover rates were examined. This was done because there is little to be gained by estimating LGC models for variables that

might not be correlated in the first place. This was the case for all the diversity variables and turnover rates. In particular, the mean same-period correlation between each of the 11 4-week period turnover rates and racial, age, and sex diversity was .03, .01 and .01, respectively. In addition, the average correlation among the monthly turnover rates was -.04, and the maximum correlation was only .02. This does not support the notion that there is a readily identifiable latent variable driving the period turnover rates. Indeed, it was not possible to achieve a good fit for a LGC model with these data, as several of the fit indices were stuck below .90. Despite this, it nonetheless seemed worthwhile to investigate whether turnover rates aggregated over a longer time period might yield more positive results, especially given earlier studies (e.g., McElroy, Morrow, & Rude, 2001; Shaw et al., 1998) that have done so. Table 16 shows the correlations among the aggregated turnover rate and diversity variables. As can be seen in the table, the correlations of interest are all significant, though relatively small. As predicted, higher diversity of all three types was associated with higher turnover rates. To assess the joint impact of these diversity variables on turnover rates, a multiple regression analysis was conducted where all three diversity variables were entered in a single step with the turnover rate serving as the dependent variable. As can be seen in Table 17, the partial regression coefficients for race and sex diversity were significant ($\beta_{racediv}$ =.07; β_{sexdiv} = .08; ΔR^2 = .015, $p<.05$), though the partial effects of age diversity were no longer significant ($\beta_{racediv}$ =.02). Thus, hypotheses 15, 18, and 21 received bivariate support, although the LGC methodology originally proposed to test the hypotheses was found to be inappropriate due to the lack of dimensionality underlying the 4-week turnover rates.

Returning to the correlation matrices underlying the proposed LGC models, for sex diversity and controllable profit the average same-period correlation was only .02, and inspection of the scatter plots did not reveal any discernible patterns (e.g., nonlinearity). Thus, hypothesis 22 was not supported. In contrast, there were larger and more consistently significant relationships between age and racial diversity and controllable profit. Thus, the following sections describe the development of the individual growth models for these three variables. After each of these models is developed, two combined models are discussed that examine whether the intercept and slope parameters of the diversity variables are significantly correlated with those of controllable profit.

LGC Model for Racial Diversity

Recall that racial diversity is operationalized as $1-\sum P_i^2$ where P is the proportion of members in i groups. This index was calculated for each of the 11 time periods. The means, standard deviations and intercorrelations of these values were used as LISREL input in fitting the initial model. These data, presented in Table 18, indicate that the mean and standard deviation are almost constant across time. However, a plot of the individual store trajectories showed substantial variability in the nature of change in racial diversity over time. The change in this index over time for a random sample of 15 restaurants is presented graphically in Figure 5. As can be seen in the figure, there is no easily identifiable function that might be used to represent the slope in these data; thus, the last nine of the eleven slope parameters were freely estimated. Similarly, there appears to be reasonable variability in the intercept and slopes across the stores. If this variability can be explained, the data may be more interesting than the constant pattern of means and standard deviations in Table 18 suggests. Another pattern worthy of note is

that the data in Table 18 exhibit serial autocorrelation in that the correlations are substantially higher as the periods become closer in time. After parameters in an initial model were estimated, and after the MIs were considered in light of these data, it was decided that adjacent error covariances of the 11 indicator variables should be freely estimated to improve model fit. In addition, the MI for the error covariance between P3 and P5 was 284, so this parameter was allowed to be freely estimated as well. This yielded a final model depicted in Figure 6. The fit indices for this model suggested a good fit to the data (NFI = .99, NNFI = .99, CFI = .99, AGFI = .93, RMSEA = .08, SRMR = .05, χ^2 = 593.53, df = 41, $p<.05$). In this model the pattern of slope factor loadings supports the notion that neither a linear nor quadratic slope would accurately represent these data. A plot of the standardized LISREL estimates of the slope factor loadings is presented in Figure 7. These might be more appropriately called "shape patterns" (Chan, 2002) or "linear splines" (Meredith & Tisak, 1990), but the original "slope" designation is used here to be more broadly consistent with the LGC literature.

Chan, Ramey, Ramey, and Schmitt (2000) note that much like in regression, in LGC analyses, one can test the incremental fit provided by estimating a freeform slope as compared to a linear or quadratic slope. Thus, models identical to the one described immediately above were tested, with the slope factor loadings constrained to either a linear or quadratic form. The relative fit of these models was tested with the commonly used χ^2 difference test, although differences in the practical fit indices were examined as well. The model with a linear slope (i.e., one with slope factor loadings fixed to 0 – 10 for the 11 time periods) fit the data relatively well as indicated by the practical fit indices (NFI = .98, NNFI = .98, CFI = .98, AGFI = .87, RMSEA = .10, SRMR = .05). The χ^2

value for this model was 1132.44 with 50 degrees of freedom. Thus, the $\Delta\chi^2$ was 538.91, with $\Delta df = 9$ ($p < .05$). This indicates this straight-line trajectory model fits the data significantly less well than the freeform slope described above. This is consistent with the slight decrement in the practical fit indices. An additional model was estimated that had two latent slope factors – one with a linear trajectory and another that was quadratic. The practical fit indices for this model were nearly identical to that of the initial model (NFI = .99, NNFI = .99, CFI = .99, AGFI = .92, RMSEA = .08, SRMR = .01), though the χ^2 value was somewhat larger ($\chi^2 = 728.17$, $df = 46$), a difference that was statistically significant ($\Delta\chi^2 = 134.64$, with $\Delta df = 5$, $p < .05$). This also supported the notion that the freeform slope model provided the best fit to these data.

LGC Model for Age Diversity

Age diversity was operationalized as the standard deviation of employee age in days (age was calculated when the employee was hired). LISREL could not calculate starting values with the resulting large numbers in the fitted covariance matrix, and so this variable was divided by 100, which eliminated the problem. The data input to LISREL, after applying this transformation, are presented in Table 19. As with race diversity, the means and standard deviations show a very consistent pattern across time. To better understand the nature of the changes in age diversity over time a plot of a random sample of 15 restaurants was examined (see Figure 8). Seeing a similarly unclear pattern as with the racial diversity profile, the form of the slope parameter was again freely estimated. This yielded a slope factor depicted in Figure 9. As in the previous models, based on the MIs, the error covariances of adjacent time points were allowed to vary, as were those between several variables only two time periods apart: P13 and P2, P11 and P13, and P3

and P5. The resultant model, shown in Figure 10, fit the data very well (NFI = .98, NNFI = .98, CFI = .98, AGFI = .94, RMSEA = .08, SRMR = .08, χ^2 = 543.89, df = 39, p<.05). The same model with a linear slope fit the data significantly less well (NFI = .97, NNFI = .96, CFI = .97, AGFI = .89, RMSEA = .09, SRMR = .08, χ^2 = 1,085.77, df = 48; $\Delta\chi^2$ = 541.88, with Δdf = 9, p < .05). The model with two latent slopes – with linear and quadratic shapes – also fit the data significantly less well as compared to the original model (NFI = .98, NNFI = .98, CFI = .98, AGFI = .93, RMSEA = .08, SRMR = .04, χ^2 = 666.61, df = 44; $\Delta\chi^2$ = 122.72, with Δdf = 5, p < .05). This supports the use of the freeform slope model for age diversity.

LGC Model for Controllable Profit

As with age diversity, LISREL was unable to obtain starting values due to the scale of controllable profit. Thus, this variable was divided by 100 to facilitate the computation of starting values. The LISREL input is provided in Table 20. As with the previous variables, the trajectories of a random sample of 15 cases was examined (Figure 11). The figure suggests that estimating a free-form slope would be appropriate in this situation. After examining an initial model with error covariances constrained to zero, it was decided that adjacent error covariances should be freely estimated. The resulting model, shown in Figure 12, provided a good fit to the data without any additional modifications (NFI = .98, NNFI = .98, CFI = .99, AGFI = .93, RMSEA = .08, SRMR = .02, χ^2 = 643.16, df = 42, p < .05). Both in relative and absolute terms, a linear growth model fit the data very poorly (NFI = .88, NNFI = .88, CFI = .89, AGFI = .75, RMSEA = .21, SRMR = .05, χ^2 = 4,536.89, df = 51; $\Delta\chi^2$ = 3,893.73, with Δdf = 9, p < .05), as did a model with linear and quadratic slope factors (NFI = .94, NNFI = .93, CFI = .94, AGFI =

.90, RMSEA = .15, SRMR = .04, χ^2 = 2,406.88, df = 47; $\Delta\chi^2$ = 1,763.72, with Δdf = 5, p < .05). This supports the use of the freeform slope model for controllable profit (see Figure 13).

Combined LGC Model for Racial Diversity and Controllable Profit

Table 21 presents the intercorrelations between racial diversity and controllable profit over the 11 time periods. As can be seen in the table, these correlations are all significant, and the average correlation is .13. This supports the notion that investigating a combined model where the slopes and intercepts of these two variables are allowed to covary is worthwhile. The data provided in Tables 18, 20, and 21 were used to estimate such a model. However, there were estimation problems with the freely estimated factor loadings of the racial diversity slope that prevented LISREL from achieving stable starting values. Fixing the racial diversity slope factor loading for P1 and P2 to the maximum likelihood LISREL (i.e., unstandardized) estimates obtained in the separate LGC race model depicted in Figure 6 solved this problem. The resulting model fit well but the MI for the error covariance of racial diversity at P1 and P12 was 184. Because this was a large MI for variables that were only two time periods away, this parameter was freely estimated. The final model allowed the four latent variables to covary, while the separate racial diversity and profit LGC models were identical to those discussed earlier except for the two fixed racial diversity slope factor loadings and the freely estimated error covariance mentioned directly above.

The fit indices indicated this model fit the data well (NFI = .99, NNFI = .99, CFI = .99, AGFI = .95, RMSEA = .05, SRMR = .03, χ^2 = 1,178.78, df = 201, p <.05). The associated parameter estimates are presented in Figure 14. As can be seen in the figure,

the race intercept is significantly associated with both profitability parameters. In particular, it is positively associated with the controllable profit intercept (standardized effect size estimate = .11, $p < .05$), but negatively associated with the slope (standardized effect size estimate = -.16, $p < .05$). This means that at the beginning of data collection racially diverse stores also tended to be slightly higher in terms of profitability. However, racial diversity was also associated with a negative change in profitability over time.

To put these effects in more concrete terms, the impact of racial diversity on profit trajectories was calculated for stores at -1, -.5, 0, +.5, and +1 standard deviation (SD) units from the mean of racial diversity using the equation below:

$$\hat{P}_{RDSDk} = \left[(.11RDSD_k) + \overline{P}_0\right] + \left[\sum_{j=1}^{10}\left(\lambda_{PSj+1} - .16RDSD_k\right)\left(PSD_{j+1}\right) + \overline{P}_{j+1}\right] \qquad \text{(EQ1)}$$

Where:

$\hat{P}_{RDSDk}$ = the total predicted profitability across the 11 time periods for a given standing on racial diversity in k SD units;

$RDSD_k$ = the deviation from the racial diversity mean in k SD units where k = -1, -.5, 0, +.5, and +1.0;

$\overline{P}_0$ = the mean profit at the first time period

j = the 10 intervals separating the 11 time periods;

λ_{PSj} = the completely standardized factor loading for the profitability slope at time j+1;

PSD_{j+1} = the SD of profit at time j+1;

$\overline{P}_{j+1}$ = the mean profit at time j+1.

This equation was used to calculate five predicted totals and trajectories for controllable profit, each associated with one of the k values of racial diversity expressed in SD units. As can be seen in the equation, the positive effect of racial diversity on the profitability intercept is taken into account by the terms on the left of the summation. In particular, the SD units for a given level of racial diversity is multiplied by .11, the correlation between the racial diversity and profitability intercepts, which is then added to the mean profit at time one (P8). So, for a restaurant 1 SD above the mean on racial diversity the predicted starting point on profitability is (.11*1SD*$9,041/SD) + $19,921 = $20,915.

The terms to the right of the summation represent the predicted profit for a given level of racial diversity in SD units for each of the 10 time intervals. In the first set of parentheses after the summation, the predicted profit slope factor loading is calculated based on the initial slope factor loading, the racial diversity SD, and the correlation between the racial diversity intercept and profit slope (-.16). Thus, for the first interval (i.e., in between P8 and P9), where the profit slope factor loading is .04, the predicted profit slope factor loading for a restaurant 1 SD above the mean on racial diversity is (-.16*1SD)+.04 = -.12 profitability SD units. This value is then scaled by the profit SD for a given time point, and added to the mean profit at that time point to obtain the predicted profit for that time point given the standing on racial diversity. Thus, for the example using +1 SD units above the racial diversity mean, the terms for the first and last time periods are:

$$[(-.16)(1SD)+-.11](\$8{,}555/\text{SD})+\$17{,}929$$

and

$$[(-.16)(1SD)+.04](\$8{,}970/\text{SD})+\$20{,}654$$

Summing across the 10 time intervals and taking into account the initial starting point, discussed above, thus yields the total predicted profit for the 11 time points for the given level of racial diversity in SD units.

Figure 15 presents the five predicted profitability trajectories associated with these levels of racial diversity. The figure shows that stores higher in racial diversity start off at higher levels of profitability, but that the effect on the slopes offsets this effect after the first time period (P9). Stores at higher levels of racial diversity are projected to continue at lower levels of profitability throughout the data collection period. Using equation 1, the total profits were calculated for the time period under study for these five levels of racial diversity. Figure 16 shows the linear association between standardized racial diversity levels and the predicted total predicted profits over the course of a year. A 1 SD difference in racial diversity amounts to $14,656 per restaurant. To make the impact of the racial diversity metric more concrete, several possible store compositions were examined to assess the predicted effects of racial diversity on profitability over the course of one year (the predicted values calculated using EQ1 were multiplied by 52/44 to scale the 44-week predicted total to an annual predicted total). Figure 17 presents these results graphically. Each pattern represents a different racial group and the smallest slice (e.g., one of the two small slices in the Chart 2 of Figure 17) represents one person. These charts reinforce the notion that small differences in racial diversity are predicted to make sizable differences in controllable profit (again, note that these charts show the predicted

effects of racial diversity, rather than those associated with the proportion of a particular racial group). For instance, the difference between Charts 5 and 6 of Figure 17 is two members from one racial group, but the difference in predicted annual profit between the two compositions is almost $13,000. Similarly, Chart 6 has only one member from the non-majority group in the restaurant whereas Chart 4 depicts a restaurant where half the members are of the same race and the remainder is made up of three different races. This translates into an estimated annual difference in profit of over $41,000. Applying these numbers across the entire organization of over 2,000 stores results in differences that are practically meaningful changes in controllable profit.

Combined LGC Model for Age Diversity and Controllable Profit

Table 22 presents the intercorrelations between age diversity and controllable profit over the 11 time periods. As can be seen in the table, these correlations are generally significant, though small. This suggests that investigating a combined model where the slopes and intercepts of these two variables are allowed to covary is worthwhile. The data provided in Tables 19, 20, and 22 provided the data used to estimate this model. As with the combined model for racial diversity, there were estimation problems with the freely estimated factor loadings of the age diversity slope that prevented LISREL from achieving stable estimates, and so two loadings were fixed using their LISREL estimates from the age diversity model. The combined model fit the data very well (NFI = .98, NNFI = .98, CFI = .98, AGFI = .94, RMSEA = .05, SRMR = .05, χ^2 = 1,336.37, df = 200, $p < .05$). Figure 18 provides a representation of this model along with the standardized effect size estimates. As can be seen in the figure, the age diversity intercept was positively and significantly associated with the profitability intercept

(standardized effect size estimate = .10), though none of the other parameters of interest were significant. Thus, hypothesis 19 was not supported.

Relationship Between Demographic Fit Within a Store and the Probability of Turnover

Hypotheses 14, 17, and 19 stated that the race, age, and sex-based fit of an individual within a store are expected to predict individual turnover probabilities. Earlier, a number of reasons were discussed as to why survival analysis is well suited toward the prediction of such probabilities. Here, this approach is extended in a multilevel framework where store composition is used as a moderator of the individual-level relationships between these demographic variables and turnover. This was done using hierarchical nonlinear modeling (HnLM) using the HLM software package (Bryk, Raudenbush & Congdon, 1996). At the individual level, the logit link function was used to model the relationship between a given demographic characteristic and turnover within each store. The variability in the individual-level parameters is then modeled as a function of store-level composition variables. If these store-level composition variables do indeed predict the variability in the individual level parameters (i.e., the extent to which a given demographic variable is related to turnover probabilities), then the form of the relationship will be inspected to determine whether it is consistent with the hypotheses.

Because continuous time multilevel survival analysis models are not as fully developed in comparison to discrete time models, the latter approach was taken here. Transforming continuous measurement of time (here it was measured in days; i.e., the day hired and terminated) is referred to as coarsening in the event history modeling literature. A number of researchers have noted that coarsening does not pose serious estimation or bias problems as long as the coarsening occurs at random and no more than

roughly 5% of the events (i.e., turnover) occur at a given interval (Allison, 1984). Both of these conditions are met here. For the purposes of this study, a period of two weeks was chosen. This was done because this was a relatively high turnover job, and an inspection of the survival curve which suggested that a longer interval would lose some information about the survival probability over time. In addition, with this approach there is actually relatively little coarsening because many employees' terminations were registered at two-week intervals (i.e., at the end of a pay period). In other words, if one were to examine a bar chart plotting the number of terminations over time there would be a series of small spikes spaced two weeks apart.

This approach resulted in 35 periods (i.e., 70 weeks) from which individual hire and termination data were available. If an employee did not turnover by the end of the 35th period, his or her tenure is censored (i.e., turnover did not occur by the end of data collection). Singer and Willett (1993; Willett & Singer, 1991) indicate that standard logistic regression software yields maximum likelihood estimates that can be used to test the hypotheses associated with survival analysis. However, to allow for this, the cases must be restructured into a person-period format. An example of this data structure is provided in Table 23. The table contains records for three hypothetical individuals who worked in the same store for a total of 40 two-week periods for which data were available. As can be seen in the table, there is one record for each of these two-week periods. Based on the third column, one can see that employees 1 and 2 turned over after working for three and two time periods, respectively; employee 3 worked for the entire span of data collection and was censored because the turnover event had not occurred at the end of data collection. In the fourth column the sex of each employee is entered on

all the records for that employee. In the last two columns, we have the store sex composition calculated for each period the employee worked, and the product between the individual's sex and the store sex composition (the treatment of these last two variables in the analysis is discussed later). The sex compositions are not the same for employees 1 and 2 because the time periods are not the same – rather than representing an absolute point in time, they represent, in sequence, the number of periods a given employee worked (i.e., each employee's time starts at one). This was done to ensure that the effects of time are modeled consistently across employees. Before proceeding with the multilevel survival analyses, however, a single level analysis was conducted to better understand the form of the survival function. This was done by conducting a logistic regression on the individual-level person-period dataset structured as described above (N = 1,319,180 person-periods; N = 177,873 employees; censored N =29,515; these figures also describe the sample size for the demographics-based multilevel analyses reported below). In this analysis turnover was regressed on period and period^2, and predicted turnover hazards at each of the t periods were calculated using the following equation (Tabachnick & Fidell, 1996):

$$P(Turnover)t = \frac{e^{A+B_1 Period_t + B_2 Period_t^2}}{1+e^{A+B_1 Period_t + B_2 Period_t^2}} \qquad \text{(EQ2)}$$

The results of this analysis are shown graphically in Figure 19 (P(Turnover) = -1.85 - .04*(Period) +.001*(Period^2), -2 Log Likelihood = 924,926; all parameters significant at $p<.05$). This figure plots cumulative survival as a function of time. As can be seen in the

figure, the survival curve closely resembles a quadratic function. The shape of the curve and the significance of the quadratic term suggest that a quadratic term should be used in the multilevel analyses.

One of the challenges alluded to above is how to handle the dynamic nature of the store compositions. Fortunately, survival analysis at the individual level can easily handle covariates that change over time. Similarly, multilevel survival analyses can handle time-varying covariates, either at the individual or group level, though implementations of the latter are extremely rare. Barber, Murphy, Axinn, and Maples (2000) present the only discussion of which I am aware that deals with how to handle this latter situation. Because this is not yet commonly done, a brief explanation of the analysis is provided here, though the reader is referred to Barber et al. (2000) for a more detailed discussion.

For this example the sex will be used because a dichotomous variable simplifies the the explanation. Based on the hypotheses formulated earlier in this paper, one would expect to find that the store-level sex composition (i.e., proportion of employees of a given sex) moderates the relationship between sex of the employee and the probability of turnover. Recall that effects such as these are commonly referred to as cross-level interactions. In addition, it is necessary to include variables establishing the baseline survival function. One way to do this is to enter periods and the periods^2 to allow for a

quadratically shaped survival curve.[1] The conceptual model in a traditional two-level form, as in Bryk and Raudenbush (1992), is as follows

$$Logit(p_{ijk}) = \beta_{ok} + \beta_1 Period_{ij} + \beta_2 Period_{ij}^2 + \beta_{3k} Sex_{jk} \qquad \text{(EQ3)}$$

$$\beta_{0k} = \gamma_{00} + \varepsilon_{0K} \qquad \text{(EQ4)}$$

$$\beta_1 = \gamma_{10} \qquad \text{(EQ5)}$$

$$\beta_2 = \gamma_{20} \qquad \text{(EQ6)}$$

$$\beta_{3k} = \gamma_{30} + \gamma_{31} StoreSex_{tk} + \varepsilon_{3K} \qquad \text{(EQ7)}$$

Where:

β_{0k} = the overall turnover probability in store k

β_1 = the effects of time on the turnover probability for person j at time t

β_2 = the effects of time^2 on the turnover probability for person j at time t

β_{3k} = the sex effect on the turnover probability for person j, which varies as a function of the store sex composition in store k at time t

$\varepsilon_{0K}, \varepsilon_{1K}$ = the unobserved random effects error terms (i.e., that vary across k stores)

[1] Willett and Singer (1993) recommend using a dichotomous variable representing each time period in discrete time survival analyses. Their recommended approach has the benefit of describing the survival function more precisely because a hazard is calculated for each time period under study. This poses problems in the present research, however, because interactions with each of the 35 time periods would require adding several dozen new variables to the HnLMs (over 200 in the case of the analyses of racial similarity due to the use of dummy variables to represent the racial groups). This, combined with the shape of the survival curve, which closely resembled a quadratic function, led to the decision to use a linear and quadratic factor in the multilevel analyses. Other researchers have also used this approach (e.g., Barber et al., 2000; Horney, Osgood, and Marshall, 1995). Thus, parsimony dictated the use of the Period and Period^2 variables as opposed to the indicators.

Substituting terms to express these equations as a one-level model yields the following:

$$Logit(p_{ijk}) = (\gamma_{00} + \varepsilon_{0K}) + (\gamma_{10} Period_{ij}) + (\gamma_{20} Period_{ij}^2)$$

$$+ (\gamma_{30} + \gamma_{31} StoreSex_{tk} + \varepsilon_{3K}) \quad \text{(EQ8)}$$

This single level equation can now be reformulated as a two-level HnLM where $StoreSex_{tk}$ is treated as an individual level covariate that does not vary across stores:

$$Logit(p_{ijk}) = \beta_{ok} + \beta_1 Period_{ij} + \beta_2 Period_{ij}^2 + \beta_{3k} Sex_j + \beta_4 StoreSex_{tj}$$

$$+ \beta_5 StoreSex_t Sex_j \quad \text{(EQ9)}$$

$$\beta_{0k} = \gamma_{00} + \varepsilon_{0K} \quad \text{(EQ10)}$$

$$\beta_1 = \gamma_{10} \quad \text{(EQ11)}$$

$$\beta_2 = \gamma_{20} \quad \text{(EQ12)}$$

$$\beta_{3k} = \gamma_{30} + \varepsilon_{3K} \quad \text{(EQ13)}$$

$$\beta_4 = \gamma_{40} \quad \text{(EQ14)}$$

$$\beta_5 = \gamma_{50} \quad \text{(EQ15)}$$

Thus, as can be seen above, the main effect and corresponding cross-level interaction between the time-varying $StoreSex_{tk}$ covariate and the employees' sex_j is now at the individual level, without a corresponding random error component in EQ14 or EQ15. This allows for proper estimation of this effect. The results of these models are described below.

Effects of Store Sex Composition on the Relationship Between Sex and Turnover

Equations 9-15 above represent the initial model fitted to the data.[2] Table 24 summarizes the results of this analysis. The first column provides the variable name (and indicates the coding in the case of the dummy and composition variables), followed by the gamma designation and associated random error components (when applicable), and then the parameter estimate, standard error, the t coefficient and Exp (B). All the substantive parameters of interest were significant predictors of survival probabilities. A predicted turnover hazard was calculated for each time period for males and females working in two different store sex compositions using the formula given in EQ2 which was expanded to include all the coefficients included in this analysis. The results of these calculations reveals the form of the sex X store sex composition interaction (see Figure 20). As can be seen in the figure, shifting to a store with a higher female composition slightly reduces the likelihood of turnover for a female employee. An opposite, though somewhat larger effect is seen for males. Thus, hypothesis 20 was supported because increased sex similarity reduces the hazard of turnover. Note that while the differences between the curves on the chart are relatively small, the plotted hazards are not cumulative, but rather are calculated separately for each time period. Cumulative survival probabilities would reveal larger differences than those depicted in Figure 20.

The curves in Figure 20 are parallel because this model does not include any interactions with time. Thus, additional analyses were conducted to see if the sex X store

[2] Models with time and time^2 as random effects were also estimated, which allows the particular shape of the survival function to vary across stores. The results obtained in these analyses were very similar to those reported here (where these variables were modeled as fixed). However, due to ill-conditioning, in some of the analyses several thousand restaurants were dropped. While this effect could likely be mitigated by centering the level 1 predictors, the intercept in the models for race and sex is inherently meaningful where the level 1 predictors are in their raw score form, and so it was decided that the two time parameters should be modeled as fixed to make better use of the entire dataset.

sex composition interaction might vary as a (linear) function of time. This was done because some of the relational demography literature reviewed in the introduction suggests that these effects occur earlier on in groups (Hypothesis 23). Because the groups studied here were very dynamic, these effects were modeled as a function of individual tenure rather than by looking at the start or end of a group. This analysis thus added three cross-level effects to the model: Sex X period, store sex composition X period, and sex X period X store sex composition. Table 25 summarizes the results of these analyses. As can be seen in the table, all three interactions were significant. These results are also presented graphically in Figure 21. As can be seen in the figure, the curves are no longer parallel, resulting from the significant interactions with time. Clearly, there is support for the notion that the sex X store sex composition interaction is initially important but that this effect fades over time. This is evident after examining the change in the vertical displacement of the two curves for males over time. At the end of the first two week time period (P1), a male working in a 90% female store has a predicted hazard rate of 16%, as compared to a predicted hazard of 11% working in a 10% female store. By P20, however, the curves cross and the form of the relationship between sex and store sex composition is reversed. A similar effect is seen for females, except that the curves cross much earlier, indicating that sex similarity effects become a less important predictor of turnover hazard more quickly as compared to males. This supports Hypothesis 23.

Effects of Store Race Composition on the Relationship Between Race and Turnover

The model for race was the same as the one for sex except that an additional individual level dummy variable was necessary because three race categories were

examined for these analyses: Blacks, Hispanics, and Whites. Similarly, another store composition variable was necessary to represent the additional proportion which served as another indicator of store racial composition. As with sex, only the two indicator variables for individual employees' race were allowed to vary across stores. The results of these analyses are presented in Table 26. As can be seen in the table, all of the coefficients are significant. These results, used in conjunction with an appropriately expanded version of the formula presented in EQ2, are represented graphically in Figure 22. (Note that because some of the curves overlap, the underlying data are presented in Table 27). The figure shows predicted turnover hazards at various points in time as a function of individual race within three different store racial compositions. As can be seen in the figure and the accompanying table, the turnover hazard for Whites is highest in a store composition that is 10% White, 10% Hispanic, and 80% Black. In contrast, it is lowest when the composition is 80% White, 10% Hispanic, and 10% Black. A similar pattern is seen for Hispanics and Blacks in that their predicted turnover hazard is lowest when their own race is most highly represented in the store composition. Thus, H14 was supported.

As with sex, additional analyses were conducted to see if these effects vary as a function of time. The results of these analyses are presented in Table 28. As can be seen in the table, several interactions with time were statistically significant. These results (with the non-significant coefficients set to zero in the calculation of predicted turnover hazard), are presented graphically in Figure 23 (accompanying data are in Table 29). For Whites, there is some evidence to support the hypothesized effect in that the interaction between time, the individual level White dummy variable and the two store composition

variables were significant. Indeed, the figure shows that the difference in predicted turnover hazard between the 80% White, 10% Black, 10% Hispanic and the other two store compositions for Whites becomes smaller as time goes on. A similar effect is seen for Hispanics, though this is only seen in the difference between the 80% Hispanic, 10% Black and 10% White line and the 80% White, 10% Black and 10% Hispanic lines. The curves for Blacks are similar in that the protective effects of a predominantly Black store composition on turnover hazard actually reverses in pattern with one of the other compositions at P16. Thus, these data provide some support for the notion that the impact of racial similarity on turnover risk is higher when individuals are first hired as compared to later on in their tenure. This supports Hypothesis 23.

Effects of Store Age Composition on the Relationship Between Age and Turnover

The HnLM for age was the same as sex except that the former was substituted for the latter at both the individual level and the restaurant level. Age was continuously scaled to years by dividing the age in days by 365. The results of this analysis are summarized in Table 30. As can be seen in the table, all three interaction effects were significant. Predicted turnover hazards were calculated at approximately the following values: Average age (19 years old), 10th percentile (16 years old) and 90th percentile (31 years old). These predicted values are shown graphically in Figure 24, with the underlying data presented in Table 31. As can be seen in the figure, at P1 a 31-year-old employee has a predicted turnover hazard of roughly 15% in a store with an average age of 16, whereas the predicted hazard for a 16-year-old employee in a store with the same average age is slightly less than 12%. Conversely, in a store whose employees average 31-years old, the

pattern is reversed. This supports Hypothesis 17 because age similarity is related to a reduced turnover hazard.

As with the previous two demographic variables supplemental analyses were conducted where additional linear interaction terms with time were added. The coefficients for this analysis are presented in Table 32; predicted values across time are found in Table 33, which are depicted graphically in Figure 25. The data represented in the figure show that the curve with the steepest slope is that of a 31-year old in a store with an average age of 16 years old. Similarly, the differences in the predicted hazard values for a store with an average age of 31 are largest earlier on. This supports the notion that these demographic effects are stronger earlier on in employees' tenure. This analysis, along with those presented earlier, support the assertion of Hypothesis 23 in that demographic similarity had a stronger effect on turnover hazard at lower tenure levels.

The Effect of Extraversion on the Demographic Similarity-Turnover Relationship

Hypothesis 24 stated that extraverts should be less susceptible to demographic similarity effects because this trait should facilitate making judgments based on deeper aspects of similarity rather than demographic similarity. For these analyses, the N was 48,440 (censored N = 10,211). These analyses were conducted using the HnLMs described earlier except that extraversion was added at the individual level as a random effect, while the following interaction effects were also added at the individual level as fixed to accommodate their time-varying components: Extraversion X demographics; extraversion X demographic composition; and extraversion X demographics X demographic composition. The significance of this last interaction term is the focus of this hypothesis. Like the results described earlier, separate analyses were conducted for

each of the three types of demographic similarity studied here to keep the number of variables in each analysis manageable. These results are summarized in Tables 34, 35, and 36 for age, race, and sex, respectively. As can be seen at the bottom of Table 34 the age X age composition X extraversion interaction was not significant. Thus, this hypothesis was not supported for age. On the other hand, Table 35 indicates that Hispanic X race composition X extraversion interactions were statistically significant. A plot of the predicted hazards associated with two different levels of extraversion (3 and 5), two different races (Black and Hispanic), and two different restaurant racial compositions is presented in Figure 26, with the accompanying data presented in Table 37. As can be seen in the figure and the accompanying table, higher levels of extraversion were associated with an increased turnover hazard when in a restaurant racial composition inconsistent with one's own race. Thus, this hypothesis was not supported for race. As can be seen in Table 36, the sex X sex composition X extraversion interaction was not significant. Thus, hypothesis 24 was not supported for any of the demographic variables.

DISCUSSION

The research reported here investigated several key relationships: Between mean levels of personality and ability test scores and demographic diversity and financial performance; between personality-based fit and turnover; and between demographic-based fit and turnover. These questions were addressed using a very large sample of applicant test scores and demographics from over 2,000 quick service restaurants across the country. The results, summarized in Table 38 and more theoretically in Figure 27, yielded evidence of a positive association between ability and team financial

performance. In contrast, the personality variables were not related to team financial performance in the predicted direction, and neither the personality variables nor ability had any meaningful relationship with turnover. The results concerning demographic fit and diversity generally supported the hypotheses in that evidence of relationships to team financial performance, team turnover rates, and individual turnover were found. The following sections describe these results in more detail, and discuss key practical, methodological, and theoretical implications of each.

Team Ability and Personality Composition and Team Performance

The hypothesized relationship between mean levels of team personality and ability is depicted in Figure 2. In short, mean levels of ability and conscientiousness were hypothesized to be positively related to team performance, whereas extraversion was hypothesized to have an inverted-U relationship to performance. Team performance was operationalized as a change in controllable profit for the restaurant in a given month. The only support for any of these hypotheses was with regard to ability, which was positively though weakly related to this change in profitability. Given the vast literature supporting the notion that ability is related to individual level job performance, this result was fully expected. What was surprising, however, was that this relationship was relatively weak, only accounting for a change of perhaps several thousand dollars per restaurant over the course of a year.

There are several possible explanations for this result. First, because of the large amount of missing data (i.e., the overall low proportion of employees who had test scores), the average level of ability might have been an unreliable indicator of the true level of ability within a store. The data were not consistent with this explanation,

however, because test saturation did not moderate the ability – performance relationship. If this explanation was in fact true, one would expect to find a stronger relationship in stores with higher test saturations. Second, the ability measure was quite short, which led to a somewhat less reliable indicator of ability than might be desired. This explanation seems unlikely as well, however, because aggregating the ability scores of people within a store substantially improves the reliability of the measure. Another explanation seems more likely. In particular, these jobs may have been too simple to observe much of a benefit associated with a higher mean level of ability. Given the strongly supported notion that the ability-performance relationship declines as jobs become less complex, this explanation seems very plausible. The low complexity of these jobs is exemplified by the substantial amount of resources that are directed towards making fast-food jobs as simple as possible. Thus, future research should examine this issue in more complex jobs. In addition, it would be helpful to conduct studies in which measures of individual job performance are available to see whether this relationship is at least in part mediated by average levels of individual performance. This would address the issue of whether having a high ability team merely enhances team performance via individual inputs, or whether there is an additional compositional effect that accentuates the translation of individual into team performance. Studies such as this would be especially helpful because it is important to demonstrate the impact of personnel selection at higher levels of analysis. With regard to the specific results obtained here, the fact that the standardized incremental relationship between ability and financial performance was only .03 suggests a relatively weak overall relationship between individual job performance and restaurant performance. If this is in fact the case one has to seriously question the usefulness of

ability in contexts such as this, especially given the substantial adverse impact that will invariably result.

Given the low level of job complexity studied here, motivational variables would be expected to be more important than ability variables. Surprisingly, however, mean levels of conscientiousness, the motivational variable that has received the most attention in selection research over the last 10 years, were not related to the change in controllable profit in the month-long period under study. The most obvious potential explanation for this null result is that individuals' conscientiousness scores were rendered somewhat meaningless by faking. This is consistent with the highly elevated mean and relative lack of variance in these scores. The scores were distorted both in comparison to the student sample that had little motivation to fake, as well as in comparison to many incumbent samples using similar items developed by the same consulting firm that developed the test used here. Given that the conscientiousness scale exhibited strong construct validity based on the correlations with the IPI items, one is left either with this conclusion or the alternative that conscientiousness simply doesn't matter at the team level. This latter explanation seems difficult to justify, as it is easy to think of numerous ways in which careless employees can reduce the profitability of a restaurant. Of course this explanation is in conflict with predictive designs that report significant associations between conscientiousness and job performance, but few other explanations seem plausible. On the other hand, it is important to note that the few studies cited in the introduction that have found mean levels of conscientiousness relate to team performance all used concurrent rather than predictive designs. Last, as with ability, it is also important to note that because the test saturation X conscientiousness interaction still did not yield a

relationship to controllable profit, it is difficult to ascribe these null results to missing conscientiousness data at the individual level.

The same results were found for agreeableness and extraversion. These results were contrary to the hypotheses, which were formulated based on empirical studies on teams, and on meta analyses of research examining the role of these variables in customer service or team-based work systems. Although there is certainly less evidence supporting the validity of these constructs for the prediction of job performance at the individual level as compared to conscientiousness, the existing research and theory strongly supports the notion that they should relate to the outcomes studied here. The discussion above concerning conscientiousness seems applicable here as well, in that the scores on these scales showed strong evidence of faking vis-à-vis the means and standard deviations reported in Table 7. Additional support for this notion can be found in that the intercorrelations at both the restaurant and individual applicant level suggest a single personality factor, whereas the individual level student sample correlations suggest distinct personality factors. These results provide especially strong support for the notion that at least some faking is going on because the student sample is actually much more likely to be representative of the applicant sample in this job as compared to other jobs. Similarly, it should be noted that the meta analyses cited in the introduction do not clearly indicate whether the included studies are predictive or concurrent, so it is difficult to assess whether the results relied upon would be expected to transfer to applicant settings. The likelihood of and potential implications associated with faking have received extensive attention elsewhere, and so they will not be reiterated here; however, the results

reported with the very large sample studied here clearly imply that this is indeed an important problem.

Relationships Between Ability and Personality and Turnover

Figure 3, link I, shows that in general it was expected that there would be a relationship between personality and ability and turnover at the individual level. There was no support found for any of these hypotheses in that test scores were unrelated to turnover probability. This is despite that fact that time to turnover was studied using survival analysis rather than treating it is a dichotomous variable, and that this was a very high turnover job, presumably allowing for ample variability in this measure. Similarly, a distinction was made between voluntary and involuntary turnover, and also between avoidable and unavoidable turnover. It was thought that examining turnover at a granular level would increase the likelihood that significant predictors would be identified, though again, there were still null results.

Based on the person-job fit literature, it was thought that ability would be positively related to turnover probabilities. This is because the jobs studied here are extremely simple, suggesting that high ability employees would find a mismatch between their own level of ability and that required on the job. The level of the ability items suggest that one possible explanation for a lack of a relationship is that these items might exhibit ceiling effects. Indeed, the consulting firm that created the items refers to them as "fundamental skills" items because they are aimed at assessing a very low baseline of cognitive skills. The test score distribution, however, provided very little evidence of ceiling effects. This suggests that either there is a true null effect in these data or that perhaps the ability items are seriously lacking in construct validity. Given the large

sample size, even a somewhat construct-deficient measure should have likely yielded at least some detectable relationship to turnover if one in fact existed. Thus, it seems more likely that the ability-turnover relationship is extremely small or nonexistent in these data. Perhaps these jobs were so simple that it is plainly obvious to any high ability employees that they do not fit in with the cognitive demands of the job, but that these employees are taking the position anyway due to financial hardship that does not allow them to turnover quickly.

Similarly, there was also no evidence that conscientiousness was related to any type of turnover. This was true both at the individual level, and in terms of conscientiousness diversity at the group level. As mentioned in the introduction, there is relatively little research addressing this issue. It seems as if holding onto even menial jobs would be associated with key aspects of conscientiousness as it is currently understood. Again, given the very large sample size, appropriate analytic techniques, and construct validity evidence on based on the research conducted in the student sample, it seems either that the null hypothesis is true or that responses on the conscientiousness scale were so distorted that it is impossible to find any meaningful correlates of these scores.

These results also make it difficult to draw any conclusions about the usefulness of the involuntary and (avoidable and unavoidable) voluntary turnover. This is because null relationships were found between the ability and the personality variables and all types of turnover. While these distinctions are seen as important to some researchers, others have criticized them for a number of reasons. Indeed, it seems as if one may be able to generate plausible hypotheses that apply to all types of turnover, rendering useless the distinctions that were attempted here. For instance, it is easy to see how a motivational

variable might lead one to turnover voluntarily due to dissatisfaction on the job, but also to be terminated (i.e., to turnover involuntarily) because of a lack of motivation and thus poor job performance.

Effects of Demographic Diversity on Team Performance and Team Turnover Rates

The research reported here tested several hypotheses concerning the effects of team demographic diversity on team performance and turnover rates. The hypotheses concerning turnover were supported for all three demographic variables in that higher diversity was also associated with higher turnover rates. This is consistent with the relational demography literature discussed earlier in the paper. Interestingly, no LGC analyses of the trend in turnover rates were possible due to the lack of identifiable dimensionality underlying the turnover rates over time. This lack of stability over time calls into question the nature of turnover as a construct and is broadly consistent with the large body of literature that suggests its etiology is complex and multifaceted. Thus, while having obvious practical importance, this finding highlights the challenges researchers may have in more theoretically-oriented studies of turnover rates.

The LGC analyses found that racial diversity was negatively associated with changes in the trend in controllable profit over time. The fact that the racial diversity intercept was positively associated with the controllable profit intercept suggests that the former finding cannot be easily discounted by the explanation that some other third variable leads to poor profitability. In fact, the LGC analyses control for the influence of the controllable profit intercept on the relationship between racial diversity intercept and the controllable profit slope. This means that it is also difficult to dismiss these findings with

the argument that racially diverse restaurants show declines in profitability simply because (for some unknown reason) they start off so high to begin with.

These findings have a number of important implications. First, they are consistent with relational demography theory, which states that increased demographic diversity should be associated with negative outcomes at work. This clearly provides some strong evidence for the relevance of relational demography theory to organizational science, moving beyond studies that examine "softer" criteria such as communication or attachment to an organization. It is important to note, however, that these results were not obtained for sex or age diversity. Due to the sample size and nature of the measures used here, these null results cannot be easily ascribed to a lack or power or traditional conceptions of measurement error. On the other hand, it is likely that there is some noise in these indices of diversity because all the employees' demographics were equally weighted, regardless of whether the employee worked only a few hours a week as compared to 30, for instance. Taking these results at face value, however, suggests two important implications. First, it suggests that relational demography theory may be too broad and that perhaps certain types of demographic diversity may be salient in different situations. Second, it suggests that racial diversity is associated with something that is much more powerful and organizationally relevant as compared to the two other types of diversity studied here. Though it is easy to ascribe this to the overall relevance of race to everyday life, future research should seek to identify why this might be the case.

This is the first study of which I am aware that has successfully linked racial diversity to organizational profitability. If this finding can be replicated in other settings it presents yet another quandary for organizations that seek to become more racially

diverse. This is underscored by the fact that the measures of racial diversity, like the other measures in this study, underestimate the true effect sizes due to missing information about how many hours each employee worked at a given restaurant, and more importantly, which employees actually worked together. Thus, while to some it may seem politically or ethically correct to seek racially diverse workforces, results such as these suggest that this is likely to negatively impact an organization's bottom line. Organizational decision makers are thus faced with a choice between two desirable goals that are mutually exclusive to some extent. It is important to emphasize, however, that these results speak to the impact of racial diversity per se, rather than representation of a particular racial group.

Effects of Demographic Similarity on Turnover

The relational demography literature reviewed in the introduction suggested that demographic similarity would be associated with a reduced likelihood of turnover. There was strong support for this set of hypotheses for all three demographic variables studied here – age, race, and sex. In particular, being similar to the restaurant's composition on each of these variables was associated with a lower likelihood of turnover. Like the results discussed earlier, this supports the notion that demographic similarity is associated with important outcomes at work. In contrast to the results obtained for turnover rates at the restaurant level of analysis, these results were obtained using a cross-level approach. These results taken together support the notion that relational demography and fit research more broadly can be construed either as a cross-level or higher-level diversity effect, depending on the outcome of interest. The multilevel approach to assessing these issues as cross-level effects represents a useful yet somewhat novel approach to this

issue, though other researchers have recently explored approaches that bear some similarities to the one used here (Davison, Kwak, Seok Seo, & Choi, 2002; Reise, 2000). It should be noted that other fit studies may not necessarily find the same results using both cross-level and diversity approaches with other variables; however, a careful consideration of theory, research designs, and levels of analysis issues should assist researchers in predicting when these two approaches would be expected to yield different results.

The results also indicated that demographic similarity is related to higher turnover hazards earlier in an employee's tenure, and that this effect fades, and in some cases, reverses direction over time. This is support for a potentially key qualifier of relational demography theory – that demographic similarity is perhaps most important at the initial stages on the job. The theoretical implication of this finding is consistent with the literature from social psychology discussed earlier suggesting that people make salient categorizations based on the information available to them at the time. Practically, this suggests that during the first few weeks or months on the job employees might be considered to be at the highest risk for demographic-similarity related turnover and that interventions might be directed towards this time frame.

In contrast, the hypotheses concerning the extraversion X demographic similarity interaction on turnover were not supported. Based on several studies reviewed in the introduction, it was expected that extraverts would be at a lower risk for turnover as a function of demographic similarity. These null results are somewhat surprising in light of those concerning time in that if the mechanism underlying the influence of time is getting to know one's coworkers, being high on extraversion should accelerate this process.

Thus, either this may not be the underlying mechanism, or again, perhaps some faking occurred that distorted the meaning of applicants' extraversion scores. Future research should be directed at clarifying this issue.

Methodological Implications for the Dynamic Studies of Teams

As was mentioned in the introduction, a number of authors have called for research that takes into account the dynamic nature of teams. Team research poses serious challenges as compared to individual level research because the composition of teams changes over time. This was especially true given the high rate of turnover in the jobs studied here. This presented serious challenges because there are no obvious approaches for handling situations such as these. Two of the analytic techniques used here are uniquely well suited toward handling this problem, though to my knowledge neither has been applied in published teams research in this way. First, the LGC methodology allowed the changes in demographic diversity to be easily captured and then related to changes in controllable profit over time. This approach was especially valuable because the zero-order correlations between the monthly racial diversity index and controllable profit were actually positive. However, when analyzed using the LGC approach a much different pattern emerged, underscoring the complexity involved in studying the correlates of change over time. Similarly, because store-level outcomes such as profitability are likely subject to a range of forces beyond the immediate control of the store employees, studying changes over time undercuts most arguments that the obtained results are due to a third, uncontrolled variable that varies widely across stores. Thus, the usefulness of this approach in studies of dynamic team composition is clear, and

researchers should strongly consider this approach when studying dynamic team processes.

Second, the multilevel survival analyses captured the dynamic nature of team demographic composition by taking advantage of the fact that survival analysis can easily handle time-varying covariates. The multilevel implementation of this approach is somewhat more complex than traditional multilevel analyses but can be accommodated in standard multilevel software such as MLwin, HLM, and SAS proc NLMIXED. Researchers studying the effects of dynamic team compositions on dichotomous outcomes should strongly consider the benefits of this approach.

Although this study applied new approaches to novel problems in I/O psychology, there also remains a key statistical issue that would benefit from additional research. This issue pertains to the unknown impact of the restriction of range that occurred in the individual level personality scores. There appears to be no published research on how this is expected to impact the magnitude of the correlations observed in designs like the one used here. Generating some data indicated that the effects are likely to be very complex because of the wide variability in restaurant size and selection ratios used in each restaurant. Thus, this issue is ripe for simulation or analytic studies. Specifically, it would be helpful to know how this range restriction impacts studies such as this as a function of the number level 2 units as well as the number of cases in each one, the strength of the true correlation between the aggregated variable and the other level 2 variable, and the magnitude and variability of the range restriction. This would be especially insightful because studies with these designs are likely to become increasingly

common in personnel selection, where restriction of range is often an issue in both predictive and concurrent designs.

Theoretical and Methodological Implications for Relational Demography and PO Fit

One of the major themes throughout this paper was the emphasis on looking at the effects of team composition on both individual and team-level outcomes. Although similar results were obtained for the cross- and team-level analyses, the latter were only evident when turnover rates were aggregated over a considerable time period. This highlights the complexities involved in conducting multi-level longitudinal research, in that different results can be obtained depending on the time scale and level of analysis. This suggests that researchers should thoroughly evaluate the soundness of their choices along these lines.

The results reported here also have implications for the way in which similarity and PO fit are related. In particular, this research showed how similarity can be conceptualized as fit between an individual and a composition in which that individual is embedded. HnLMs were then used to see if the consistency between that composition and the individuals' attributes predicted an important outcome of interest. Taking similar approaches would again help to clarify the confusion surrounding levels of analysis in the relational demography literature, and also would provide researchers an alternative to the complexities involved in polynomial regression analysis advocated by Edwards (e.g., 2002), often the approach of choice for PO fit studies. In addition, this research demonstrated how fit can also be construed as a higher-level diversity effect. This approach should be pursued in other studies in that enables PO fit and other types of fit

researchers to broaden the nomological network of variables surrounding the notion of fit.

Conclusion

The research reported here found evidence that diversity is negatively related to organizational and individual level outcomes. Because diversity is mandatory rather than optional in most organizations, future research should investigate what can be done to ameliorate these effects. First, research that measures the variables that are thought to connect diversity with these outcomes would be helpful. Second, examining the efficacy of programs or interventions targeting these intervening variables would be helpful. Because the relationships were found to operate at multiple levels of analysis researchers and practitioners should consider leverage points at both levels as well. Not only will this further our knowledge about the particular phenomena studied here, but it will also enhance our ability to conceptualize and operationalize multilevel studies of complex human behavior at work.

REFERENCES

Ackerman, P.L. & Humphreys, L.G. (1990). Individual differences in industrial and organizational psychology. In M.D. Dunnette & L.M. Hough (Eds.), Handbook of industrial and organizational psychology, 2nd edition (Vol 1, pp. 223-282).

Agresti, A. (1996). An introduction to categorical data analysis. New York: Wiley.

Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. In S. Leinhardt (Ed.). Sociological Methodology (pp. 61-98). Jossey Bass: San Francisco.

Allison, P. D. (1984). Event history analysis: Regression for longitudinal event data. Beverly Hills, CA: Sage.

Ancona, D. & Caldwell, D. (1992). Demography and design: Predictors of new product team performance. Organization Science, 3, 321-341.

Ancona, D. G., & Caldwell, D. F. (1998). Rethinking team composition from the outside in, Gruenfeld, Deborah H (Ed). (1998). Composition. Research on managing groups and teams, Vol. 1. (pp. 21-37). Stamford, CT, US: JAI Press, Inc.

Antonioni, D., & Park, H. (2001). The effects of personality similarity on peer ratings of contextual work behaviors. Personnel Psychology, 54(2), 331-360.

Arrow, H. (1998). Standing out and fitting in: Composition effects on newcomer socialization. In Gruenfeld, D. H. (Ed.), Research on Managing Groups and Teams: Composition (pp. 59-80). Stamford, CT: JAI Press.

Arthur, W., Tubre, T., & Hanson, A. (2000). g-based optimal pairing strategies in a team training protocol. Paper presented at the Annual Society for Industrial/Organizational Psychology, New Orleans, LA.

Arvey, R. D. & Faley, R. H. (1984). Fairness in Selecting Employees. Reading Mass: Addison-Wesley.

Ashkanasy, N. M. & O'Connor, C. (1997). Value congruence in leader-member exchange. Journal of Social Psychology, 137, 647-662.

Austin, J. T., & Villanova, P. (1992). The criterion problem: 1917-1992. Journal of Applied Psychology, 77(6), 836-874.

Barber, J. S., Murphy, S. A., Axinn, W. A., & Maples, J. (2001). Discrete-time multilevel hazard analysis. Sociological Methodology, 31, 201- 235.

Baron, R. A. (1996). Interpersonal relations in organizations. In K. R. Murphy (Ed.). Individual Differences and Behavior in Organizations. (pp. 334-370). San Francisco: Jossey-Bass.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. Journal of Applied Psychology, 81, 261-272.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. Journal of Applied Psychology, 81(3), 261-272.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next. International Journal of Selection and Assessment, 9(1-2), 9-30

Barrick, M. R., Stewart, G. L., & Piotrowski, M. (2002). Personality and job performance: Test of the mediating effects of motivation among sales representatives. Journal of Applied Psychology, Applied-Psychology.

Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. Journal of Applied Psychology, 83, 377-391.

Barrick, M.R. & Mount, M.K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. Personnel Psychology, 44, 1-26.

Barry, B., & Stewart, G. L. (1997). Composition, process, and performance in self-managed groups: The role of personality. Journal of Applied Psychology, 82(1), 62-78.

Barsade, S. G., Ward, A. J., Turner, J. D. F., & Sonnenfeld, J. A. (2000). To your heart's content: A model of affective diversity in top management teams. Administrative Science Quarterly, 45(4), 802-836.

Baugh, S. G., & Graen, G. B. (1997). Effects of team gender and racial composition on perceptions of team performance in cross-functional teams. Group and Organization Management, 22(3), 366-383.

Bedeian, A. G., Kemery, E. R., & Mossholder, K. W. (1989). Testing for cross-level interactions: An empirical demonstration. Behavioral Science, 34(1), 70-78.

Bentler, P. M. & Bonnett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88, 588-606.

Berscheid, E. & Walster, H. (1978). Interpersonal Attraction. Reading, MA: Addison-Wesley.

Bettenhausen, K. L. (1991). Five years of groups research: What we have learned and what needs to be addressed. Journal of Management, 17, 345-381

Blau, P. (1977). Inequality and Heterogeneity. New York: Free Press.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis, K. J. Klein & S. W. J Kozlowski (Eds).

(2000). Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions. (pp. 349-381). San Francisco, CA, US: Jossey Bass Inc, Publishers.

Bliese, P. D. (2002). Multilevel random coefficient modeling in organizational research: Examples using SAS and S-PLUS. In F. Drasgow & N. Schmitt (Eds.), Measuring and Analyzing Behavior in Organizations: Advances in Measurement and Data Analysis. San Fransisco: Jossey-Bass.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. Psychological Bulletin, 117, 187-215

Bond, M. H., & Shiu, W. Y. F. (1997). The relationship between a group's personality resources and the two dimensions of its group process. Small Group Research, 28(2), 194-217.

Borman, W. C., & Motowidlo, S. J. (1997). Task performance and contextual performance: The meaning for personnel selection research. Human Performance, 10(2), 99-109.

Borman, W.C. & Motowidlo, S.J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt, W.C. Borman & Associates (Eds.), Personnel selection in organizations (pp. 71-98). San Francisco: Jossey-Bass Publishers.

Bowers, C. A., Jentsch, F., Salas, E., & Braun, C. C. (1998). Analyzing communication sequences for team training needs assessment. Human Factors, Human-Factors.

Bretz, R. D., & Judge, T. A. (1994). Person-organization fit and the Theory of Work Adjustment: Implications for satisfaction, tenure, and career success. Journal of Vocational Behavior, 44(1), 32-54

Brockner, J., & et al. (1986). Layoffs, equity theory, and work performance: Further evidence of the impact of survivor guilt. Academy of Management Journal, 29(2), 373-384.

Brockner, J., Grover, S. L., & Blonder, M. D. (1988). Predictors of survivors' job involvement following layoffs: A field study. Journal of Applied Psychology, 73(3), 436-442.

Brockner, J., Grover, S., Reed, T., DeWitt, R., & et al. (1987). Survivors' reactions to layoffs: We get by with a little help for our friends. Administrative Science Quarterly, 32(4), 526-541.

Brockner, J., Konovsky, M., Cooper Schneider, R., Folger, R., & et al. (1994). Interactive effects of procedural justice and outcome negativity on victims and survivors of job loss. Academy of Management Journal, 37(2), 397-409.

Bryk, A. S., & Raudenbush, S. W. (1992). Hierarchical linear models: Applications and data analysis methods. Thousand Oaks, CA: Sage.

Bryk, A., Raudenbush, S., & Congdon, R. (1996). Hierarchical linear and nonlinear modeling with the HLM/2L and HLM/3L Programs. Chicago: SSI.

Bryne, D. E. (1971). The Attraction Paradigm. New York: Academic Press.

Cameron, K. S. (1998). Strategic downsizing: An extreme case. In L. L. Cummings & B. M. Staw (Eds.), Research in Organizational Behavior. Volume 20 (pp. 185-229). Greenwich, CT: JAI Press.

Campbell, J. P., Blake Glasser, M., & Oswald, F. L. (1996). The substantive nature of job performance variability. In K. R. Murphy (Ed.). Individual Differences and Behavior in Organizations. (pp. 258-299). San Francisco: Jossey-Bass.

Campbell, J.P., McCloy, R.A., Oppler, S.H. & Sager, C.E. (1993). A theory of performance. In N. Schmitt, W.C. Borman & Associates (Eds.), Personnel selection in organizations (pp. 35-70). San Francisco: Jossey-Bass Publishers.

Campion, M. A., Medsker, G. J., & Higgs, A. C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. Personnel Psychology, 46(4), 823-850.

Canon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1992). Defining competencies and establishing team training requirements. In R. R. Guzzo & E. Salas and Associates (Eds.), Team Effectiveness and Decision Making in Organizations. (pp. 333-380). San Fransisco: Jossey-Bass.

Carroll, J.B. (1993). Human cognitive abilities: A survey of factor-analytic studies. New York: Cambridge University Press.

Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incporating longitudinal means and covariance structures analyis (LMACS) and multiple indicator latent growth modelling (MLGM). Organizational Research Methods, 1, 421-483.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. Journal of Applied Psychology, 83, 234-246.

Chan, D., Ramey, S., Ramey, C., & Schmitt, N. (2000). Modeling intraindividual changes in children's social skills at home and at school: A multivariate latent growth approach to understanding between-settings differences in children's social skills development. Multivariate Behavioral Research, 35, 365-396.

Chatman, J. A. (1991). Matching people and organizations: Selection and socialization in public accounting firms. Administrative Science Quarterly, 36(3), 459-484.

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(5), 956-974.

Chatman, J. A., Polzer, J. T., Barsade, S. G., & Neale, M. A. (1998). Being different yet feeling similar: The influence of demographic composition and organizational culture on work processes and outcomes. Administrative Science Quarterly, 43(4), 749-780.

Chattopadhyay, P. (1999). Beyond direct and symmetrical effects: The influence of demographic dissimilarity on organizational citizenship behavior. Academy of Management Journal, 42, 273-287.

Clevenger, J., Pereira, G. M., Wiechmann, D., Schmitt, N., & Harvey, V. S. (2001). Incremental validity of situational judgment tests. Journal of Applied Psychology, 86(3), 410-417.

Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. Journal of Management, 23, 239-290

Conway, J. M. (1996). Additional construct validity evidence for the task/contextual performance distinction. Human Performance, 9(4), 309-329.

Conway, J. M. (1999). Distinguishing contextual performance from task performance for managerial jobs. Journal of Applied Psychology, 84(1), 3-13.

Costa, P.T, Jr. & McCrae, R.R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. Journal of Personality and Social Psychology, 54, 853-863.

Costa, P.T, Jr. & McCrae, R.R. (1990). Personality in adulthood. New York, NY, US: The Guilford Press

Costa, P.T. & McCrae, R.R. (1995). Domains and facts: Hierarchical personality assessment using the Revised NEO Personality Inventory. Journal of Personality Assessment, 64, 21-50.

Cox, D. R. (1972). Regression models and life tables. Journal of the Royal Statistical Society, 34, 187-220.

Davison, M. L., Kwak, N., Seok Seo, Y., & Choi, J. (2002). Using Hierarchical Linear Models to Examine Moderator Effects: Person-by-Organization Interactions. Organizational Research Methods, 5, 231-254.

Day, D. V., & Bedeian, A. G. (1995). Personality similarity and work-related outcomes among African-American nursing personnel: A test of the supplementary model of person environment congruence. Journal of Vocational Behavior, 46(1), 55-70.

De Corte, W. (1999). Weighing job performance predictors to both maximize the quality of the selected workforce and control the level of adverse impact. Journal of Applied Psychology, 84(5), 695-702.

Devine, D. J. (1999). Effects of cognitive ability, task knowledge, information sharing, and conflict on group decision-making effectiveness. Small Group Research, 30, 608-634

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. Annual Review of Psychology, 41, 417-440.

Digman, J.M. (1989). Five robust trait dimensions: Development, stability and utility. Journal of Personality, 57, 195-214.

Digman, J.M. & Inouye, J. (1986). Further specification of the five robust factors or personality. Journal of Personality and Social Psychology, 50, 116-123.

Dipboye, R. L. (1994). Structured and unstructured selection interviews: Beyond the job-fit model. In G. Ferris (ed.), Research in Personnel and Human Resources Management 12, Greenwich, CT: JAI Press, pp. 79-123.

Driskell, J. E. & Salas, E. (1992). Can you study real teams in contrived settings? The value of small group research to understanding teams. In R. W. Swezey & E. Salas. Teams: Their Training and Performance. (pp. 101-126). USA: Ablex.

Edwards J. (1994). The study of congruence in organizational behavior research: Critique and a proposed alternative. Organizational Behavior and Human Decision Processes, 58,51-100.

Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow & N. Schmitt (Eds.), Measuring and Analyzing Behavior in Organizations: Advances in Measurement and Data Analysis. San Fransisco: Jossey-Bass.

Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. Journal of Applied Psychology, 84, 155-166.

Engle, E. M., & Lord, R. G. (1997). Implicit theories, self-schemas, and leader-member exchange. Academy of Management Journal, 40(4), 988-1010.

Eysenck, H. J. (1947). Dimensions of Personality. New York: Praeger.

Feeley, T. J., & Barnett, G. A. (1997). Predicting employee turnover from communication networks. Human Communication Research, 23, 370-387

Ferris, G. R., Judge, T. A., Chachere, J. G., & Liden, R. C. (1991). The age context of performance-evaluation decisions. Psychology and Aging, 6(4), 616-622.

Ferris, G. R., Youngblood, S. A., & Yates, V. L. (1985). Personality, training performance, and withdrawal: A test of the person-group fit hypothesis for organizational newcomers. Journal of Vocational Behavior, 27(3), 377-388.

Fleishman, E. A. & Zaccaro, S. J. (1992). Toward a taxonomy of team performance functions. In R. W. Swezey & E. Salas. Teams: Their Training and Performance. (pp. 31-56). USA: Ablex.

Flynn, L. P. & Shore, L M. (1994). Gender and context: The moderating effect of work group gender composition and occupation sex-type on work attitudes and behaviors. In M. Schnake (Ed.), Proceedings of the Southern Management Association.

Forsyth, D. R. (1999). Group Dynamics. USA: Wadsworth.

Frei, R. L., & McDaniel, M. A. (1998). Validity of customer service measures in personnel selection: A review of criterion and construct evidence. Human Performance, 11, 1-27

George, J. M. (1990). Personality, affect, and behavior in groups. Journal of Applied Psychology, 75(2), 107-116.

George, J. M. (1996). Trait and state affect. In K. R. Murphy (Ed.). Individual Differences and Behavior in Organizations. (pp. 145-174). San Francisco: Jossey-Bass.

George, J. M., & Bettenhausen, K. (1990). Understanding prosocial behavior, sales performance, and turnover: A group-level analysis in a service context. Journal of Applied Psychology, 75(6), 698-709.

Ghiselli, G. E. (1966). The Validity of Occupational Aptitude Tests. New York: Wiley.

Gladstein, D. (1984). Groups in context: A model of task group effectiveness. Administrative Science Quarterly, 29, 499-517.

Glick, W. H., & Roberts, K. H. (1984). Hypothesized interdependence, assumed independence. Academy of Management Review, 9, 722-735

Goldberg, L.R. (1990). An alternative "description of personality": The Big-Five factor structure. Journal of Personality and Social Psychology, 59, 1216-1229.

Goldberg, L.R. (1993). The structure of phenotypic personality traits. American Psychologist, 48, 26-34.

Gowan, M. A., Riordan, C. M., & Gatewood, R. D. (1999). Test of a model of coping with involuntary job loss following a company closing. Journal of Applied Psychology, 84(1), 75-86.

Green, S. G., Anderson, S. E., & Shivers, S. L. (1996). Demographic and organizational influences on leader-member exchange and related work attitudes. Organizational Behavior and Human Decision Processes, 66(2), 203-214.

Greguras, G. J., & Robie, C. (1998). A new look at within-source interrater reliability of 360-degree feedback ratings. Journal of Applied Psychology, 83(6), 960-968.

Griffeth, R. & Hom, P. (1994). Employee Turnover. USA: South-Western College Publishing.

Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. Journal of Management, 26(3), 463-488.

Guastello, S. J., & Guastello, D. D. (1998). Origins of coordination and team effectiveness: A perspective from game theory and nonlinear dynamics. Journal of Applied Psychology, 83(3), 423-437.

Guion, R.M. & Gottier, R.F. (1965). Validity of personality measure in personnel selection. Personnel Psychology, 18, 135-164.

Gully, S. M., Devine, D. J., & Whitney, D. J. (1995). A meta-analysis of cohesion and performance: Effects of levels of analysis and task interdependence. Small Group Research, 26(4), 497-520.

Guzzo, R.A. & Dickson, M.W. (1996). Teams in organizations: Recent research on performance and effectiveness. Annual Review of Psychology, 46, 307-338.

Guzzo, R.A. & Shea, G.P. (1992). Group performance and intergroup relations in organizations. In M.D. Dunnette and L.M. Hough (Eds.), Handbook of Industrial and Organizational Psychology (Vol 3, pp. 269-313). Palo Alto, CA: Consulting Psychologists Press.

Hackman, J. R. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. In L. Berkowitz (Ed.). Advances in Experimental Social Psychology. New York: Academic Press.

Hackman, J. R. (1987). The design of work teams. In J. W. Lorsch (Ed.), Handbook of Organizational Behavior (pp. 315-342). Englewood Cliffs, NJ: Prentice-Hall.

Hackman, J. R. (1992). Group influences on individuals in organizations. In Dunnette, Marvin D. (Ed); Hough, Leaetta M. (Ed). (1992). Handbook of industrial and organizational psychology, Vol. 3 (2nd ed.). (pp. 199-267).

Hackman, J.R. & Morris, C.G. (1975). Group task, group interaction process, and group performance effectiveness: A review and proposed integration. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology, (Vol 8). New York: Academic Press.

Hannan, M. T. (1971). Aggregation and Disaggregation in Sociology. Lexington, MA: Health.

Hannan, M. T. (1990). Aggregation and Disaggregation in the Social Sciences. Lexington, MA: Lexington Books.

Harrison, D. A. (2002). Structure and timing in limited range dependent variables: Regression models for predicting if and when. In F. Drasgow & N. Schmitt (Eds.). Measuring and Analyzing Behavior in Organizations. Jossey-Bass: San Francisco.

Harrison, D. A. (2002). Structure and timing in limited range dependent variables: Regression models for predicting if and when. In. F. Drasgow & N. Schmitt (Eds.), Measuring and Analyzing Behavior in Organizations: Advances in Measurement and Data Analysis. San Fransisco: Jossey Bass.

Harrison, D. A., Price, K. H., & Bell, M. P. (1998). Beyond relational demography: Time

and the effects of surface- and deep-level diversity on work group cohesion. Academy of Management Journal, 41(1), 96-107.

Hartigan, J. A., & Wigdor, A. K. (1989). Fairness in employment testing: Validity generalization, minority issues, and the General Aptitude Test Battery..

Hattrup, K., O' Connell, M. S., & Wingate, P. H. (1998). Prediction of multidimensional criteria: Distinguishing task and contextual performance. Human Performance, 11, 305-319

Hattrup, K., O' Connell, M. S., & Wingate, P. H. (1998). Prediction of multidimensional criteria: Distinguishing task and contextual performance. Human Performance, 11(4), 305-319.

Hattrup, K., Rock, J., & Scalia, C. (1997). The effects of varying conceptualizations of job performance on adverse impact, minority hiring, and predicted performance. Journal of Applied Psychology, 82, 656-664.

Heck, R. H., & Thomas, S. L. (2000). An introduction to multilevel modeling techniques. Mahwah, NJ, US: Lawrence Erlbaum Associates, Inc., Publishers.

Hedeker, D. & Gibbons, R. D. (1996).. MIXREG: A computer program for mixed-effects regression analysis for autocorrelated errors. Computer Methods and Programs in Biomedicine, 49, 229-252.

Hedeker, D., Siddiqui, O., & Hu, F.B. (2000). Random-effects regression analysis of correlated grouped-time survival data. Statistical Methods in Medical Research, 9:161-179

Herold, D. M .(1978). Improving the performance effectiveness of groups through a task-congingent selection of intervention strategies. Academy of Management Review, 3, 315-325.

Hinds, P. J., Carley, K. M., Krackhardt, D., & Wholey, D. (2000). Choosing work group members: Balancing similarity, competence, and familiarity. Organizational Behavior and Human Decision Processes, 81(2), 226-251.

Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. Journal of Management, 23(6), 723-744.

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. Journal of Management, 24, 623-641

Hogan, J., Rybicki, S. L., Motowidlo, S. J., & Borman, W. C. (1998). Relations between contextual performance, personality, and occupational advancement. Human Performance, 11(2-3), 189-207

Hogan, R., & Hogan, J. (1995). Hogan Personality Inventory Manual. Tulsa, OK: Hogan Assessment Systems.

Hogan, R., Raza, S., & Driskell, J. E. (1988). Personality, team performance, and organizational context. In P. Whitney & Ochsman, R. B. (Eds.), Psychology and Productivity. New York: Plenum.

Holland, J. L. (1985). Making Vocational Choices: A Theory of Vocational Personalities and Work Environments. (2nd Ed.). Englewood Cliffs, NJ: Prentice-Hall.

Horney, J., Osgood, D. W., & Marshall, I. H. (1995). Criminal careers in the short-term: Intra-individual variability crime and its relation to local life circumstances. American Sociological Review, 60, 655-673.

Hough, L. M. & Schneider, R. J. (1996). Personality traits, taxonomies, and applications in organizations. In K. R. Murphy (Ed.). Individual Differences and Behavior in Organizations. (pp. 31-88). San Francisco: Jossey-Bass.

Hough, L. M. (1992). The "Big Five" personality variables--construct confusion: Description versus prediction. Human Performance, 5(1-2), 139-155.

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. International Journal of Selection and Assessment, 9(1-2), 152-194

House, R., Rousseau, D.M., & Thomas-Hunt, M. (1995). The meso paradigm: A framework for the integration of micro and macro organizational behavior. In L.L. Cummings & B.Staw (Eds.), Research in organizational behavior (Vol 17, pp. 71-114). Greenwich, CT: JAI Press.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitude, job knowledge, and job performance. Journal of Vocational Behavior, 29, 340-362

Hunter, J.E. & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-98.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. Journal of Applied Psychology, 85(6), 869-879.

Ilgen, D. R. (1999). Teams embedded in organizations: Some implications. American Psychologist, 54, 129-139

Ilgen, D. R., Major, D. A., Hollenbeck, J. R., & Sego, D. J. (1995). Raising an Individual Decision-Making Model to the Team Level: A New Research Model and Paradigm. In R. A. Guzzo & E. Salas and Associates (Eds.), Team Effectiveness and Decision Making in Organizations. San Fransisco: Jossey-Bass.

Iverson, R. D., & Buttigieg, D. M. (1997). Antecedents of union commitment: The impact of union membership differences in vertical dyads and work group relationships. Human Relations, 50(12), 1485-1510.

Jackson, S. E. (1992). Team composition in organizational settings: Issues in managing an increasingly diverse work force. In S. Worchel, W. Wood, & J. A. Simpson (Eds.), Group Process and Productivity. (pp. 138-177). Newburry Park, CA: Sage.

Jackson, S. E., & Ruderman, M. N. (Eds.). (1995). Diversity in work teams: Research paradigms for a changing workplace. Washington, DC, US: American Psychological Association.

Jackson, S. E., Brett, J. F., Sessa, V. I., Cooper, D. M., & et al. (1991). Some differences make a difference: Individual dissimilarity and group heterogeneity as correlates of recruitment, promotions, and turnover. Journal of Applied Psychology, 76(5), 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 and Associates (Eds.), Team Effectiveness and Decision Making in Organizations. San Francisco: Jossey-Bass.

Jackson, S. E., Stone, V. K., & Alvarez, E. B. (1992). Socialization amidst diversity: The impact of demographics on work team oldtimers and newcomers. In L. L. Cummings & B. M. Staw (Eds.), Research in Organizational Behavior. Volume 15 (pp. 45-110). Greenwich, CT: JAI Press.

Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict, and performance in workgroups. Administrative Science Quarterly, 44(4), 741-763.

Jensen, A.R. (1980). Bias in mental testing. New York: Free Press.

Johns, G. (1981). Difference score measures of organizational behavior variables: A critique. Organizational Behavior and Human Decision Processes, 27, 443-463.

Johnson, J. W. (2001). The relative importance of task and contextual performance dimensions to supervisor judgments of overall performance. Journal of Applied Psychology, 86, 984-996.

Jöreskog, K. & Sörbom, D. (1986). LISREL 6: Analysis of linear structural relationships by maximum likelihood and least square methods. Chicago: Scientific Software International.

Jöreskog, K. & Sörbom, D. (1989). LISREL 7: A Guide to the program and applications. (2nd Edition). Chicago: Scientific Software International.

Jöreskog, K. & Sörbom, D. (1993). LISREL 8 user's reference guide. Chicago: Scientific Software International.

Jung, D. I., & Sosik, J. J. (1999). Effects of group characteristics on work group performance: A longitudinal investigation. Group Dynamics, 3(4), 279-290.

Kennedy, P. (1998). A Guide to Econometrics. USA: MIT Press.

Kenny, D. A., & Judd, C. M. (1986). Consequences of violating the independence

assumption in analysis of variance. Psychological Bulletin, 99(3), 422-431.

Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. American Psychologist, 43, 23-34

Kidwell, R. E., Jr., Mossholder, K. W., & Bennett, N. (1997). Cohesiveness and organizational citizenship behavior: A multilevel analysis using work groups and individuals. Journal of Management, 23(6), 775-793.

Kirchmeyer, C. (1995). Demographic similarity to the work group: A longitudinal study of managers at the early career stage. Journal of Organizational Behavior, 16(1), 67-83.

Klein, K. J. & Kozlowski, S. W. J. (2000). Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions. San Francisco, CA, US: Jossey Bass Inc, Publishers

Klein, K. J., Dansereau, F., & Hall, R. J. (1994). Levels issues in theory development, data collection, and analysis. Academy of Management Review, 19, 195-229.

Klimoski, R. & Jones, R.G. (1995). Staffing for effective group decision making: Key issues in matching people and teams. In R.A. Guzzo, E. Salas, & Associated (Eds.), Team effectiveness and decision making in organizations (pp. 291-332). San Francisco: Jossey-Bass.

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes, K. J. Klein & S. W. J Kozlowski (Eds). (2000). Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions. (pp. 3-90). San Francisco, CA, US: Jossey Bass Inc, Publishers

Kreft, I. & de Leeuw, J. (1998). Introducing Multilevel Modeling. Thousand Oaks, CA: Sage.

Kristof Brown, A. L., & Stevens, C. K. (2001). Goal congruence in project teams: Does the fit between members' personal mastery and performance goals matter? Journal of Applied Psychology, 86(6), 1083-1095.

Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. Personnel Psychology, 49, 1-49

Landy, F. J. & Farr, J. L. (1983). The measurement of work performance. New York: Academic Press.

Langbein, L. I. & Lichtman, A. J. (1978). Ecological Inference. Beverly Hills, CA: Sage.

LaRocco, J. M. (1983). Job attitudes, intentions, and turnover: An analysis of effects using latent variables. Human Relations, 36(9), 813-825.

Laughlin, P. R., & Adamopoulos, J. (1980). Social combination processes and individual

learning for six-person cooperative groups on an intellective task. Journal of Personality and Social Psychology, 38(6), 941-947.

Leana, C. R. & Feldman, D. C. (1994). The psychology of job loss. In G. R. Ferris (Ed.). Research in Personnel and Human Resources Management. Volume 12 (pp. 271-302). New York: JAI Press.

LePine, J. A., & Van Dyne, L. (1998). Predicting voice behavior in work groups. Journal of Applied Psychology, 83, 853-868.

LePine, J. A., & Van Dyne, L. (2001). Voice and cooperative behavior as contrasting forms of contextual performance: Evidence of differential relationships with Big Five personality characteristics and cognitive ability. Journal of Applied Psychology, 86, 326-336.

Lepine, J. A., Hanson, M. A., Borman, W. C., & Motowidlo, S. J. (2000). Contextual performance and teamwork: Implications for staffing. In G. R. Ferris (Ed.). Research in Personnel and Human Resources Management. Volume 19 (pp. 53-90). New York: JAI Press.

LePine, J. A., Hollenbeck, J. R., Ilgen, D. R., & Hedlund, J. (1997). Effects of individual differences on the performance of hierarchical decision-making teams: Much more than g. Journal of Applied Psychology, 82(5), 803-811.

Levine, J. M., & Moreland, R. L. (1990). Progress in small group research. Annual Review of Psychology, 41, 585-634.

Lofquist, L. H. & Dawis, R. (1969). Adjustment To Work: A Psychological View Of Man's Problems In A Work-Oriented Society. New York: Appleton-Century-Crofts.

Markham, S. E. (1988). Pay-for-performance dilemma revisited: Empirical example of the importance of group effects. Journal of Applied Psychology, 73(2), 172-180.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon Bowers, J. A. (2000). The influence of shared mental models on team process and performance. Journal of Applied Psychology, 85(2), 273-283.

McCain, B. E., O' Reilly, C., & Pfeffer, J. (1983). The effects of departmental demography on turnover: The case of a university. Academy of Management Journal, 26(4), 626-641.

McCormick E., DeNisi A., Staw J. (1979). Use of the position analysis questionnaire for establishing the job component validity of tests. Journal of Applied Psychology, 64, 51-56.

McCormick E., Jeanneret P., Mecham R. (1972). A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire (PAQ), (Monograph) Journal ofApplied Psychology, 56,347-368.

McCormick, E. J., DeNisi, A. S., & Shaw, J. B. (1979). Use of the position analysis

questionnaire for establishing the job component validity of tests. Journal of Applied Psychology, Applied-Psychology.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. Journal of Applied Psychology, 79(4), 599-616.

McElroy, J. C., Morrow, P. C., & Rude, S. N. (2001). Turnover and organizational performance: A comparative analysis of the effects of voluntary, involuntary, and reduction in force turnover. Journal of Applied Psychology, 86, 1294-1299.

McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across noncognitive measures. Journal of Applied Psychology, 85, 812-821.

McGrath, J. E. (1984). Groups: Interaction and Performance. Englewood Cliffs, NJ: Prentice-Hall.

McGrath, J. E. (1998). A view of group composition through a group-theoretic lens. In D. E. Gruenfeld (Ed.). Research on Managing Teams and Groups: Composition (pp. 255-272). Greenwich, CT: JAI Press.

McGrath, J. E., Berdahl, & Arrow, H. (1996). Traits, expectation, culture, and clout: The dynamics of diversity in work groups. In S. E. Jackson & M. N. Ruderman (Eds.), Diversity in Work Teams: Research Paradigms for a Changing Workplace (pp. 17-45). Washington DC: American Psychological Association.

McPherson, J. M., Popielarz, P. A., & Drobnic, S. (1992). Social networks and organizational dynamics. American Sociological Review, 57(2), 153-170.

Meglino B, Ravlin E, Adkins C. (1989). A work values approach to corporate culture: A field test of the value congruence process and its relationship to individual outcomes. Journal of Applied Psychology, 74, 424-432.

Meglino B., Ravlin E., Adkins C. (1992). The measurement of work value congruence: A field study comparison. Journal of Management,18, 33--43.

Mellor, S. (1995). Gender composition and gender representation in local unions: Relationships between women's participation in local office and women's participation in local activities. Journal of Applied Psychology, 80(6), 706-720.

Meredith, W. & Tisak, J. (1990). Latent curve analysis. Psychometrika, 55, 107-122.

Milliken, F. J. & Martins, L. L. (1996). Searching for common threads: Understanding the multiple effects of diversity in organizational groups. Academy of Management Review, 21, 402-443.

Mischel, W. (1968). Personality and assessment. New York: Wiley.

Moreland, R. L., Levine, J. M. & Wingert, M. L. (1992). Creating the ideal group: Composition effects at work.

Morgan, B. B. & Lassiter, D. L. (1992). Team composition and staffing. In R. W. Swezey & E. Salas. Teams: Their Training and Performance. (pp. 75-100). USA: Ablex.

Morgan, B.B, Glickman, A.S., Woodward, E.A., Blaiwes, A.S., et al. (1986). Measurement of team behaviors in a Navy training environment. US Naval Training Systems Center Technical Reports.

Morita, J. G., Lee, T. W., & Mowday, R. T. (1989). Introducing survival analysis to organizational researchers: A selected application to turnover research. Journal of Applied Psychology, 74, 280-292.

Morita, J. G., Lee, T. W., & Mowday, R. T. (1993). The regression-analog to survival analysis: A selected application to turnover research. Academy of Management Journal, 36, 1430-1464

Motowidlo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. Journal of Applied Psychology, 79, 475-480.

Motowidlo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. Human Performance, 10, 71-83

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-Factor Model of personality and performance in jobs involving interpersonal interactions. Human Performance, 11(2-3), 145-165

Mount, M. K., Judge, T. A., Scullen, S. E., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater and level effects in 360-degree performance ratings. Personnel Psychology, 51(3), 557-576.

Mount, M.K. & Barrick, M.R. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. In G. R. Ferris (Ed.) Research in Personnel and Human Resources Management (Volume 13) 153-200 New York: JAI.

Moynihan, L. M. & Peterson, R. S. (2001). A contingent configuration approach to understanding the role of personality in organizational groups. Research in Organizational Behavior, 23, 327-378.

Mueller, C. W., Finley, A., Iverson, R. D., & Price, J. L. (1999). The effects of group racial composition on job satisfaction, organizational commitment, and career commitment: The case of teachers. Work and Occupations, 26(2), 187-219.

Murname, R. J., Singer, J.D., & Willett, J. B. (1989). The career paths of teachers: Implications for teacher supply and methodological lessons for research. Educational Researcher, 17, 22-30.

Murphy, K. R. (1996). Individual Differences and Behavior in Organizations: Much More than g. In K. R. Murphy (Ed.). Individual Differences and Behavior in Organizations. (pp. 3-30). San Fransisco: Jossey-Bass.

Murphy, K.R. & Shiarella, A.H. (1998). Implications of the multidimensional nature of job performance for the validity of seletion tests: Multivariate frameworks for studying test validity: Errata. Personnel Psychology, 51, 822.

National Research Council (1999). The Changing Nature of Work. Washington DC: National Academy Press.

Neuman, G. A., & Wright, J. (1999). Team effectiveness: Beyond skills and cognitive ability. Journal of Applied Psychology, 84, 376-389.

Neuman, G. A., Wagner, S. H., & Christiansen, N. D. (1999). The relationship between work-team personality composition and the job performance of teams. Group and Organization Management, 24, 28-45

Nezlek, J. B., & Zyzniewski, L. E. (1998). Using hierarchical linear modeling to analyze grouped data. Group Dynamics, 2(4), 313-320.

Nieva, V. F., Fleishman, E. A., & Rieck, A. M. (1978). Team dimensions: Their identity, their measuremetn and their relationships. Washington DC: ARRO.

Nunnally, J. (1978). Psychometric Theory (2nd Ed). New York: McGraw Hill.

O' Reilly, C. A., III, Williams, K. Y., & Barsade, S. (1998). Group demography and innovation: Does diversity help?, Gruenfeld, Deborah H (Ed). (1998). Composition. Research on managing groups and teams, Vol. 1. (pp. 183 207). Stamford, CT, US: JAI Press, Inc.

O'Connor, K. M. (1998). Experiental diversity in groups: Conceptualizing and measuring variation among teammates. In Gruenfeld, D. H. (Ed.), Research on Managing Groups and Teams: Composition (pp. 167-182) Stamford, CT: JAI Press.

O'Reilly, C. A., Williams, K. Y., & Barsdale, S. (1998). Group demography and innovation: Does diversity help? In Gruenfeld, D. H. (Ed.), Research on Managing Groups and Teams: Composition (pp. 183-207) Stamford, CT: JAI Press.

O'Reilly, C. A., Caldwell, D. F., & Barnett, W. P. (1989). Work group demography, social integration, and turnover. Administrative Science Quarterly, 34(1), 21-37.

O'Reilly, C. A., Chatman, J., & Caldwell, D. F. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. Academy of Management Journal, 34(3), 487-516.

Ostroff, C. (1993). Comparing correlations based on individual-level and aggregated data. Journal of Applied Psychology, 78(4), 569-582.

Ostroff, C., & Bowen, D. E. (2000). Moving HR to a higher level: HR practices and organizational effectiveness, K. J. Klein & S. W. J Kozlowski (Eds). (2000). Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions. (pp. 211-266). San Francisco, CA, US: Jossey Bass Inc, Publishers.

Pelled, L. H. (1996). Demographic diversity, conflict, and work group outcomes: An intervening process theory. Organization Science, 7(6), 615-631.

Pelled, L. H. (1996). Relational demography and perceptions of group conflict and performance: A field investigation. International Journal of Conflict Management, 7(3), 230-246.

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), 1-28.

Pelled, L. H., Xin, K. R., & Weiss, A. M. (2001). No es como mi: Relational demography and conflict in a Mexican production facility. Journal of Occupational and Organizational Psychology, 74(1), 63-84.

Peters, L. H., & Sheridan, J. E. (1988). Turnover research methodology: A critique of traditional designs and a suggested survival model alternative. In K. M. Rowland & G. R. Ferris (Eds.), Research in Personnel and Human Resource Management (Volume 10). Grennwich, CT: JAI Press.

Pfeffer, J. (1983). Organizational demography. Research in Organizational Behavior, 5, 299-357.

Phillips, A. S., & Bedeian, A. G. (1994). Leader-follower exchange quality: The role of personal and interpersonal attributes. Academy of Management Journal, 37(4), 990-1001.

Ployhart, R. E. (1998). A multilevel perspective on personnel selection. Paper presented at the Annual Meeting of the Academy of Management, San Diego, CA.

Ployhart, R. E., & Hakel, M. D. (1998). The substantive nature of performance variability: Predicting interindividual differences in intraindividual performance. Personnel Psychology, 51(4), 859-901.

Pugh, S. D. (2001). Service with a smile: Emotional contagion in the service encounter. Academy of Management Journal, 44(5), 1018-1027.

Ragins, B. R., & Cornwell, J. M. (2001). Pink triangles: Antecedents and consequences of perceived workplace discrimination against gay and lesbian employees. Journal of Applied Psychology, 86(6), 1244-1261.

Ree, M.J. & Earles, J.A. (1992). Intelligence is the best predictor of job performance. Current Directions in Psychological Science, 1, 86-89.

Reise, S. P. (2000). Using multilevel logistic regression to evaluate person-fit in IRT models. Multivariate Behavioral Research, 35, 543-568.

Rentcsh, J. R. & Hall, R. J. (1994). Members of great teams think alike: A model of team effectiveness and schema similarity among team members. In M. M. Beyerlein & D. A. Johnson. (Eds.), Advances in Interdisciplinary Studies of Work Teams: Theories of Self-Managing Work Teams. Volume 1 (pp. 223-262). Greenwich, CT: JAI Press.

Richard, O. C. (2000). Racial diversity, business strategy, and firm performance: A resource-based view. Academy of Management Journal, 43(2), 164-177.

Riordan, C. M & Holliday-Wayne, J. (1998). Work group gender composition attitudes: A test of competing hypotheses. Proceedings of the Southern Management Association.

Riordan, C. M. & Weatherly, E. W. (1999). Relational demography within groups: An empirical test of a theoretical model. Paper presented at the annual meeting of the Academy of Management, Chicago.

Riordan, C. M. (1997). Advancing relational demography theory: A construct validity study of three measures of demographic similarity. In L. N. Doeler & J. B. Keys (Eds.). Academy of Management Best Paper Proceedings.

Riordan, C. M. (2000). Relational demography within groups: Past developments, contradictions, and new directions. In G. R. Ferris (Ed.). Research in Personnel and Human Resources Management. Volume 19 (pp. 131-174). New York: JAI.

Riordan, C. M., & Shore, L. M. (1997). Demographic diversity and employee attitudes: An empirical examination of relational demography within work units. Journal of Applied Psychology, 82(3), 342-358.

Roberts, K. H. & Burstein, K. (1980). Issues of aggregation. In K. H. Roberts & K. Burstein (Eds.), New directions for methodology in the social sciences. (Volume 6). San Fransisco: Jossey-Bass.

Roberts, K. H., Hulin, C. L., & Rousseau, D. M. (1978). Developing an interdisciplinary science of organizations. San Fransisco: Jossey-Bass.

Rosse, J. G. & Noel, T. W. (1996). Leaving the organization. In K. R. Murphy (Ed.). Individual Differences and Behavior in Organizations. (pp. 451-506). San Francisco: Jossey-Bass.

Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. Journal of Applied Psychology, 83(4), 634-644.

Roth, P. L., Bobko, P., Switzer, F. S., III, & Dean, M. A. (2001). Prior selection causes biased estimates of standardized ethnic group differences: Simulation and analysis. Personnel Psychology, 54(3), 591-617.

Rousseau, D. M. (1985). Issues of Level in Organizational Research: Multi-Level and Cross-Level Perpsectives. In L. L. Cummings & B. M. Staw (Eds.), Research in Organizational Behavior. Volume 7 (pp. 1-37). Greenwich, CT: JAI Press.

Rousseau, D. M., & House, R. J. (1994). MESO organizational behavior: Avoiding the fundamental biases. In C. L. Cooper & D. M. Rousseau (Eds.), Trends in Organizational Behavior Volume 1 (pp. 13-30). London: Wiley.

Rousseau, D.M. (1977). Technological differences in job characteristics, employee satisfaction, and motivation: A synthesis of job design research and sociotechnical systems theory. Organizational Behavior and Human Decision Processes, 19, 18-42.

Rousseau, D.M. (1985). Issues of level in organizational research: Multilevel and cross-level perspectives. In L.L. Cummings & B.Staw (Eds.), Research in organizational behavior (Vol 7, pp. 1-37). Greenwich, CT: JAI Press.

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. Personnel Psychology, 50(3), 707-721.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas. Teams: Their Training and Performance. (pp. 3-30). USA: Ablex.

Schmidt, F. L. (1993). Personnel psychology at the cutting edge. In N. Schmitt & W. C. Borman and Associates, Personnel Selection in Organizations. (pp. 497-516). San Fransisco: Jossey-Bass.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychological Bulletin, 124, 262-274

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychological Bulletin, 124(2), 262-274.

Schmidt, F. L., Ones, D. S., & Hunter, J. E. (1992). Personnel selection. Annual Review of Psychology, 43, 627-670.

Schmitt, N. & Chan, D. (1998). Personnel selection: A theoretical approach. Thousand Oaks, CA: Sage.

Schmitt, N. (1989). Fairness in employment selection.

Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency of various predictor combinations. Journal of Applied Psychology, 82, 719-730.

Schneider, B. (1985). Organizational Behavior. Annual Review of Psychology, 36, 573-611.

Schneider, B. (1987). The people make the place. Personnel Psychology, 40(3), 437-453.

Schneider, B. (1996). When individual differences aren't. In K. R. Murphy (Ed.). Individual Differences and Behavior in Organizations. (pp. 548-572). San Francisco: Jossey-Bass.

Schneider, B., Smith, D. B., & Sipe, W. P. (2000). Personnel selection psychology:

Multilevel considerations, K. J. Klein & S. W. J Kozlowski (Eds). (2000). Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions. (pp. 910120). San Francisco, CA, US: Jossey Bass Inc, Publishers.

Scullen, S. E., Mount, M. K., & Goff, M. (2000). Understanding the latent structure of job performance ratings. Journal of Applied Psychology, 85(6), 956-970.

Sessa, V. I. & Jackson, S. E. (1995). Diversity in decision-making teams: All differences are not created equal. In M. M. Chemers, M. A. Costanzo, M. A. & S. Oskamp (Eds.), Diversity in Organizations: New Perspectives for a Changing Workplace. (pp. 157-190). Thousand Oaks, CA: Sage.

Shaw, J. B., & Barrett Power, E. (1998). The effects of diversity on small work group processes and performance. Human Relations, 51(10), 1307-1325.

Shaw, J. D., Delery, J. E., Jenkins, G. D., & Gupta, N. (1998). An organization-level analysis of voluntary and involuntary turnover. Academy of Management Journal, 41, 511-525.

Shaw, M.E. (1971). Group dynamics, the psychology of small goup behavior. New York: McGraw-Hill.

Sherif, M., Harvey, O. J., White, B. J., Hood, W R., & Sherif, C. W. (1961). Intergroup conflict and cooperation. The Robbers Cave Experiment. Norman, OK: Institute of Group Relations.

Simon, H. A. (1973). The organization of complex systems. In H. H. Patee (Ed.). Hierarchy Theory (pp.1-27). New York: Braziller.

Singer, J. D. & Willett, J. B. (1991). Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. Psychological Bulletin, 110, 268-290.

Singer, J. D., & Willett, J. B. (1991). Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. Psychological Bulletin, 110(2), 268-290.

Smith, P. C. (1976). Behavior, results, and organizational effectiveness: The problem of criteria. In M. D. Dunnette (Ed.), Handbook of Industrial and Organizational Psychology (pp. 745-775). Chicago: Rand McNally.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. Multivariate Behavioral Research, 25, 173-180.

Steinder, I.D. (1972). Group process and productivity. New York: Acadmic Press.

Stevens, M. J., & Campion, M. A. (1994). The knowledge, skill, and ability requirements for teamwork: Implications for human resource management. Journal of Management, 20(2), 503-530.

Stoolmiller, M. (1995). Using latent growth curve models to study developmental processes. In J. Gottman (Ed.), The Analysis of Change (pp. 103-138). Hillsdale, NJ: Erlbaum.

Stumpf, S. A., & Dawley, P. K. (1981). Predicting voluntary and involuntary turnover using absenteeism and performance indices. Academy of Management Journal, 24(1), 148-163.

Sundstrom, E., de Meuse, K.P. & Futrell, D. (1990). Work teams: Applications and effectiveness. American Psychologist, 45, 120-133.

Sundstrom, E., McIntyre, M., Halfhill, T., & Richards, H. (2000). Work groups: From the Hawthorne studies to work teams of the 1990s and beyond. Group Dynamics, 4(1), 44-67.

Swezey, R.W. & Salas, E. (1992). Guidelines for use in team-training development. In R.W. Swezey & E. Salas (Eds.), Teams: Their training and performance (pp. 219-245). Stamford, CT: Ablex Publishing Corp.

Tabachnick, B. G. & Fidell (1996). Using Multivariate Statistics (3rd Ed.). New York: Harper Collins.

Tajfel, H. (1981). Human Groups and Social Categories: Studies in Social Psychology. Cambridge, England: Cambridge University Press.

Tajfel, H. (1982). Social psychology of intergroup relations. Annual Review of Psychology, 33, 1-39.

Teachman, J. D. (1980). Analysis of population diversity: Measures of qualitative variation. Sociological Methods and Research, 8, 341-362.

Tellegen, A. (1985). Structure of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A.H. Tuma & J.D. Maser (Eds.), Anxiety and the anxiety disorder (pp. 681-706). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Tesluk, P. E., & Mathieu, J. E. (1999). Overcoming roadblocks to effectiveness: Incorporating management of performance barriers into models of work group effectiveness. Journal of Applied Psychology, 84, 200-217.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. Personnel Psychology, 44, 703-742

Timmerman, T. A. (2000). Racial diversity, age diversity, interdependence, and team performance. Small Group Research, 31(5), 592-606.

Tjosvold, D. & Tjosvold, M. M. (1994). Cooperation, competition, and constructive controversey: Knowledge to empower self-managing work teams. In M. M. eyerlein & D. A. Johnson. (Eds.) Advances in Interdisciplinary Studies of Work Teams: Theories of Self-Managing Work Teams. Volume 1 (pp. 119-144). Greenwich, CT: JAI Press.

Tjosvold, D. (1995). Cooperation theory, constructive controversy, and effectiveness: Learning from crisis. In R. A. Guzzo & E. Salas and Associates (Eds.), Team Effectiveness and Decision Making in Organizations. San Fransisco: Jossey-Bass.

Tosi, H. (1992). The Environment/Organization/Person Contingency Model: A Meso Approach to the Study of Organizations. Greenwich, CT: JAI Press.

Trevor, C. O., Gerhart, B., & Boudreau, J. W. (1997). Voluntary turnover and job performance: Curvilinearity and the moderating influences of salary growth and promotions. Journal of Applied Psychology, 82(1), 44-61.

Tsui, A. S., & O' Reilly, C. A. (1989). Beyond simple demographic effects: The importance of relational demography in superior-subordinate dyads. Academy of Management Journal, 32(2), 402-423.

Tsui, A. S., Egan, T. D., & O' Reilly, C. A. (1992). Being different: Relational demography and organizational attachment. Administrative Science Quarterly, 37(4), 549-579.

Tsui, A. S., Egan. T. D., & Xin, K. R. (1995). Diversity in organizations: Lessons from demography research. In M. M. Chemers, M. A. Costanzo, M. A. & S. Oskamp (Eds.), Diversity in Organizations: New Perspectives for a Changing Workplace. (pp. 191-219). Thousand Oaks, CA: Sage.

Tsui, A. S., Xin, K. R., & Egan, T. D. (1995). Relational demography: The missing link in vertical dyad linkage, Jackson, Susan E. (Ed); Ruderman, Marian N. (Ed). (1995). Diversity in work teams: Research paradigms for a changing workplace. (pp. 97 129). Washington, DC, US: American Psychological Association.

Turban, D. B., & Jones, A. P. (1988). Supervisor-subordinate similarity: Types, effects, and mechanisms. Journal of Applied Psychology, 73(2), 228-234.

Turner, J. (1982). Toward a cognitive definition of the group. In H. Tajfel (Ed.). Social Identity and Intergroup Relations. Cambridge: Cambridge University press.

Turner, J. (1987). Rediscovering The Social Group: A Social Categorization Theory. Oxford, UK: Blackwell.

Tziner, A., & Eden, D. (1985). Effects of crew composition on crew performance: Does the whole equal the sum of its parts? Journal of Applied Psychology, 70(1), 85-93.

Van der Vegt, G., Emans, B. J. M., & Van De Vliert, E. (2001). Patterns of interdependence in work teams: A two-level investigation of the relations with job and team satisfaction. Personnel Psychology, 54(1), 51-69.

Van Scotter, J. R., & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. Journal of Applied Psychology, 81, 525-531.

Vandenberghe, C. (1999). Organizational culture, person-culture fit, and turnover: A

replication in the health care industry. Journal of Organizational Behavior, Organizational-Behavior.

Viswesvaran, C. & Ones, D.S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. Educational and Psychological Measurement, 59, 197-210.

Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. Journal of Applied Psychology, 81(5), 557-574.

Voiers, W. D. (1956). Bombing accuracy as a function of the ground-school proficiency structure of the B-29 bomb team. (Research Report AFDTRC-TN-56-4) Lackland Air Force Base, TX: Air Force Personnel and Training Research Center.

Wagner, W. G., Pfeffer, J., & O' Reilly, C. A. (1984). Organizational demography and turnover in top-management groups. Administrative Science Quarterly, 29(1), 74-92.

Wanous, J. P., Stumpf, S. A., & Bedrosian, H. (1979). Job survival of new employees. Personnel Psychology.

Watson, W. E., Kumar, K., & Michaelsen, L. K. (1993). Cultural diversity's impact on interaction process and performance: Comparing homogeneous and diverse task groups. Academy of Management Journal, 36(3), 590-602.

Weiner, B. (1979). A theory of motivation for some classroom experiences. Journal of Educational Psychology, 71, 3-25.

Weiner, B. (1985). An attributional theory of motivation and emotion. Psychological Review, 92, 548-573.

Weingart, L. R. (1997). How did they do that? The ways and means of studying group process. In L. L. Cummings & B. M. Staw (Eds.), Research in Organizational Behavior. Volume 19 (pp. 189-240). Greenwich, CT: JAI Press.

Wells, D. L., & Muchinsky, P. M. (1985). Performance antecedents of voluntary and involuntary managerial turnover. Journal of Applied Psychology, 70(2), 329-336.

Werbel, J. D. & Johnson, D. J. (2001). The use of person-group fit for employee selection: A missing link in person-environment fit. Human resource Management, 40, 227-240.

Werbel, J. D., & Gilliland, S. W. (1999). Person-environment fit in the selection process.

Wesolowski, M. A., & Mossholder, K. W. (1997). Relational demography in supervisor-subordinate dyads: Impact on subordinate job satisfaction, burnout, and perceived procedural justice. Journal of Organizational Behavior, 18(4), 351-362.

Westphal, J. D., & Zajac, E. J. (1995). Who shall govern? CEO/board power, demographic similarity, and new director selection. Administrative Science Quarterly, 40(1), 60-83.

Wharton, A. S., Rotolo, T., & Bird, S. R. (2000). Social context at work: A multilevel analysis of job satisfaction. Sociological Forum, 15(1), 65-90.

Wiersema, M. F., & Bird, A. (1993). Organizational demography in Japanese firms: Group heterogeneity, individual dissimilarity, and top management team turnover. Academy of Management Journal, 36(5), 996-1025.

Wilk S. L., Desmarais L. B., Sackett P. R. (1995). Gravitation to jobs commensurate with ability: Longitudinal and cross-sectional tests. Journal of Applied Psychology, 80, 79-85.

Wilk, S. L., & Sackett, P. R. (1996). Longitudinal analysis of ability-job complexity fit and job change. Personnel Psychology, 49(4), 937-967.

Willet, J. B & Singer, J. D. (1995). It's Déjà vu All Over Again: Using multiple-spell discrete time survival analysis. Journal of Educational and Behavioral Statistics, 20, 41-67.

Willett, J. B. & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. Psychological Bulletin, 116, 363-381.

Willett, J. B. & Singer, J. D. (1993). Investigating onset, cessation, relapse, and recovery: Why you should, and how you can, use discrete-time survival analysis to examine event occurrence. Journal of Consulting and Clinical Psychology, 61, 952-965.

Williams, K. Y. & O'Reilly, C. A. (1998). Demography and diversity in organizations: A review of 40 years of research. Research in Organizational Behavior, 20, 77-140.

Williams, W. M., & Sternberg, R. J. (1988). Group intelligence: Why some groups are better than others. Intelligence, 12, 351-377.

Yammarino (199x). A conceptual-empirical approach for testing meso and Multi-Level Theories. In H. L. Tosi (Ed.) Extensions of the Environment/Organization/Person Model (Volume 2). Greenwich, CT: JAI Press.

Zalesny, M. D., Salas, E., & Prince, C. (1995). Conceptual and measurement issues in coordination: Implications for team behavior and performance. In G. R. Ferris (Ed.). Research in Personnel and Human Resources Management. Volume 13 (pp. 81-116). New York: JAI Press.

Zeidner, J. & Johnson, C. D. (1991). The Economic Benefits of Predicting Job Performance: Selection Utility (Volume 1). New York: Praeger.

Zenger, T. R., & Lawrence, B. S. (1989). Organizational demography: The differential effects of age and tenure distributions on technical communication. Academy of Management Journal, 32(2), 353-376.

Table 1. Summary of Hypotheses and Planned Analytic Methods

Type of Hypothesis	Hypothesis	Analysis Method
Individual Level	H8. Conscientiousness will be negatively related to VT.	Cox Regression
Individual Level	H9. Conscientiousness will be negatively related to IVT.	Cox Regression
Individual Level	H10. High ability employees will be more likely to exhibit VT as opposed to IVT after controlling for Conscientiousness.	Cox Regression
Team Level	H1. The team mean ability score will be positively related to team performance.	Regression
Team Level	H2. The team mean conscientiousness score will be positively related to team performance.	Regression
Team Level	H3. The team mean conscientiousness score will be positively related to team performance after controlling for the effects of team cognitive ability.	Regression
Team Level	H4. Mean level of team Agreeableness will relate to team performance	Regression
Team Level	H5. Mean level of team agreeableness will relate to team performance after controlling for cognitive ability.	Regression
Team Level	H6. Mean levels of team extraversion will exhibit a nonlinear (inverted U-shaped) relationship with team performance.	Regression
Team Level	H7. Team manager tenure will moderate the relationship between mean team KSAO levels and team performance such that there is a stronger relationship between KSAO levels and performance for teams with high-tenure managers than with those with low tenure managers.	Regression
Team Level	H11. Increased team level diversity in conscientiousness should be associated with lower team performance.	Regression
Team Level	H12. Increased diversity in conscientiousness should be associated in higher team voluntary turnover (but not involuntary turnover).	Cox Regression

Table 1 (cont'd)

Team Level	H15. Increased team racial diversity will be related to increased team turnover rates. This effect will be evident after controlling for the relevant personality constructs and ability.	Latent Growth Curve Structural Equation Modelling
Team Level	H16. Increased racial diversity will be negatively related to team performance. This effect will be evident after controlling for the relevant personality constructs and ability.	Latent Growth Curve Structural Equation Modelling
Team Level	H18. Increased team age diversity will be related to increased team turnover rates. This effect will be evident after controlling for the relevant personality constructs and ability.	Latent Growth Curve Structural Equation Modelling
Team Level	H19. Increased age diversity will be negatively related to team performance. This effect will be evident after controlling for the relevant personality constructs and ability.	Latent Growth Curve Structural Equation Modelling
Team Level	H21 Increased team sex diversity will be related to increased team turnover rates	Latent Growth Curve Structural Equation Modelling
Team Level	H22. Increased sex diversity will be negatively related to team performance	Latent Growth Curve Structural Equation Modelling
Cross-Level	H13. Increased conscientiousness-based person-team fit should be associated with a lower likelihood of individual level voluntary turnover (but not involuntary turnover).	Multilevel Survival Analysis
Cross-Level	H14. The racial fit of the individual within the team will predict individual turnover such that misfit is associated with a higher likelihood of turnover. This relationship will be more evident in Whites working in teams with higher minority representations rather than vice-versa.	Multilevel Survival Analysis

Table 1 (cont'd)

Cross-Level	H17. The age-based fit of the individual within the team will predict individual turnover such that misfit is associated with a higher likelihood of turnover.	Multilevel Survival Analysis
Cross-Level	H20. The sex-based fit of the individual within the team will predict individual turnover such that misfit is associated with a higher likelihood of turnover.	Multilevel Survival Analysis
Cross-Level	H23. The fit between an individual's demographics and that of the team will be more strongly related to turnover at lower levels of tenure.	Multilevel Survival Analysis
Cross-Level	H24. The effects of demographic fit within a team will depend on an individual's extraversion such that high extraversion individuals that do not demographically fit will be less likely to turnover than low extraversion individuals who do not fit.	Multilevel Survival Analysis

Table 2. Descriptive statistics, internal consistency estimates, uncorrected and corrected correlations between the CRI and IPI scales (N = 135).

	Mean	SD	CRICON	CRIEX	CRIAG	CRIAG-R	IPIEX	IPIAG	IPICON
CRICON	3.69	.47	(.81)	.18*	-.01	.09	.07	.14	.66*
CRIEX	3.61	.65	.23	(.73)	.21*	.35*	.67*	.34*	.04
CRIAG	3.39	.52	-.01	.27	(.82)	.82*	.03	.19*	.02
CRIAG-R	3.55	.47	.12	.51	1.11	(.66)	.20*	.44*	.07
IPIEX	3.42	.65	.08	.82	.03	.25	(.91)	.36*	-.02
IPIAG	4.15	.41	.17	.45	.24	.61	.43	(.79)	.05
IPICON	4.13	.44	.80	.05	.02	.09	-.02	.06	(.85)

Notes: CRICON = CRI Conscientiousness; CRIEX = CRI Extraversion; CRIAG = CRI Agreeableness; CRIAG-R = CRI Agreeableness Revised; IPIEX = IPI Extraversion; IPIAG = IPI Agreeableness; IPICON = IPI Conscientiousness. Coefficient alpha internal consistency estimates are on the diagonol, uncorrected correlations are above the diagonol, and correlations corrected for internal consistency unreliability are below the diagonol. * $p<.05$.

Table 3. Test-retest correlations of CRI and IPI scales at a one-week interval (N = 29).

	1-RT	2-RT	3-RT	4-RT	5-RT	6-RT	7-RT
1. CRICON	.83*	.08	-.10	.05	-.16	-.03	.62*
2. CRIEX	.10	.82*	.31	.43*	.62*	.30	.11
3. CRIAG	-.18	.28	.88*	.72*	.08	.20	.09
4. CRIAGR	-.13	.28	.74*	.83*	.19	.51*	.11
5. IPIEX	-.10	.67*	.08	.24	.91*	.30	-.19
6. IPIAG	.13	.35*	.18	.42*	.19	.83*	-.08
7. IPICN	.68*	.18	.17	.25	-.19	-.03	.89*

Notes: RT = Retest; CRICON = CRI Conscientiousness; CRIEX = CRI Extraversion; CRIAG = CRI Agreeableness; CRIAG-R = CRI Agreeableness Revised; IPIEX = IPI Extraversion; IPIAG = IPI Agreeableness; IPICN = IPI Conscientiousness. *p<.05.

Table 4 . Race and Sex Composition of the Sample.

	Black	American Indian/Alaskan	Asian/Pacific Islander	Hispanic	White	Other	Total
Female	31,995	1,957	2,418	16,202	44,897	71	97,540
Male	26,038	1,825	2,586	18,372	43,730	65	92,616
Total	58,033	3,782	5,004	34,574	88,627	136	190,156

Table 5. Age composition of the sample (Valid N = 251,355)

Age Range	N	%
<20	103,180	54.3%
20-24	42,095	22.1%
25-29	18,996	10.0%
30-34	11,066	5.8%
35-39	6,658	3.5%
40-44	3,278	1.7%
45-49	1,688	0.9%
50-54	895	0.5%
55-59	435	0.2%
>59	1,025	0.5%
Total	189,257	100%

Table 6. Categories of Turnover and Incidence

Reason	N	%
Voluntary		
Other Job	36,948	26.2%
Moved	6,111	4.3%
Personal	34,291	24.3%
Voluntary Other	13,400	9.5%
Subtotal	90,750	64.4%
Involuntary Avoidable		
Unacceptable Attendance	26,099	18.5%
Substandard Performance	2,920	2.1%
Violation of Policy	4,791	3.4%
Gross Misconduct	2,489	1.8%
Job Abandonment	726	0.5%
Involuntary Other	2,510	1.8%
Resigned in Lieu of Discharge	39,535	28.1%
Subtotal	26,099	18.5%
Involuntary Unavoidable		
Death	269	0.2%
No Work Available/Layoff	17	0.0%
Unavailable for Scheduled Work Hours	91	0.1%
Military	1,563	1.1%
Transfer	262	0.2%
Other	8,416	6.0%
Subtotal	10,618	7.5%
Total	140,903	100%

Table 7. Descriptives statistics and intercorrelations among restaurant-level variables.

	Mean	SD	1.	2.	3.	4.	5.	6.	7.
1. CP PD4	$16,415	$8,280	1						
2. CP PD5	$17,612	$8,585	.91*	1					
3. CRI Saturation PD5	.39	.20	-.06*	-.04*	1				
4. Sales TXs PD5	73,705	21,544	.90*	.92*	-.09*	1			
5. Ability PD5	6.21	1.38	.02	.04*	.15*	.00	1		
6. Extraversion PD5	4.26	.30	-.03	-.04*	.04*	-.04*	.19*	1	
7. Cons. PD5	4.43	.22	-.01	-.02	.10*	-.03	.21*	.50*	1
8. Agreeableness PD5	4.26	.25	.04*	.03	.07*	.02	.20*	.56*	.66*

Notes: CP = Controllable Profit; PD4 = Period 4; PD5 = Period 5; TXs = Transactions; Cons = Conscientiouensss; N = 2205 – 2207 for personality variables and CRI saturation; N = 2500 – 2502 for financial variables. *p < .05. Figures for test scores are averages within each store.

Table 8. Hierarchical regressions for mean ability levels as a predictor of controllable profit (N =2,206).

Step and Predictors	Statistics for Step					Statistics for Predictors	
	R^2	df	ΔR^2	Δdf	ΔF	β	t
Step 1	.854	3	—	—	4,287.26		
Profit PD 4						.64*	39.89
CRI Saturation PD 5						.04*	4.62
# of Sales PD5						.32*	19.92
Step2	.855	4	.001	1	13.13		
Ability PD 5						.03*	3.62
Step 3	.855	5	.000	1	.592		
CRI Saturation X Ability						.03	.77

Notes: * $p < .05$. Dependent variable is Controllable Profit for Period 5. PD = Period. Measures collected at the individual-level are averages calculated within the stores.

Table 9. Hierarchical regressions for mean conscientiousness levels as a predictor of controllable profit (N =2,207).

Step and Predictors	Statistics for Step					Statistics for Predictors	
	R^2	df	ΔR^2	Δdf	ΔF	β	t
Step 1	.854	3	-	-	4,290.31		
Profit PD 4						.64*	39.89
CRI Saturation PD 5						.04*	4.63
# of Sales PD5						.32*	19.92
Step2	.854	4	.000	1	2.31		
Conscientiousness PD 5						-.01	-1.52
Step 3	.854	5	.000	1	.04		
CRI Saturation X Conscientiousness						.04	.21

Notes: * $p < .05$. Dependent variable is Controllable Profit for Period 5. PD = Period. Measures collected at the individual-level are averages calculated within the stores.

Table 10. Hierarchical regressions for mean agreeableness levels as a predictor of controllable profit (N =2,207).

Step and Predictors	Statistics for Step					Statistics for Predictors	
	R^2	df	ΔR^2	Δdf	ΔF	β	t
Step 1	.854	3	-	-	4,290.31		
Profit PD 4						.64*	39.89
CRI Saturation PD 5						.04*	4.63
# of Sales PD5						.32*	19.92
Step2	.855	4	.000	1	3.41		
Agreeableness PD 5						-.02	-1.85
Step 3	.855	5	.000	1	.391		
CRI Saturation X Agreeableness						-.09	-.62

Notes: * $p < .05$. Dependent variable is Controllable Profit for Period 5. PD = Period. Measures collected at the individual-level are averages calculated within the stores.

Table 11. Hierarchical regressions for mean extraversion and extraversion^2 levels as a predictor of controllable profit (N =2,205).

Step and Predictors	Statistics for Step					Statistics for Predictors	
	R^2	df	ΔR^2	Δdf	ΔF	β	T
Step 1	.854	3	—	—	3228.06		
Profit PD 4						.64*	39.96
CRI Saturation PD 5						.04*	4.79
# of Sales PD5						.32*	19.94
Tenure PD5						.02*	2.78
Step2	.854	4	.00	1	7.33		
Extraversion PD 5						-.02*	-2.71
Step 3	.854	5	.00	1	.366		
Extraversion^2						.08	.61

Notes: * $p < .05$. Dependent variable is Controllable Profit for Period 5. PD = Period. Measures collected at the individual-level are averages calculated within the stores

Table 12. Hierarchical regressions for mean conscientiousness X manager tenure levels as a predictor of controllable profit (N =2,206).

Step and Predictors	Statistics for Step					Statistics for Predictors	
	R^2	df	ΔR^2	Δdf	ΔF	β	t
Step 1	.854	3	—	—	4,290.31		
Profit PD 4						.64*	39.89
CRI Saturation PD 5						.04*	4.63
# of Sales PD5						.32*	19.92
Step2	.856	5	.001	2	9.86		
Manager Tenure						.03*	4.17
Conscientiousness						-.01	-1.54
Step 3	.856	6	.000	1	.989		
Manager Tenure x Conscientiousness						.16	1.00

Notes: * $p < .05$. Dependent variable is Controllable Profit for Period 5. PD = Period. Measures collected at the individual-level are averages calculated within the stores.

Table 13. Hierarchical regressions for mean agreeableness X manager tenure levels as a predictor of controllable profit (N =2,206).

Step and Predictors	Statistics for Step					Statistics for Predictors	
	R^2	df	ΔR^2	Δdf	ΔF	β	t
Step 1	.854	3	—	—	4,290.31		
Profit PD 4						.64*	39.89
CRI Saturation PD 5						.04*	4.63
# of Sales PD5						.32*	19.92
Step2	.855	5	.001	2	10.16		
Manager Tenure						.03*	4.11
Agreeableness						-.01	-1.73
Step 3	.855	6	.000	1	.583		
Mgr. Tenure x Agreeableness						.11	.76

Notes: * $p < .05$. Dependent variable is Controllable Profit for Period 5. PD = Period. Measures collected at the individual-level are averages calculated within the stores.

Table 14. Hierarchical regressions for mean ability X manager tenure levels as a predictor of controllable profit (N =2,206).

Step and Predictors	Statistics for Step					Statistics for Predictors	
	R^2	df	ΔR^2	Δdf	ΔF	β	t
Step 1	.854	3	—	—	4,290.31		
Profit PD 4						.64*	39.89
CRI Saturation PD 5						.04*	4.63
# of Sales PD5						.32*	19.92
Step2	.856	5	.002	2	14.45		
Manager Tenure						.03*	3.96
Ability						.03*	3.40
Step 3	.856	6	.000	1	-.02		
Mgr. Tenure x Ability						-.02	-.481

Notes: * $p < .05$. Dependent variable is Controllable Profit for Period 5. PD = Period. Measures collected at the individual-level are averages calculated within the stores.

Table 15. Hierarchical regressions for coefficient of variability in conscientiousness as a predictor of controllable profit (N =2,029).

Step and Predictors	Statistics for Step					Statistics for Predictors	
	R^2	df	ΔR^2	Δdf	ΔF	β	t
Step 1	.853	3	—	—	3,924.45		
Profit PD 4						.64*	38.15
CRI Saturation PD 5						.04*	4.13
# of Sales PD5						.32*	19.25
Step2	.853	4	.000	1	2.30		
Conscientiousness COV						.01	1.51

Notes: * $p < .05$. Dependent variable is Controllable Profit for Period 5. PD = Period. Measures collected at the individual-level are averages calculated within the stores. COV = Coefficient of variation, the standard deviation divided by the mean for a given store.

Table 16. Correlations among 11-month aggregates of restaurant level demographic diversity measures and turnover rate (N = 2496-2498).

	Mean	SD	1.	2.	3.	4.
1. Race Diversity	3.76	1.96	1			
2. Sex Diversity	4.84	.46	.16	1		
3. Age Diversity	26,738	9,934	.14	.17	1	
4. Turnover	.56	.09	.09	.10	.04	1

Notes: All correlations are significant at $p < .05$. The demographic diversity measures were aggregated across the time periods by summing the values, whereas the turnover rate was averaged.

Table 17. Multiple regression analyses for the prediction of 11-month aggregate of turnover rates with 11-month aggregates of age, sex, and demographic diversity.

Step and Predictors	Statistics for Step					Statistics for Predictors	
	R^2	df	ΔR^2	Δdf	ΔF	β	t
Step 1	.015	3	—	—	12.81		
Age Diversity						.02	.78
Race Diversity						.07	3.64*
Sex Diversity						.08	4.03*

Notes: * $p < .05$.

Table 18. Means, standard deviations, and intercorrelations for racial diversity across the 11 time periods (N = 2,388).

Variable	Mean	SD	RD8	RD9	RD10	RD11	RD12	RD13	RD1	RD2	RD3	RD4	RD5
RD8	.38514	.20747	1.00										
RD9	.38545	.20692	.93	1.00									
RD10	.38452	.20801	.87	.94	1.00								
RD11	.38346	.20982	.84	.89	.95	1.00							
RD12	.38305	.20859	.82	.86	.90	.95	1.00						
RD13	.37994	.20900	.81	.84	.88	.91	.96	1.00					
RD1	.37802	.20956	.78	.82	.85	.88	.91	.96	1.00				
RD2	.37736	.20818	.77	.80	.83	.85	.88	.91	.95	1.00			
RD3	.37822	.20682	.75	.79	.81	.83	.85	.88	.90	.94	1.00		
RD4	.38013	.20853	.74	.77	.79	.81	.83	.85	.87	.89	.94	1.00	
RD5	.37547	.20900	.72	.75	.78	.80	.81	.83	.85	.87	.91	.96	1.00

Notes: RD = Racial Diversity. All correlations are significant at $p < .05$. RD8 is the first time period (i.e., Period 8 in 1995), followed by RD9, etc.

Table 19. Means, standard deviations, and intercorrelations for age diversity across the 11 time periods (N = 2,388).

Variable	Mean	SD	AD8	AD9	AD10	AD11	AD12	AD13	AD1	AD2	AD3	AD4	AD5
AD8	25.54	12.58	1.00										
AD9	26.36	12.60	.79	1.00									
AD10	26.98	12.15	.63	.80	1.00								
AD11	27.45	12.36	.55	.67	.83	1.00							
AD12	27.44	11.76	.47	.57	.71	.86	1.00						
AD13	27.45	11.73	.40	.49	.59	.70	.82	1.00					
AD1	27.72	12.10	.37	.45	.54	.62	.71	.88	1.00				
AD2	27.94	12.30	.34	.44	.50	.58	.66	.78	.89	1.00			
AD3	27.58	11.81	.33	.42	.48	.55	.62	.71	.78	.88	1.00		
AD4	27.55	11.64	.32	.40	.46	.51	.58	.64	.70	.75	.86	1.00	
AD5	27.71	12.07	.31	.37	.43	.49	.55	.61	.66	.70	.79	.90	1.00

Notes: AD = Age Diversity (SD of age in days within a store/100). All correlations are significant at $p < .05$. AD8 is the first time period (i.e., Period 8 in 1995), followed by AD9, etc.

Table 20. Means, standard deviations, and intercorrelations for controllable profit across the 11 time periods (N = 2,388).

Variable	Mean	SD	CP8	CP9	CP10	CP11	CP12	CP13	CP1	CP2	CP3	CP4	CP5
CP8	$199.21	$90.41	1.00										
CP9	$206.54	$89.70	.85	1.00									
CP10	$178.91	$84.80	.80	.85	1.00								
CP11	$183.72	$84.42	.79	.84	.86	1.00							
CP12	$158.29	$81.36	.79	.84	.87	.86	1.00						
CP13	$165.78	$88.00	.73	.78	.82	.82	.86	1.00					
CP1	$136.59	$78.42	.75	.77	.83	.83	.84	.86	1.00				
CP2	$150.06	$79.41	.76	.79	.84	.84	.84	.82	.87	1.00			
CP3	$165.67	$81.21	.74	.79	.84	.83	.83	.80	.84	.88	1.00		
CP4	$167.40	$82.38	.78	.80	.83	.83	.84	.82	.85	.88	.88	1.00	
CP5	$179.29	$85.55	.77	.81	.83	.83	.83	.79	.81	.86	.88	.91	1.00

Notes: CP = Controllable Profit/100. All correlations are significant at $p < .05$. CP 8 is the first time period (i.e., Period 8 in 1995), followed by CP9, etc.

Table 21. Intercorrelations for controllable profit and racial diversity across the 11 time periods (N = 2,388).

Variable	CP8	CP9	CP10	CP11	CP12	CP13	CP1	CP2	CP3	CP4	CP5
RD8	.10	.08	.12	.10	.12	.13	.17	.17	.14	.16	.15
RD9	.10	.09	.13	.10	.12	.13	.17	.17	.15	.16	.15
RD10	.10	.09	.12	.10	.12	.13	.17	.17	.15	.16	.15
RD11	.10	.09	.12	.09	.11	.13	.17	.17	.15	.16	.15
RD12	.10	.09	.11	.09	.10	.12	.16	.16	.14	.14	.14
RD13	.10	.09	.11	.09	.10	.12	.16	.16	.14	.15	.14
RD1	.11	.09	.12	.09	.11	.12	.16	.17	.14	.16	.14
RD2	.10	.09	.11	.09	.11	.12	.16	.16	.14	.15	.14
RD3	.11	.09	.12	.09	.11	.12	.16	.16	.14	.15	.13
RD4	.10	.09	.11	.08	.11	.12	.16	.15	.13	.14	.13
RD5	.10	.09	.11	.09	.11	.12	.15	.16	.14	.14	.13

Notes: CP = Controllable Profit/100; RD = Racial Diversity. All correlations are significant at $p < .05$. CP8 (RD8) is the first time period (i.e., Period 8 in 1995), followed by CP9 (RD9), etc.

Table 22. Intercorrelations for controllable profit and age diversity across the 11 time periods (N = 2,388).

Variable	CP8	CP9	CP10	CP11	CP12	CP13	CP1	CP2	CP3	CP4	CP5
AD8	.03	.05	.06	.08	.06	.06	.05	.08	.09	.07	.09
AD9	.05	.05	.07	.07	.06	.06	.06	.09	.09	.07	.09
AD10	.06	.07	.07	.08	.07	.06	.06	.09	.10	.08	.10
AD11	.06	.06	.06	.07	.07	.05	.05	.08	.08	.06	.08
AD12	.04	.05	.06	.06	.06	.04	.03	.07	.07	.05	.07
AD13	.05	.05	.06	.06	.05	.04	.04	.06	.08	.06	.08
AD1	.07	.07	.08	.09	.08	.07	.07	.08	.09	.08	.09
AD2	.06	.05	.07	.07	.07	.06	.06	.08	.08	.06	.08
AD3	.05	.05	.06	.07	.06	.05	.05	.07	.07	.06	.07
AD4	.05	.06	.06	.06	.06	.03	.04	.06	.07	.06	.07
AD5	.07	.07	.08	.08	.07	.05	.07	.09	.09	.07	.08

Notes: CP = Controllable Profit/100; AD = Age Diversity (SD of age in days within a store/100). Correlations ≥ .04 are significant at $p < .05$. CP8 (AD8) is the first time period (i.e., Period 8 in 1995), followed by CP9 (AD9), etc.

Table 23. Example of the period-person data structure used for the multilevel survival analyses.

Store	Employee	Period	Turnover	Sex	Store Sex	Sex X Store Sex
1	1	1	0	1	.50	.50
1	1	2	0	1	.50	.50
1	1	3	1	1	.48	.48
1	2	1	0	0	.36	0
1	2	2	1	0	.35	0
1	3	1	0	1	.31	.31
1	3	2	0	1	.32	.32
:	:	:	:	:	:	:
:	:	:	:	:	:	:
1	3	35	0	1	.45	.45

Table 24. Results of HnLM examining the effect of individual sex X restaurant sex composition interaction on turnover probabilities.

Variable	Parameter(s)	B	SE	t-ratio	Exp(B)
Intercept	$\gamma_{00} + \varepsilon_{0k}$	-2.066	.021	-100.72*	.126
Period	γ_{10}	-0.035	.002	-18.10*	.965
Period^2	γ_{20}	0.001	.0001	10.10*	1.001
Female	$\gamma_{30} + \varepsilon_{3k}$	0.207	.02	9.84*	1.23
Female Composition	γ_{40}	.422	.04	11.52*	1.53
Female X Female Composition	γ_{50}	-.532	.04	-13.62*	.588

Notes: *p < .05. SE = Standard error.

Table 25. Results of HnLM examining the effect of individual sex X restaurant sex composition X time interaction on turnover probabilities.

Variable	Parameter(s)	B	SE	t-ratio	Exp(B)
Intercept	$\gamma_{00} + \varepsilon_{0k}$	-.216	.026	-84.39*	.115
Period	γ_{10}	-.0189	.003	-6.186*	.981
Period^2	γ_{20}	.0008	.00008	10.154*	1.001
Female	$\gamma_{30} + \varepsilon_{3k}$	.405	.0297	13.63*	1.820
Female Composition	γ_{40}	.599	.0487	12.303*	1.499
Female X Female Composition	γ_{50}	.405	.029	13.61*	.421
Female X Period	γ_{60}	-.865	.054	-15.95*	.969
Female Composition X Period	γ_{70}	-.03	.004	-6.172*	.970
Female X Female Composition X Period	γ_{80}	.053	.006	8.684*	1.055

Notes: *p < .05. SE = Standard error.

Table 26. Results of HnLM examining the effect of individual race X restaurant race composition interaction on turnover probabilities.

Variable	Parameter(s)	B	SE	t-ratio	Exp(B)
Intercept	$\gamma_{00} + \varepsilon_{0k}$	-1.846	0.022	-91.864	0.158
Period	γ_{10}	-0.043	0.003	-8.431	0.958
Period^2	γ_{20}	0.001	0.000	9.403	1.001
White	$\gamma_{30} + \varepsilon_{3k}$	0.254	0.041	10.217	1.289
Hispanic	$\gamma_{40} + \varepsilon_{4k}$	0.204	0.054	3.757	1.226
White Composition	γ_{50}	0.394	0.030	19.335	1.482
Hispanic Composition	γ_{60}	0.174	0.053	4.082	1.190
White X White Composition	γ_{70}	-0.678	0.048	-19.195	0.508
White X Hispanic Composition	γ_{80}	-0.362	0.063	-8.233	0.696
Hispanic X White Composition	γ_{90}	-0.606	0.081	-8.024	0.545
Hispanic X Hispanic Composition	γ_{100}	-0.373	0.073	-6.089	0.689

Notes: *p < .05. SE = Standard error.

Table 27. Predicted turnover hazards as a function of employee race and store racial composition.

	Store Racial Composition								
	80% White 10% Hispanic 10% Black			10% White 80% Hispanic 10% Black			10% White 10% Hispanic 80% Black		
	Employee Race			Employee Race			Employee Race		
P	White	Hisp.	Black	White	Hisp.	Black	White	Hisp.	Black
1	13.6%	13.6%	15.9%	15.9%	11.4%	15.5%	17.5%	14.6%	13.7%
2	13.1%	13.1%	15.4%	15.4%	11.0%	14.9%	16.9%	14.1%	13.2%
3	12.7%	12.7%	14.9%	14.9%	10.7%	14.5%	16.4%	13.7%	12.8%
4	12.3%	12.3%	14.5%	14.5%	10.3%	14.0%	15.9%	13.2%	12.4%
5	11.9%	12.0%	14.0%	14.0%	10.0%	13.6%	15.4%	12.9%	12.0%
6	11.6%	11.6%	13.7%	13.7%	9.7%	13.2%	15.0%	12.5%	11.7%
7	11.3%	11.3%	13.3%	13.3%	9.5%	12.9%	14.6%	12.2%	11.4%
8	11.0%	11.0%	13.0%	13.0%	9.2%	12.6%	14.3%	11.9%	11.1%
9	10.8%	10.8%	12.7%	12.7%	9.0%	12.3%	14.0%	11.6%	10.8%
10	10.6%	10.6%	12.4%	12.4%	8.8%	12.1%	13.7%	11.4%	10.6%
11	10.3%	10.3%	12.2%	12.2%	8.6%	11.8%	13.4%	11.1%	10.4%
12	10.2%	10.2%	12.0%	12.0%	8.5%	11.6%	13.2%	10.9%	10.2%
13	10.0%	10.0%	11.8%	11.8%	8.3%	11.4%	13.0%	10.8%	10.1%
14	9.9%	9.9%	11.6%	11.6%	8.2%	11.3%	12.8%	10.6%	9.9%
15	9.7%	9.7%	11.5%	11.5%	8.1%	11.1%	12.7%	10.5%	9.8%
16	9.6%	9.6%	11.4%	11.4%	8.0%	11.0%	12.5%	10.4%	9.7%
17	9.5%	9.5%	11.3%	11.3%	8.0%	10.9%	12.4%	10.3%	9.6%
18	9.5%	9.5%	11.2%	11.2%	7.9%	10.8%	12.3%	10.2%	9.5%
19	9.4%	9.4%	11.1%	11.1%	7.9%	10.8%	12.3%	10.2%	9.5%
20	9.4%	9.4%	11.1%	11.1%	7.8%	10.7%	12.2%	10.1%	9.4%
21	9.4%	9.4%	11.1%	11.1%	7.8%	10.7%	12.2%	10.1%	9.4%
22	9.4%	9.4%	11.1%	11.1%	7.8%	10.7%	12.2%	10.1%	9.4%
23	9.4%	9.4%	11.1%	11.1%	7.8%	10.7%	12.2%	10.1%	9.4%
24	9.4%	9.4%	11.1%	11.1%	7.9%	10.8%	12.3%	10.2%	9.5%
25	9.5%	9.5%	11.2%	11.2%	7.9%	10.8%	12.3%	10.2%	9.5%
26	9.5%	9.5%	11.3%	11.3%	8.0%	10.9%	12.4%	10.3%	9.6%
27	9.6%	9.6%	11.4%	11.4%	8.0%	11.0%	12.5%	10.4%	9.7%
28	9.7%	9.7%	11.5%	11.5%	8.1%	11.1%	12.7%	10.5%	9.8%
29	9.9%	9.9%	11.6%	11.6%	8.2%	11.3%	12.8%	10.6%	9.9%
30	10.0%	10.0%	11.8%	11.8%	8.3%	11.4%	13.0%	10.8%	10.1%

Table 27 (cont'd)

31	10.2%	10.2%	12.0%	12.0%	8.5%	11.6%	13.2%	10.9%	10.2%
32	10.3%	10.3%	12.2%	12.2%	8.6%	11.8%	13.4%	11.1%	10.4%
33	10.6%	10.6%	12.4%	12.4%	8.8%	12.1%	13.7%	11.4%	10.6%
34	10.8%	10.8%	12.7%	12.7%	9.0%	12.3%	14.0%	11.6%	10.8%
35	11.0%	11.0%	13.0%	13.0%	9.2%	12.6%	14.3%	11.9%	11.1%

Notes: P = Period.

Table 28. Results of HnLM examining the effect of individual race X restaurant race composition X time interaction on turnover probabilities.

Variable	Parameter(s)	B	SE	t-ratio	Exp(B)
Intercept	$\gamma_{00} + \varepsilon_{0k}$	-2.019*	.022	-91.864	.133
Period	γ_{10}	-.021*	.003	-8.431	.979
Period^2	γ_{20}	.001*	.000	9.403	1.001
White	$\gamma_{30} + \varepsilon_{3k}$	.579*	.030	19.335	1.784
Hispanic	$\gamma_{40} + \varepsilon_{4k}$	.216*	.053	4.082	1.241
White Composition	γ_{50}	.414*	.041	10.217	1.513
Hispanic Composition	γ_{60}	.202*	.054	3.757	1.224
White X White Composition	γ_{70}	-.926*	.048	-19.195	.396
White X Hispanic Composition	γ_{80}	-.516*	.063	-8.233	.597
Hispanic X White Composition	γ_{90}	-.446*	.073	-6.089	.640
Hispanic X Hispanic Composition	γ_{100}	-.651*	.081	-8.024	.521
White X Period	γ_{110}	-.030*	.003	-9.193	.970
Hispanic X Period	γ_{120}	-.006	.005	-1.215	.994
White Composition X Period	γ_{130}	-.026*	.004	-6.677	.975
Hispanic Composition X Period	γ_{140}	-.001	.005	-.105	.999
White X White Composition X Period	γ_{150}	.041*	.005	7.830	1.041
White X Hispanic Composition X Period	γ_{150}	.025*	.006	3.929	1.026
Hispanic X White Composition X Period	γ_{160}	.012*	.007	1.666	1.012
Hispanic X Hispanic Composition X Period	γ_{170}	.006	.008	.739	1.006

Notes: *p < .05. SE = Standard error.

Table 29. Predicted turnover hazards as a function of employee race and store racial composition and time.

	Store Racial Composition								
	80% White 10% Hispanic 10% Black			10% White 80% Hispanic 10% Black			10% White 10% Hispanic 80% Black		
	Employee Race			Employee Race			Employee Race		
P	White	Hisp.	Black	White	Hisp.	Black	White	Hisp.	Black
1	12.8%	12.8%	15.3%	14.5%	10.1%	13.7%	17.2%	13.3%	12.1%
2	12.4%	12.4%	14.8%	14.2%	9.9%	13.5%	16.6%	13.1%	11.9%
3	12.1%	12.0%	14.3%	13.9%	9.7%	13.2%	16.0%	12.8%	11.7%
4	11.7%	11.6%	13.9%	13.6%	9.5%	13.0%	15.4%	12.6%	11.5%
5	11.4%	11.2%	13.5%	13.3%	9.4%	12.8%	14.9%	12.4%	11.3%
6	11.1%	10.9%	13.1%	13.1%	9.2%	12.6%	14.4%	12.2%	11.1%
7	10.8%	10.6%	12.7%	12.8%	9.1%	12.5%	13.9%	12.1%	11.0%
8	10.6%	10.3%	12.4%	12.6%	9.0%	12.3%	13.5%	11.9%	10.9%
9	10.4%	10.0%	12.0%	12.4%	8.9%	12.2%	13.1%	11.8%	10.8%
10	10.1%	9.8%	11.7%	12.2%	8.8%	12.1%	12.7%	11.7%	10.6%
11	9.9%	9.5%	11.5%	12.1%	8.7%	12.0%	12.4%	11.6%	10.6%
12	9.7%	9.3%	11.2%	11.9%	8.7%	11.9%	12.0%	11.5%	10.5%
13	9.6%	9.1%	10.9%	11.8%	8.6%	11.8%	11.7%	11.5%	10.4%
14	9.4%	8.9%	10.7%	11.7%	8.6%	11.8%	11.4%	11.4%	10.4%
15	9.3%	8.7%	10.5%	11.6%	8.6%	11.7%	11.1%	11.4%	10.3%
16	9.1%	8.6%	10.3%	11.5%	8.5%	11.7%	10.9%	11.3%	10.3%
17	9.0%	8.4%	10.1%	11.5%	8.5%	11.7%	10.7%	11.3%	10.3%
18	8.9%	8.3%	10.0%	11.4%	8.5%	11.7%	10.4%	11.3%	10.3%
19	8.8%	8.2%	9.8%	11.4%	8.6%	11.7%	10.2%	11.4%	10.3%
20	8.8%	8.1%	9.7%	11.3%	8.6%	11.7%	10.0%	11.4%	10.4%
21	8.7%	8.0%	9.6%	11.3%	8.6%	11.8%	9.9%	11.4%	10.4%
22	8.6%	7.9%	9.5%	11.3%	8.7%	11.9%	9.7%	11.5%	10.5%
23	8.6%	7.8%	9.4%	11.4%	8.7%	11.9%	9.6%	11.6%	10.5%
24	8.5%	7.7%	9.3%	11.4%	8.8%	12.0%	9.5%	11.7%	10.6%
25	8.5%	7.7%	9.3%	11.4%	8.9%	12.1%	9.3%	11.8%	10.7%
26	8.5%	7.6%	9.2%	11.5%	9.0%	12.3%	9.2%	11.9%	10.8%
27	8.5%	7.6%	9.2%	11.6%	9.1%	12.4%	9.2%	12.0%	10.9%
28	8.5%	7.6%	9.1%	11.6%	9.2%	12.5%	9.1%	12.2%	11.1%

Table 29 (cont'd)

29	8.5%	7.5%	9.1%	11.8%	9.3%	12.7%	9.0%	12.3%	11.2%
30	8.6%	7.5%	9.1%	11.9%	9.5%	12.9%	9.0%	12.5%	11.4%
31	8.6%	7.5%	9.1%	12.0%	9.6%	13.1%	8.9%	12.7%	11.6%
32	8.7%	7.5%	9.1%	12.2%	9.8%	13.3%	8.9%	13.0%	11.8%
33	8.7%	7.6%	9.1%	12.3%	10.0%	13.6%	8.9%	13.2%	12.0%
34	8.8%	7.6%	9.2%	12.5%	10.2%	13.9%	8.9%	13.5%	12.3%
35	8.9%	7.6%	9.2%	12.7%	10.4%	14.2%	8.9%	13.7%	12.5%

Notes: P = Period.

Table 30. Results of HnLM examining the effect of individual age X restaurant age composition interaction on turnover probabilities.

Variable	Parameter(s)	B	SE	t-ratio	Exp(B)
Intercept	$\gamma_{00} + \varepsilon_{0k}$	-2.960	.098	-30.182	.052
Period	γ_{10}	-.034	.002	-17.788	.966
Period^2	γ_{20}	.001	.000	9.796	1.001
Age	$\gamma_{30} + \varepsilon_{3k}$	.053	.004	12.641	1.054
Age Composition	γ_{40}	.043	.004	9.869	1.044
Age X Age Composition	γ_{50}	-.002	.000	-11.834	.998

Notes: *p < .05. SE = Standard error.

Table 31. Predicted turnover hazards as a function of employee age X store age composition interaction.

	Store Age Composition (Average Age)								
	16			19			31		
	Employee Age								
	16	19	31	16	19	31	16	19	31
1	11.9%	12.5%	15.1%	12.2%	12.6%	14.3%	13.4%	12.9%	11.2%
2	11.5%	12.1%	14.7%	11.8%	12.2%	13.9%	13.0%	12.6%	10.9%
3	11.2%	11.8%	14.4%	11.5%	11.9%	13.5%	12.7%	12.2%	10.6%
4	10.9%	11.5%	14.0%	11.2%	11.6%	13.2%	12.3%	11.9%	10.4%
5	10.7%	11.2%	13.7%	10.9%	11.3%	12.9%	12.0%	11.6%	10.1%
6	10.4%	11.0%	13.4%	10.7%	11.0%	12.6%	11.8%	11.4%	9.9%
7	10.2%	10.7%	13.1%	10.4%	10.8%	12.3%	11.5%	11.1%	9.6%
8	10.0%	10.5%	12.8%	10.2%	10.6%	12.1%	11.3%	10.9%	9.4%
9	9.8%	10.3%	12.6%	10.0%	10.4%	11.8%	11.0%	10.7%	9.3%
10	9.6%	10.1%	12.3%	9.8%	10.2%	11.6%	10.8%	10.5%	9.1%
11	9.4%	9.9%	12.1%	9.7%	10.0%	11.4%	10.7%	10.3%	8.9%
12	9.3%	9.8%	11.9%	9.5%	9.8%	11.2%	10.5%	10.1%	8.8%
13	9.2%	9.6%	11.8%	9.4%	9.7%	11.1%	10.3%	10.0%	8.7%
14	9.0%	9.5%	11.6%	9.3%	9.6%	10.9%	10.2%	9.9%	8.5%
15	8.9%	9.4%	11.5%	9.1%	9.5%	10.8%	10.1%	9.7%	8.4%
16	8.8%	9.3%	11.4%	9.1%	9.4%	10.7%	10.0%	9.6%	8.3%
17	8.7%	9.2%	11.3%	9.0%	9.3%	10.6%	9.9%	9.5%	8.3%
18	8.7%	9.1%	11.2%	8.9%	9.2%	10.5%	9.8%	9.5%	8.2%
19	8.6%	9.1%	11.1%	8.8%	9.1%	10.5%	9.7%	9.4%	8.1%
20	8.6%	9.0%	11.1%	8.8%	9.1%	10.4%	9.7%	9.4%	8.1%
21	8.5%	9.0%	11.0%	8.8%	9.1%	10.4%	9.7%	9.3%	8.1%
22	8.5%	9.0%	11.0%	8.7%	9.0%	10.3%	9.6%	9.3%	8.0%
23	8.5%	9.0%	11.0%	8.7%	9.0%	10.3%	9.6%	9.3%	8.0%
24	8.5%	9.0%	11.0%	8.7%	9.0%	10.3%	9.6%	9.3%	8.0%
25	8.5%	9.0%	11.0%	8.7%	9.0%	10.3%	9.6%	9.3%	8.0%
26	8.5%	9.0%	11.0%	8.8%	9.1%	10.4%	9.7%	9.3%	8.1%
27	8.6%	9.0%	11.1%	8.8%	9.1%	10.4%	9.7%	9.4%	8.1%
28	8.6%	9.1%	11.1%	8.8%	9.1%	10.5%	9.8%	9.4%	8.1%
29	8.7%	9.1%	11.2%	8.9%	9.2%	10.5%	9.8%	9.5%	8.2%

Table 31 (cont'd)

30	8.8%	9.2%	11.3%	9.0%	9.3%	10.6%	9.9%	9.5%	8.3%
31	8.8%	9.3%	11.4%	9.1%	9.4%	10.7%	10.0%	9.6%	8.3%
32	8.9%	9.4%	11.5%	9.2%	9.5%	10.8%	10.1%	9.7%	8.4%
33	9.0%	9.5%	11.6%	9.3%	9.6%	10.9%	10.2%	9.9%	8.5%
34	9.2%	9.6%	11.8%	9.4%	9.7%	11.1%	10.3%	10.0%	8.7%
35	9.3%	9.8%	12.0%	9.5%	9.9%	11.3%	10.5%	10.1%	8.8%

Notes: P = Period.

Table 32. Results of HnLM examining the effect of individual age X restaurant age composition X time interaction on turnover probabilities.

Variable	Parameter(s)	B	SE	t-ratio	Exp(B)
Intercept	$\gamma_{00} + \varepsilon_{0k}$	-3.63	.152	-23.817	0.0265
Period	γ_{10}	.071	.012	5.893	1.0731
Period^2	γ_{20}	.00076	.000075	10.124	1.0008
Age	$\gamma_{30} + \varepsilon_{3k}$	.0870	.00030	-11.206	1.0908
Age Composition	γ_{40}	.066	.0068	12.847	1.0682
Age X Age Composition	γ_{50}	-.0033	.0069	9.623	0.9967
Age X Period	γ_{60}	-.005	.0005	-9.921	0.9949
Age Composition X Period	γ_{70}	-.0037	.0005	-6.952	0.9963
Age X Age Composition X Period	γ_{80}	.00018	.00002	8.085	1.0002

Notes: *p < .05. SE = Standard error.

Table 33. Predicted turnover hazards as a function of employee age X store age composition X time interaction.

	Store Age Composition (Average Age)								
	16			19			31		
	Employee Age								
	16	19	31	16	19	31	16	19	31
1	11.3%	12.3%	17.1%	11.7%	12.4%	15.6%	13.3%	12.8%	10.8%
2	11.1%	12.0%	16.3%	11.5%	12.1%	15.0%	12.9%	12.4%	10.6%
3	10.9%	11.7%	15.6%	11.2%	11.8%	14.4%	12.5%	12.1%	10.3%
4	10.7%	11.5%	14.9%	11.0%	11.5%	13.8%	12.2%	11.8%	10.1%
5	10.6%	11.2%	14.2%	10.8%	11.3%	13.3%	11.9%	11.5%	9.9%
6	10.4%	11.0%	13.6%	10.6%	11.0%	12.8%	11.6%	11.2%	9.7%
7	10.3%	10.8%	13.1%	10.5%	10.8%	12.3%	11.3%	11.0%	9.5%
8	10.2%	10.6%	12.6%	10.3%	10.6%	11.9%	11.1%	10.7%	9.4%
9	10.1%	10.4%	12.1%	10.2%	10.5%	11.5%	10.9%	10.5%	9.2%
10	10.0%	10.3%	11.6%	10.1%	10.3%	11.1%	10.6%	10.3%	9.1%
11	9.9%	10.2%	11.2%	10.0%	10.2%	10.7%	10.5%	10.1%	9.0%
12	9.8%	10.0%	10.8%	9.9%	10.0%	10.4%	10.3%	10.0%	8.9%
13	9.8%	9.9%	10.4%	9.9%	9.9%	10.1%	10.1%	9.8%	8.8%
14	9.8%	9.8%	10.1%	9.8%	9.8%	9.8%	10.0%	9.7%	8.7%
15	9.8%	9.8%	9.8%	9.8%	9.7%	9.5%	9.9%	9.6%	8.7%
16	9.7%	9.7%	9.5%	9.7%	9.7%	9.3%	9.7%	9.5%	8.6%
17	9.8%	9.6%	9.2%	9.7%	9.6%	9.1%	9.6%	9.4%	8.6%
18	9.8%	9.6%	8.9%	9.7%	9.6%	8.9%	9.6%	9.4%	8.6%
19	9.8%	9.6%	8.7%	9.7%	9.5%	8.7%	9.5%	9.3%	8.6%
20	9.9%	9.6%	8.5%	9.8%	9.5%	8.5%	9.4%	9.2%	8.5%
21	9.9%	9.6%	8.3%	9.8%	9.5%	8.3%	9.4%	9.2%	8.6%
22	10.0%	9.6%	8.1%	9.9%	9.5%	8.2%	9.4%	9.2%	8.6%
23	10.1%	9.6%	7.9%	9.9%	9.5%	8.0%	9.3%	9.2%	8.6%
24	10.2%	9.7%	7.7%	10.0%	9.6%	7.9%	9.3%	9.2%	8.7%
25	10.3%	9.7%	7.6%	10.1%	9.6%	7.8%	9.3%	9.2%	8.7%
26	10.5%	9.8%	7.5%	10.2%	9.7%	7.7%	9.4%	9.2%	8.8%
27	10.6%	9.9%	7.3%	10.4%	9.8%	7.6%	9.4%	9.3%	8.9%
28	10.8%	10.0%	7.2%	10.5%	9.8%	7.6%	9.4%	9.4%	9.0%

Table 33 (cont'd)

29	11.0%	10.1%	7.1%	10.7%	9.9%	7.5%	9.5%	9.4%	9.1%
30	11.2%	10.2%	7.1%	10.8%	10.1%	7.4%	9.6%	9.5%	9.2%
31	11.4%	10.3%	7.0%	11.0%	10.2%	7.4%	9.7%	9.6%	9.4%
32	11.6%	10.5%	6.9%	11.2%	10.3%	7.4%	9.8%	9.7%	9.5%
33	11.9%	10.7%	6.9%	11.5%	10.5%	7.4%	9.9%	9.8%	9.7%
34	12.2%	10.9%	6.8%	11.7%	10.7%	7.3%	10.0%	10.0%	9.9%
35	12.5%	11.1%	6.8%	12.0%	10.9%	7.3%	10.2%	10.2%	10.1%

Notes: P = Period.

Table 34. Results of HnLM examining the effect of individual age X restaurant age composition X extraversion interaction on turnover probabilities.

Variable	Parameter(s)	B	SE	t-ratio	Exp(B)
Intercept	$\gamma_{00} + \varepsilon_{0k}$	-1.916*	.065	-29.659	.147
Period	γ_{10}	-.051*	.003	-14.774	.950
Period^2	γ_{20}	.002*	.000	10.370	1.002
Age	$\gamma_{30} + \varepsilon_{3k}$	.050*	.018	2.751	1.051
Age Composition	γ_{40}	.026	.017	1.531	1.027
Age X Age Composition	γ_{50}	-.001	.001	-1.446	.999
Extraversion	$\gamma_{60} + \varepsilon_{6k}$	.093	.152	.613	1.098
Extraversion X Age	γ_{70}	-.010	.007	-1.477	.990
Extraversion X Age Composition	γ_{80}	-.005	.006	-.810	.995
Extraversion X Age X Age Composition	γ_{90}	.000	.000	1.200	1.000

Notes: *p < .05. SE = Standard error.

Table 35. Results of HnLM examining the effect of individual race X restaurant race composition X extraversion interaction on turnover probabilities.

Variable	Parameter(s)	B	SE	t-ratio	Exp(B)
Intercept	$\gamma_{00} + \varepsilon_{0k}$	-1.566	.137	-11.462*	.209
Period	γ_{10}	-.053	.003	-15.937*	.948
Period^2	γ_{20}	.002	.000	11.248*	1.002
White	$\gamma_{30} + \varepsilon_{3k}$	-.574	.291	-1.970*	.563
Hispanic	$\gamma_{40} + \varepsilon_{4k}$	-1.604	.523	-3.068*	.201
White Composition	γ_{50}	-.559	.322	-1.736*	.572
Hispanic Composition	γ_{60}	-.579	.532	-1.089	.561
White X White Composition	γ_{70}	.353	.441	.801	1.423
White X Hispanic Composition	γ_{80}	.236	.696	.339	1.266
Hispanic X White Composition	γ_{90}	2.090	.744	2.808*	8.082
Hispanic X Hispanic Composition	γ_{100}	1.282	.804	1.594	3.603
Extraversion	$\gamma_{110} + \varepsilon_{11k}$	-.051	.034	-1.500	.950
Extraversion X White	$\gamma_{120} + \varepsilon_{120k}$	.241	.074	3.234*	1.272
Extraversion X Hispanic	$\gamma_{130} + \varepsilon_{130k}$	.443	.131	3.380*	1.557
Extraversion X White Composition	γ_{140}	.205	.081	2.541*	1.228
Extraversion X Hispanic Composition	γ_{150}	.168	.132	1.275	1.183
Extraversion X White X White Composition	γ_{150}	-.238	.112	-2.134*	.788
Extraversion X Hispanic X White Composition	γ_{160}	-.610	.187	-3.264*	.544
Extraversion X White X Hispanic Composition	γ_{170}	-.154	.175	-.884	.857
Extraversion X Hispanic X Hispanic Composition	γ_{180}	-.427	.201	-2.124*	.652

Notes: *p < .05. SE = Standard error.

Table 36. Results of HnLM examining the effect of individual sex X restaurant sex composition X extraversion interaction on turnover probabilities.

Variable	Parameter(s)	B	SE	t-ratio	Exp(B)
Intercept	$\gamma_{00} + \varepsilon_{0k}$	-2.367	.224	-10.564*	.094
Period	γ_{10}	-.055	.003	-15.965*	.946
Period^2	γ_{20}	.002	.000	11.219*	1.002
Female	$\gamma_{30} + \varepsilon_{3k}$	.287	.341	.841	1.332
Female Composition	γ_{40}	.371	.447	.829	1.449
Female X Female Composition	γ_{50}	-.490	.619	-.791	.613
Extraversion	$\gamma_{60} + \varepsilon_{6k}$	.155	.057	2.705*	1.168
Female X Extraversion	γ_{70}	-.067	.086	-.779	.935
Female Composition X Extraversion	γ_{80}	-.053	.114	-.466	.948
Female X Female Composition X Extraversion	γ_{90}	.092	.157	.584	1.096

Notes: *p < .05. SE = Standard error.

Table 37. Predicted turnover hazards as a function of employee age X store age composition X extraversion interaction.

	Store Racial Compositions							
	10% White 80% Hispanic				80% White 10% Hispanic			
	Employee Race							
	Hispanic		Black		Hispanic		Black	
	EX=3	EX=5	EX=3	EX=5	EX=3	EX=5	EX=3	EX=5
1	11.3%	12.3%	17.1%	11.7%	12.4%	15.6%	13.3%	12.8%
2	11.1%	12.0%	16.3%	11.5%	12.1%	15.0%	12.9%	12.4%
3	10.9%	11.7%	15.6%	11.2%	11.8%	14.4%	12.5%	12.1%
4	10.7%	11.5%	14.9%	11.0%	11.5%	13.8%	12.2%	11.8%
5	10.6%	11.2%	14.2%	10.8%	11.3%	13.3%	11.9%	11.5%
6	10.4%	11.0%	13.6%	10.6%	11.0%	12.8%	11.6%	11.2%
7	10.3%	10.8%	13.1%	10.5%	10.8%	12.3%	11.3%	11.0%
8	10.2%	10.6%	12.6%	10.3%	10.6%	11.9%	11.1%	10.7%
9	10.1%	10.4%	12.1%	10.2%	10.5%	11.5%	10.9%	10.5%
10	10.0%	10.3%	11.6%	10.1%	10.3%	11.1%	10.6%	10.3%
11	9.9%	10.2%	11.2%	10.0%	10.2%	10.7%	10.5%	10.1%
12	9.8%	10.0%	10.8%	9.9%	10.0%	10.4%	10.3%	10.0%
13	9.8%	9.9%	10.4%	9.9%	9.9%	10.1%	10.1%	9.8%
14	9.8%	9.8%	10.1%	9.8%	9.8%	9.8%	10.0%	9.7%
15	9.8%	9.8%	9.8%	9.8%	9.7%	9.5%	9.9%	9.6%
16	9.7%	9.7%	9.5%	9.7%	9.7%	9.3%	9.7%	9.5%
17	9.8%	9.6%	9.2%	9.7%	9.6%	9.1%	9.6%	9.4%
18	9.8%	9.6%	8.9%	9.7%	9.6%	8.9%	9.6%	9.4%
19	9.8%	9.6%	8.7%	9.7%	9.5%	8.7%	9.5%	9.3%
20	9.9%	9.6%	8.5%	9.8%	9.5%	8.5%	9.4%	9.2%
21	9.9%	9.6%	8.3%	9.8%	9.5%	8.3%	9.4%	9.2%
22	10.0%	9.6%	8.1%	9.9%	9.5%	8.2%	9.4%	9.2%
23	10.1%	9.6%	7.9%	9.9%	9.5%	8.0%	9.3%	9.2%
24	10.2%	9.7%	7.7%	10.0%	9.6%	7.9%	9.3%	9.2%
25	10.3%	9.7%	7.6%	10.1%	9.6%	7.8%	9.3%	9.2%
26	10.5%	9.8%	7.5%	10.2%	9.7%	7.7%	9.4%	9.2%
27	10.6%	9.9%	7.3%	10.4%	9.8%	7.6%	9.4%	9.3%

Table 37 (cont'd)

28	10.8%	10.0%	7.2%	10.5%	9.8%	7.6%	9.4%	9.4%
29	11.0%	10.1%	7.1%	10.7%	9.9%	7.5%	9.5%	9.4%
30	11.2%	10.2%	7.1%	10.8%	10.1%	7.4%	9.6%	9.5%
31	11.4%	10.3%	7.0%	11.0%	10.2%	7.4%	9.7%	9.6%
32	11.6%	10.5%	6.9%	11.2%	10.3%	7.4%	9.8%	9.7%
33	11.9%	10.7%	6.9%	11.5%	10.5%	7.4%	9.9%	9.8%
34	12.2%	10.9%	6.8%	11.7%	10.7%	7.3%	10.0%	10.0%
35	12.5%	11.1%	6.8%	12.0%	10.9%	7.3%	10.2%	10.2%

Notes: P = Period

Table 38. Summary of Hypotheses and Results

Type of Hypothesis	Hypothesis	Support Found?
Individual Level	H8. Conscientiousness will be negatively related to VT.	No
Individual Level	H9. Conscientiousness will be negatively related to IVT.	No
Individual Level	H10. High ability employees will be more likely to exhibit VT as opposed to IVT after controlling for Conscientiousness.	No
Team Level	H1: The team mean ability score will be positively related to team performance.	Yes
Team Level	H2. The team mean conscientiousness score will be positively related to team performance.	No
Team Level	H3. The team mean conscientiousness score will be positively related to team performance after controlling for the effects of team cognitive ability.	No
Team Level	H4. Mean level of team Agreeableness will relate to team performance	No
Team Level	H5. Mean level of team agreeableness will relate to team performance after controlling for cognitive ability.	No
Team Level	H6. Mean levels of team extraversion will exhibit a nonlinear (inverted U-shaped) relationship with team performance.	No
Team Level	H7. Team manager tenure will moderate the relationship between mean team KSAO levels and team performance such that there is a stronger relationship between KSAO levels and performance for teams with high-tenure managers than with those with low tenure managers.	No
Team Level	H11. Increased team level diversity in conscientiousness should be associated with lower team performance.	No

Table 38 (cont'd)

Team Level	H12. Increased diversity in conscientiousness should be associated in higher team voluntary turnover (but not involuntary turnover).	No
Team Level	H15. Increased team racial diversity will be related to increased team turnover rates. This effect will be evident after controlling for the relevant personality constructs and ability.	No
Team Level	H16. Increased racial diversity will be negatively related to team performance. This effect will be evident after controlling for the relevant personality constructs and ability.	Yes
Team Level	H18. Increased team age diversity will be related to increased team turnover rates. This effect will be evident after controlling for the relevant personality constructs and ability.	No
Team Level	H19. Increased age diversity will be negatively related to team performance. This effect will be evident after controlling for the relevant personality constructs and ability.	No
Team Level	H21 Increased team sex diversity will be related to increased team turnover rates	No
Team Level	H22. Increased sex diversity will be negatively related to team performance	No
Cross-Level	H13. Increased conscientiousness-based person-team fit should be associated with a lower likelihood of individual level voluntary turnover (but not involuntary turnover).	No
Cross-Level	H14. The racial fit of the individual within the team will predict individual turnover such that misfit is associated with a higher likelihood of turnover. This relationship will be more evident in Whites working in teams with higher minority representations rather than vice-versa.	Yes

Table 38 (cont'd)

Cross-Level	H17. The age-based fit of the individual within the team will predict individual turnover such that misfit is associated with a higher likelihood of turnover.	Yes
Cross-Level	H20. The sex-based fit of the individual within the team will predict individual turnover such that misfit is associated with a higher likelihood of turnover.	Yes
Cross-Level	H23. The fit between an individual's demographics and that of the team will be more strongly related to turnover at lower levels of tenure.	Yes
Cross-Level	H24. The effects of demographic fit within a team will depend on an individual's extraversion such that high extraversion individuals that do not demographically fit will be less likely to turnover than low extraversion individuals who do not fit.	No

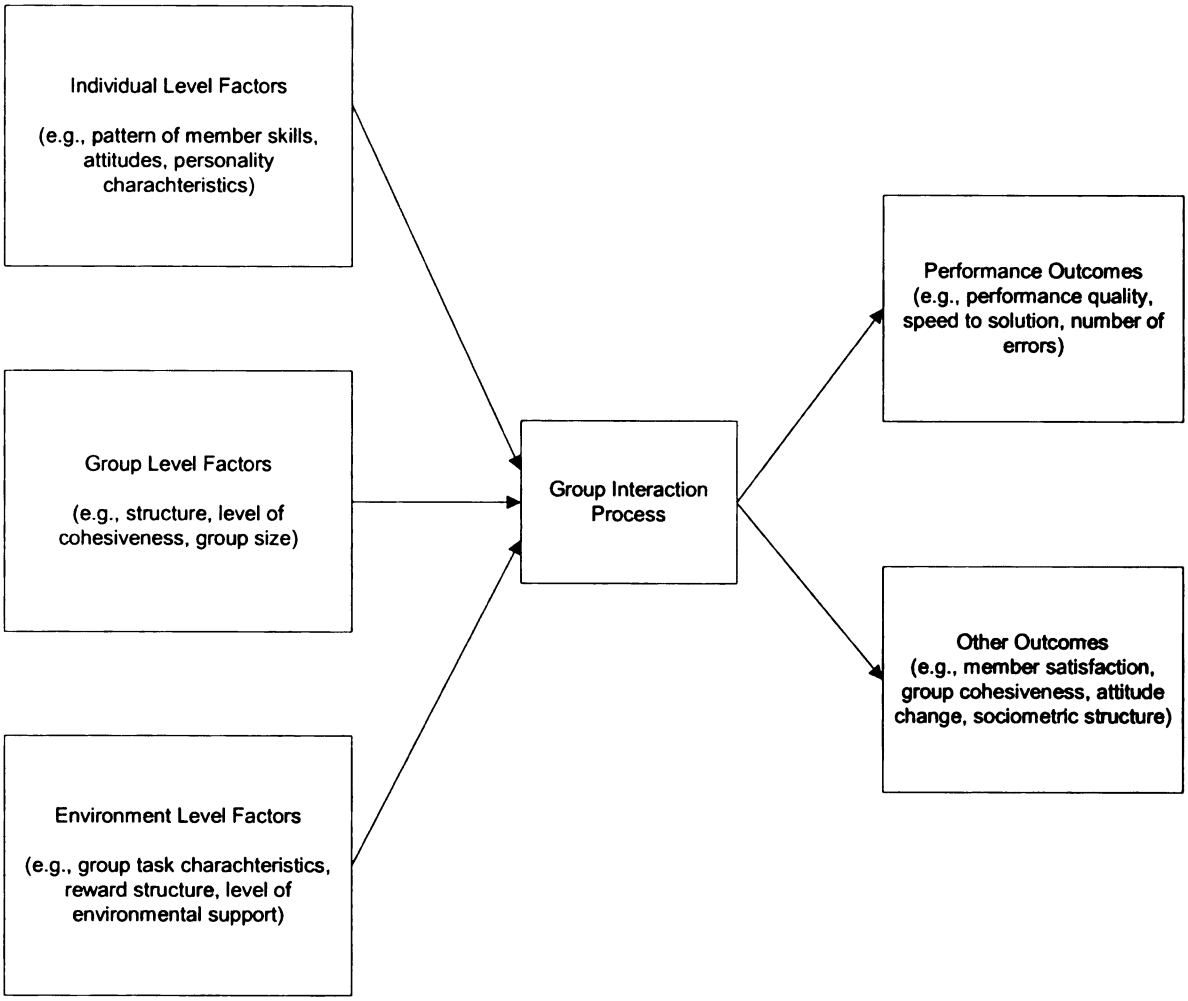

Figure 1. Hackman's IPO approach to teams

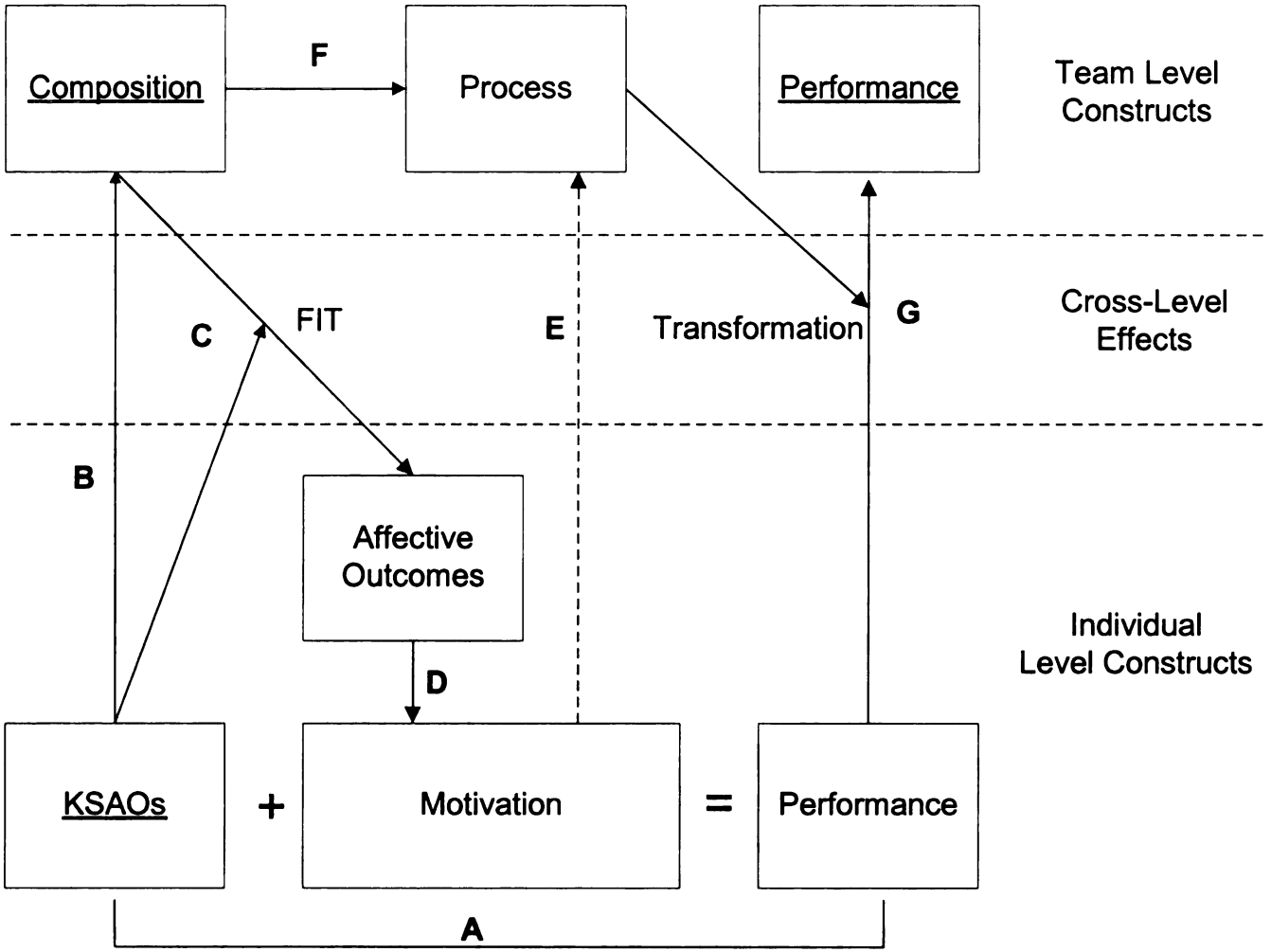

Figure 2. Theoretical model underlying the prediction of team performance.

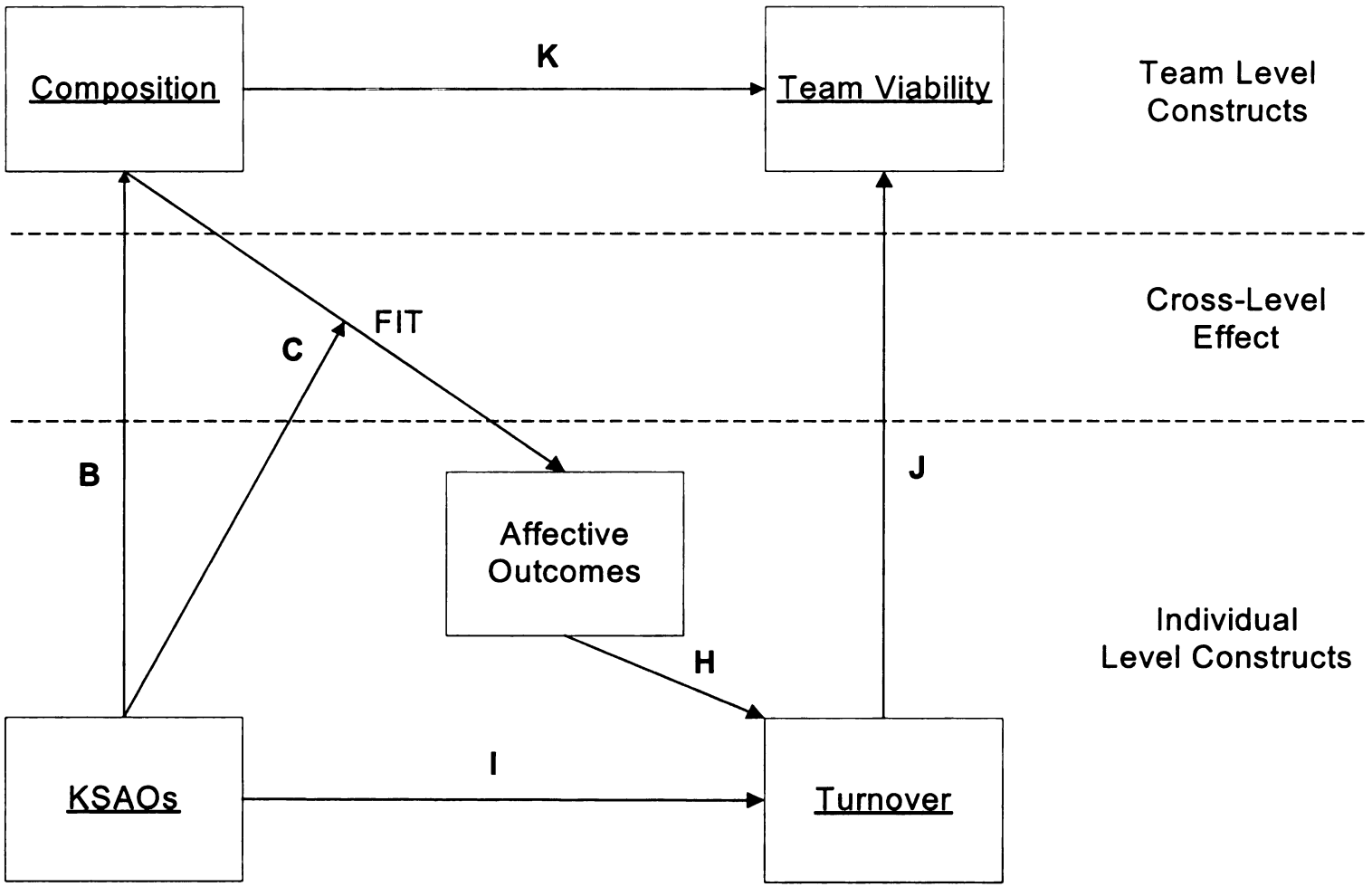

Figure 3. Theoretical model underlying the prediction of turnover.

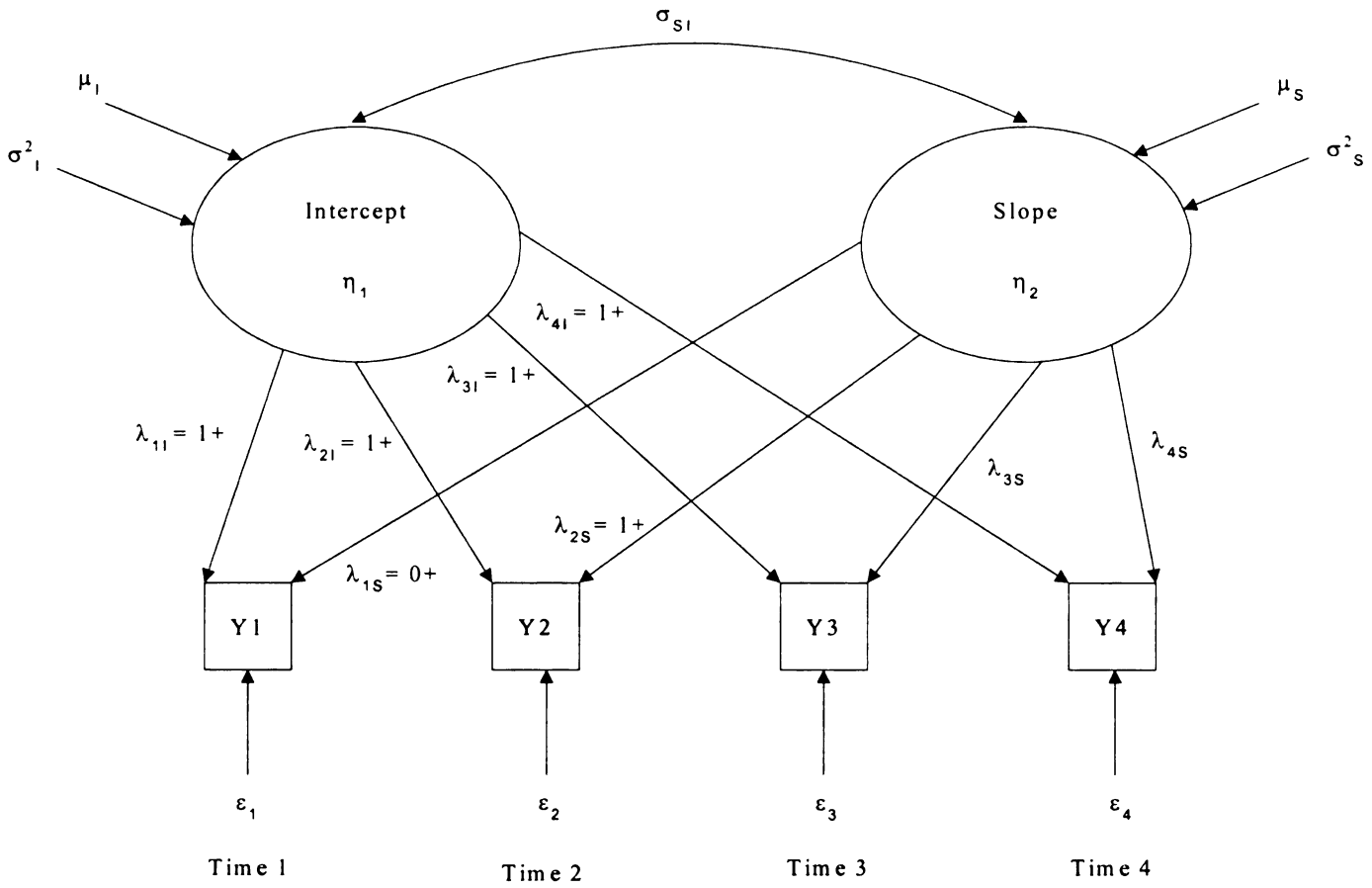

Figure 4. Representation of a simple latent growth curve structural equation model.

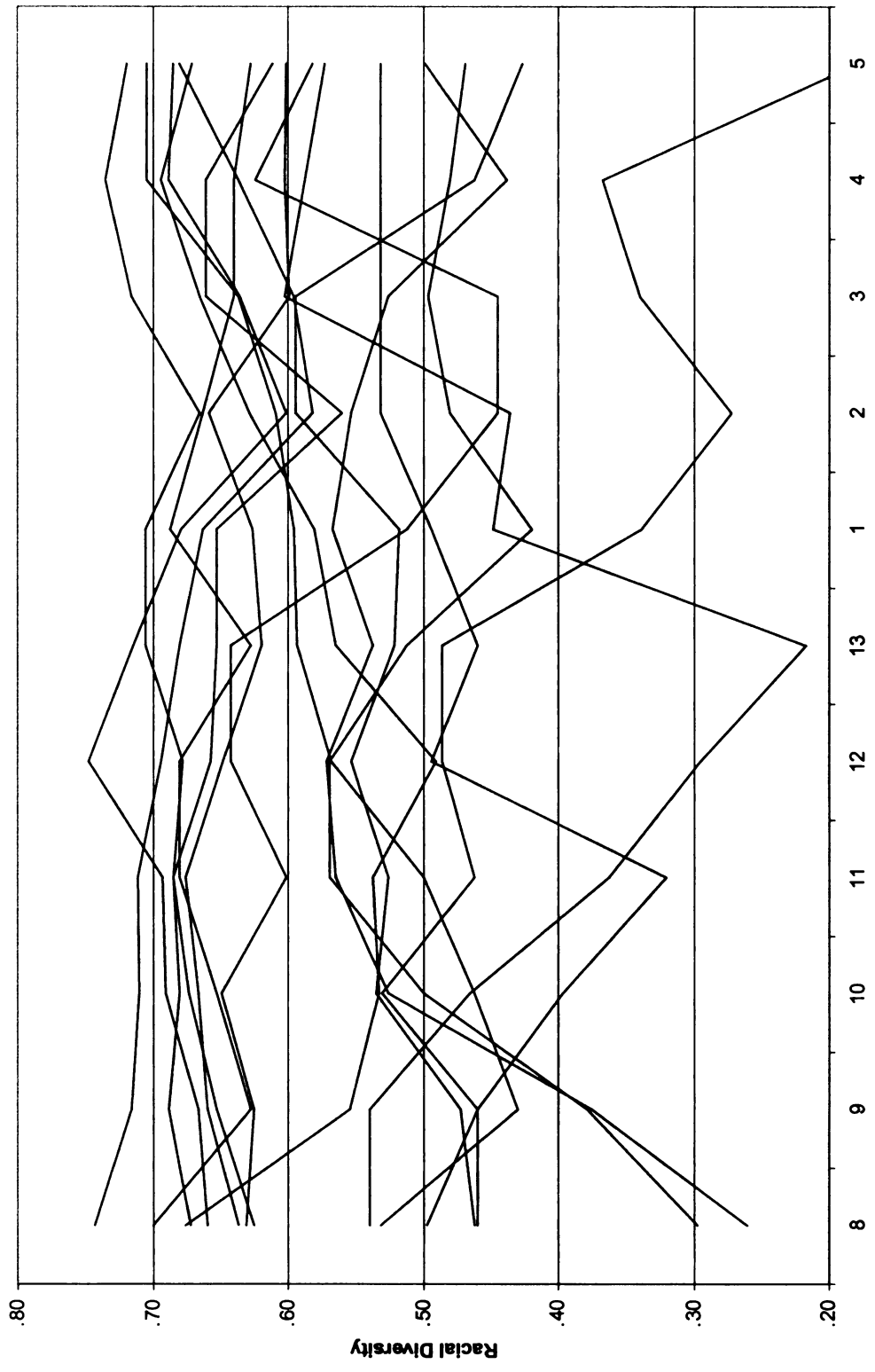

Figure 5. Trends in racial diversity for a sample of 15 restaurants.

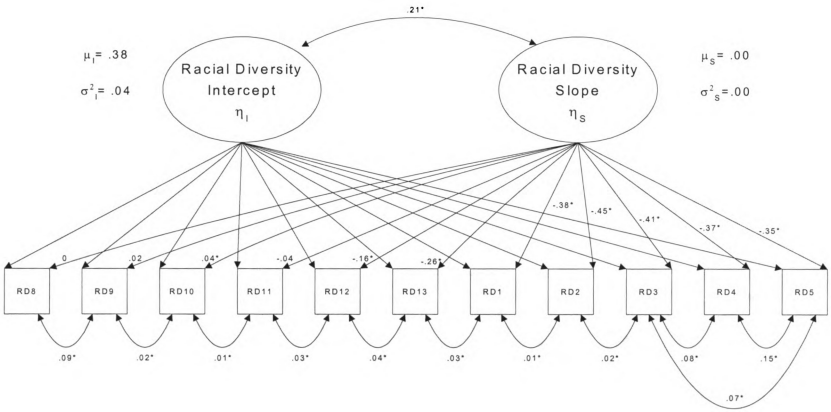

Figure 6 – Latent growth curve model for racial diversity. N = 2388 restaurants.

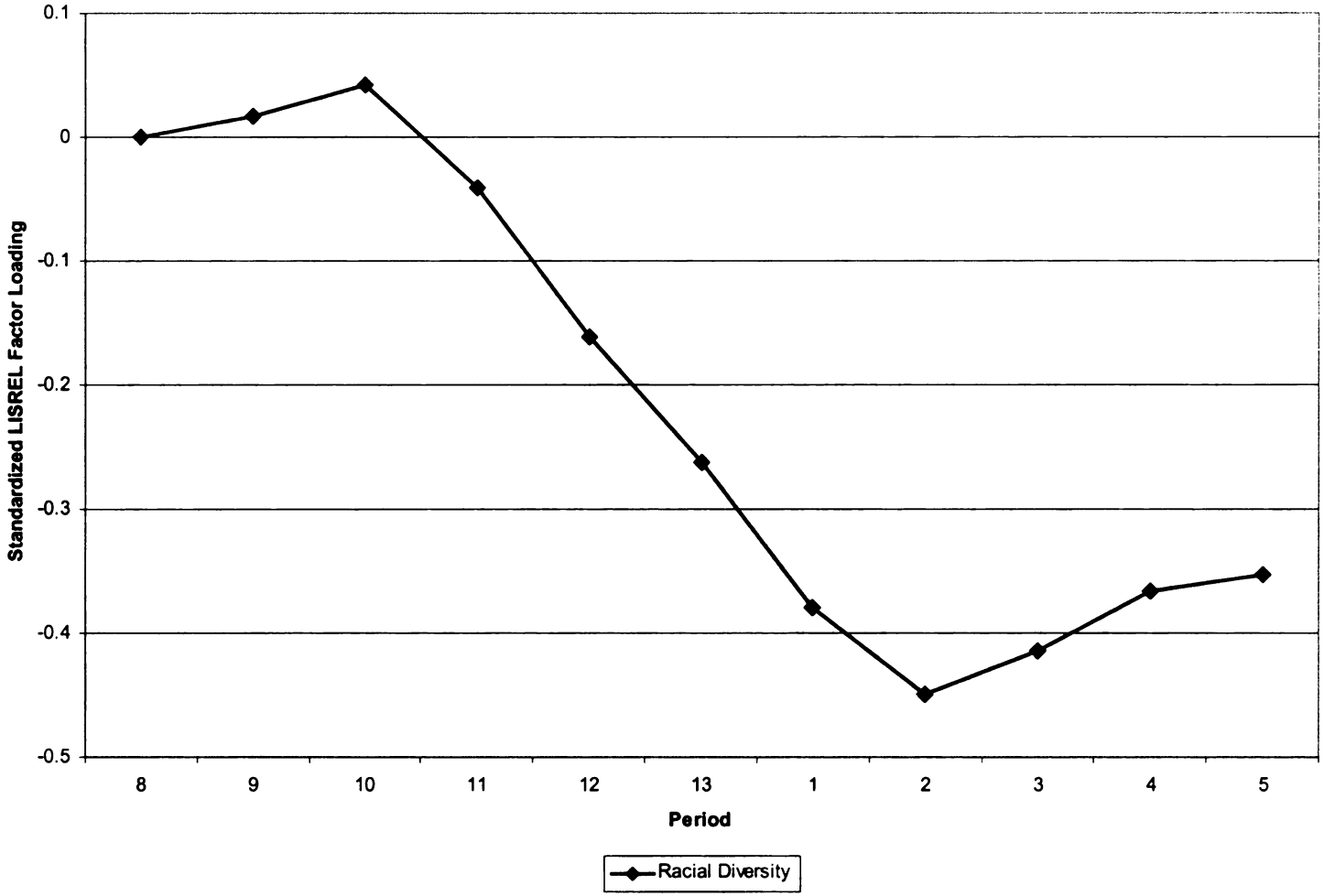

Figure 7 – Plot of standardized LISREL estimates of racial diversity slope factor loadings

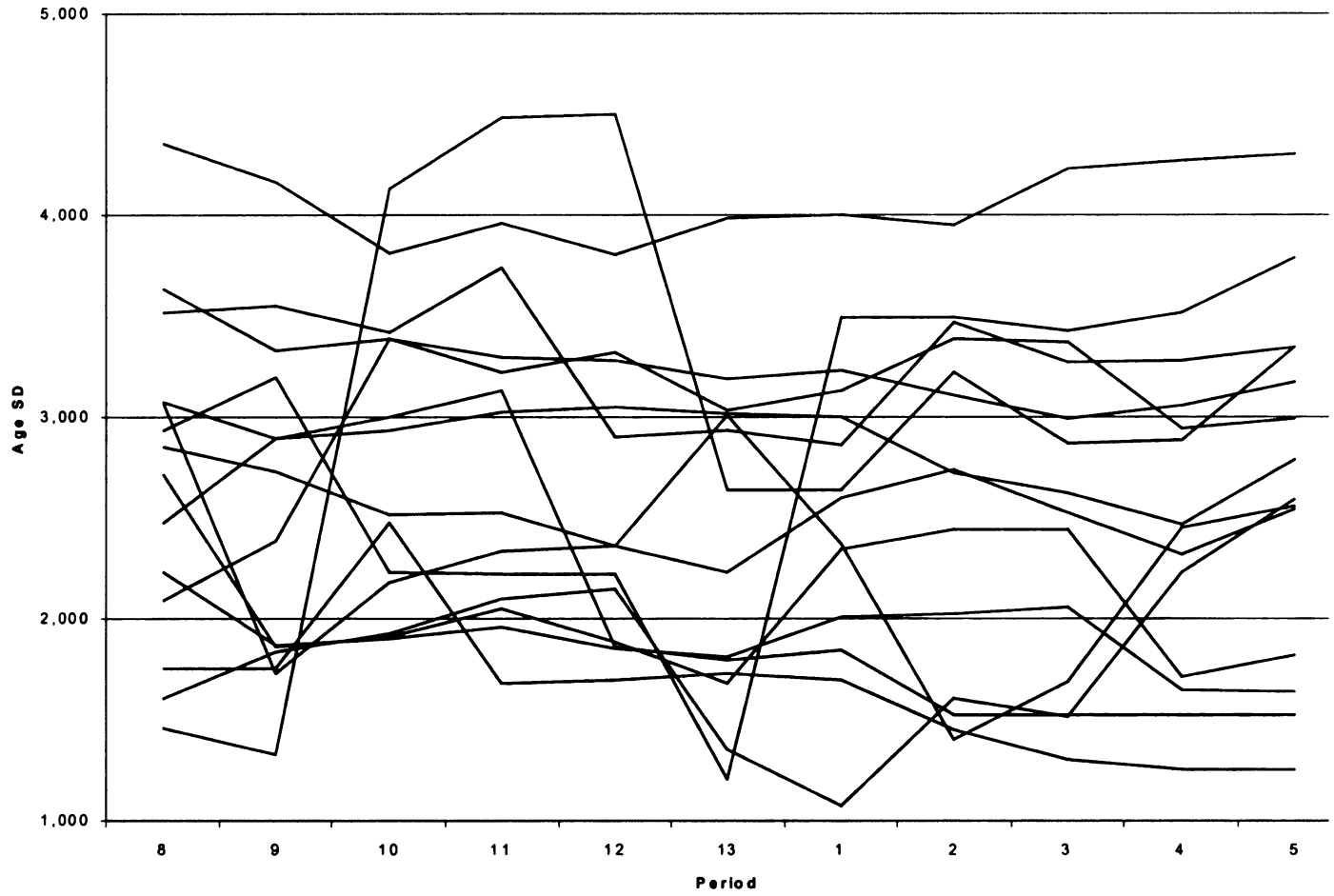

Figure 8 - Trends in age diversity for a sample of 15 restaurants.

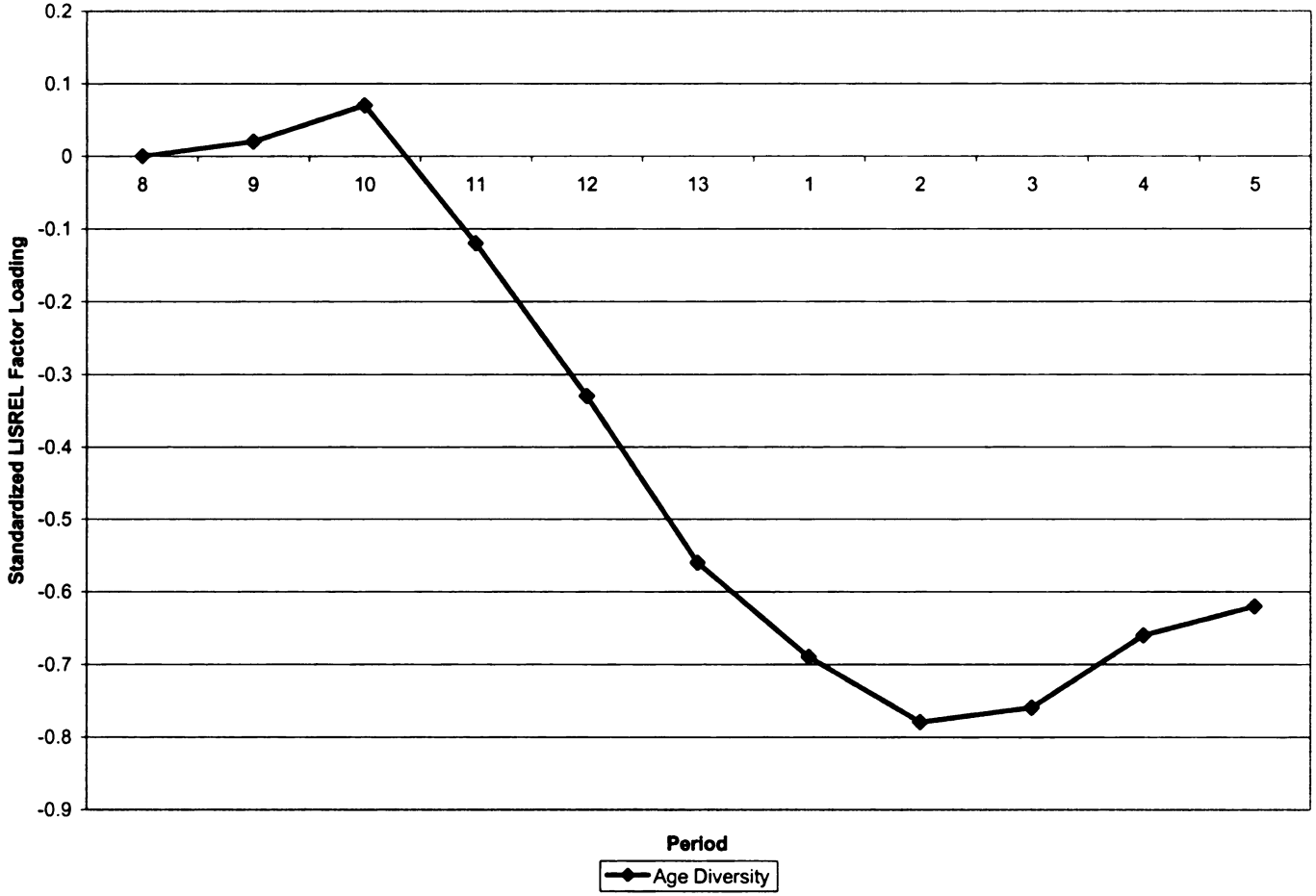

Figure 9 – Plot of standardized LISREL estimates of age diversity slope factor loadings

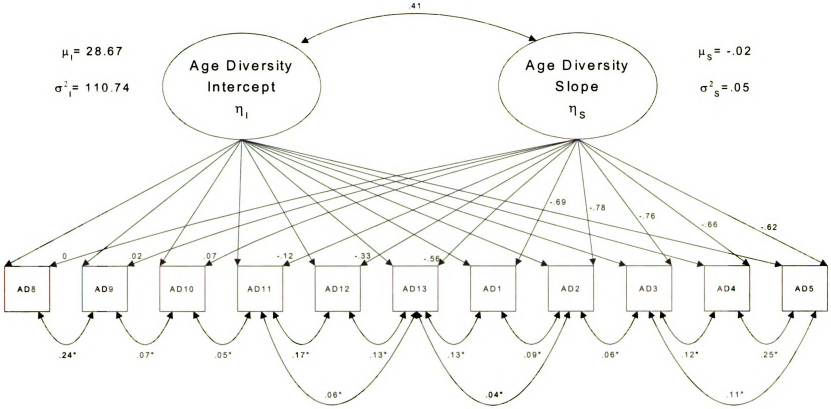

Figure 10 - Latent growth curve model for age diversity. N = 2388 restaurants.

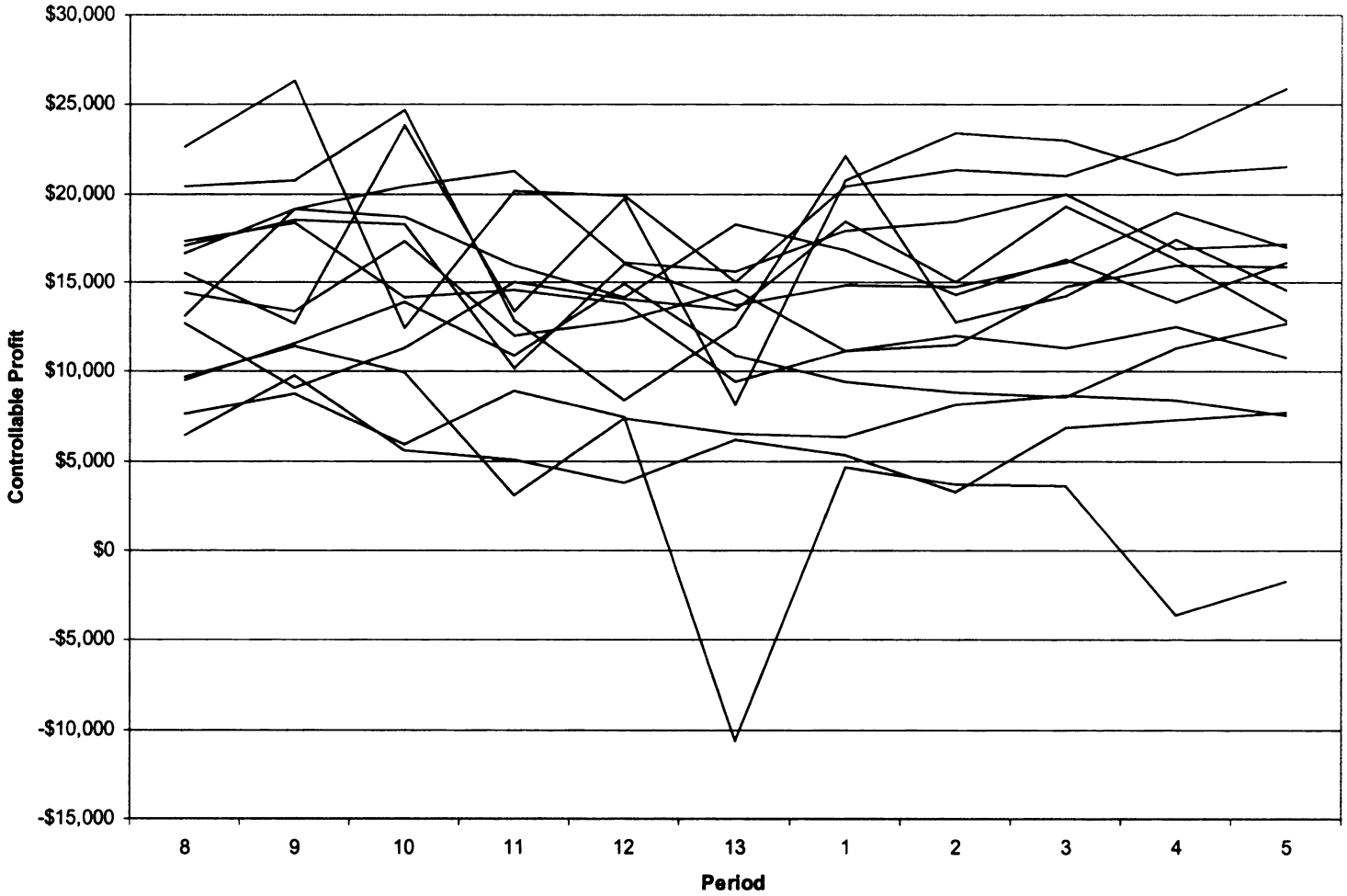

Figure 11 - Trends in controllable profit for a sample of 15 restaurants

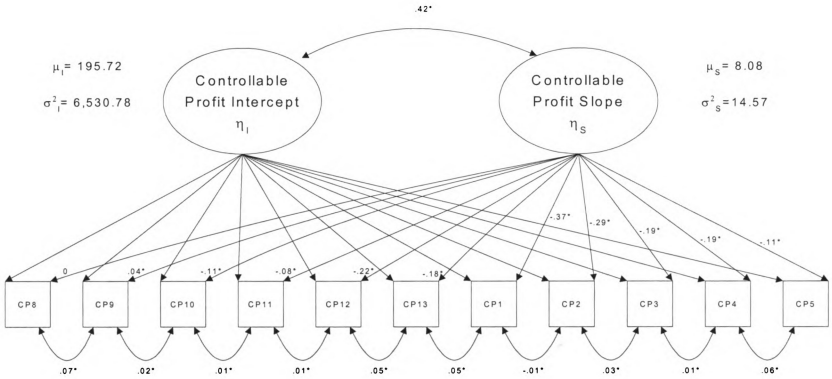

Figure 12 - Latent growth curve model for controllable profit. N = 2388 restaurants.

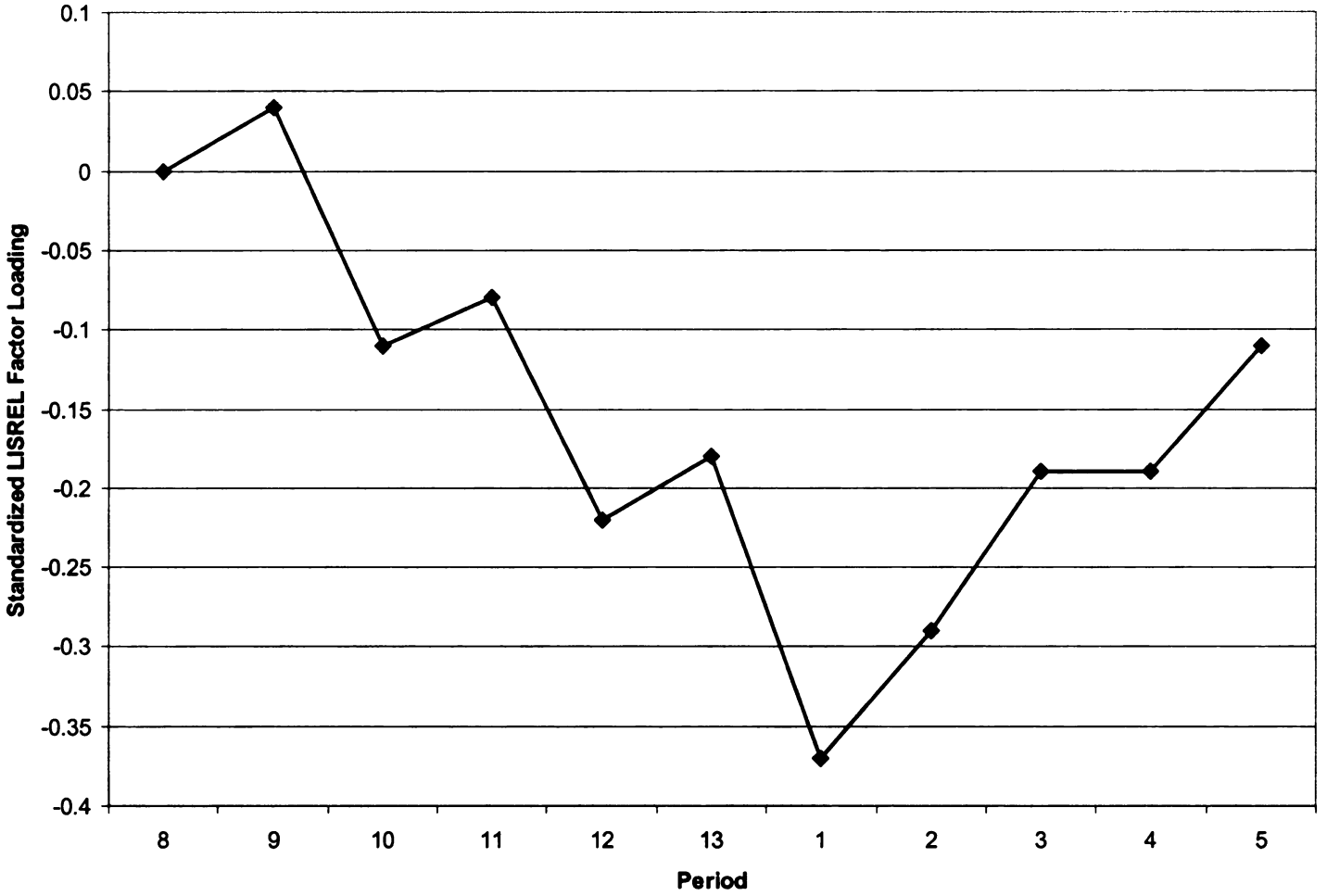

Figure 13 – Plot of standardized LISREL estimates of controllable profit slope factor loadings

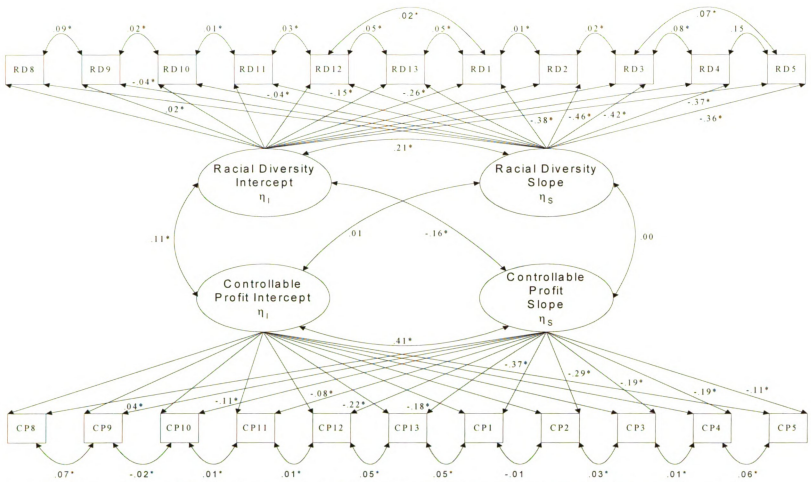

Figure 14 – Combined latent growth curve model for racial diversity and controllable profit. N = 2388 restaurants.

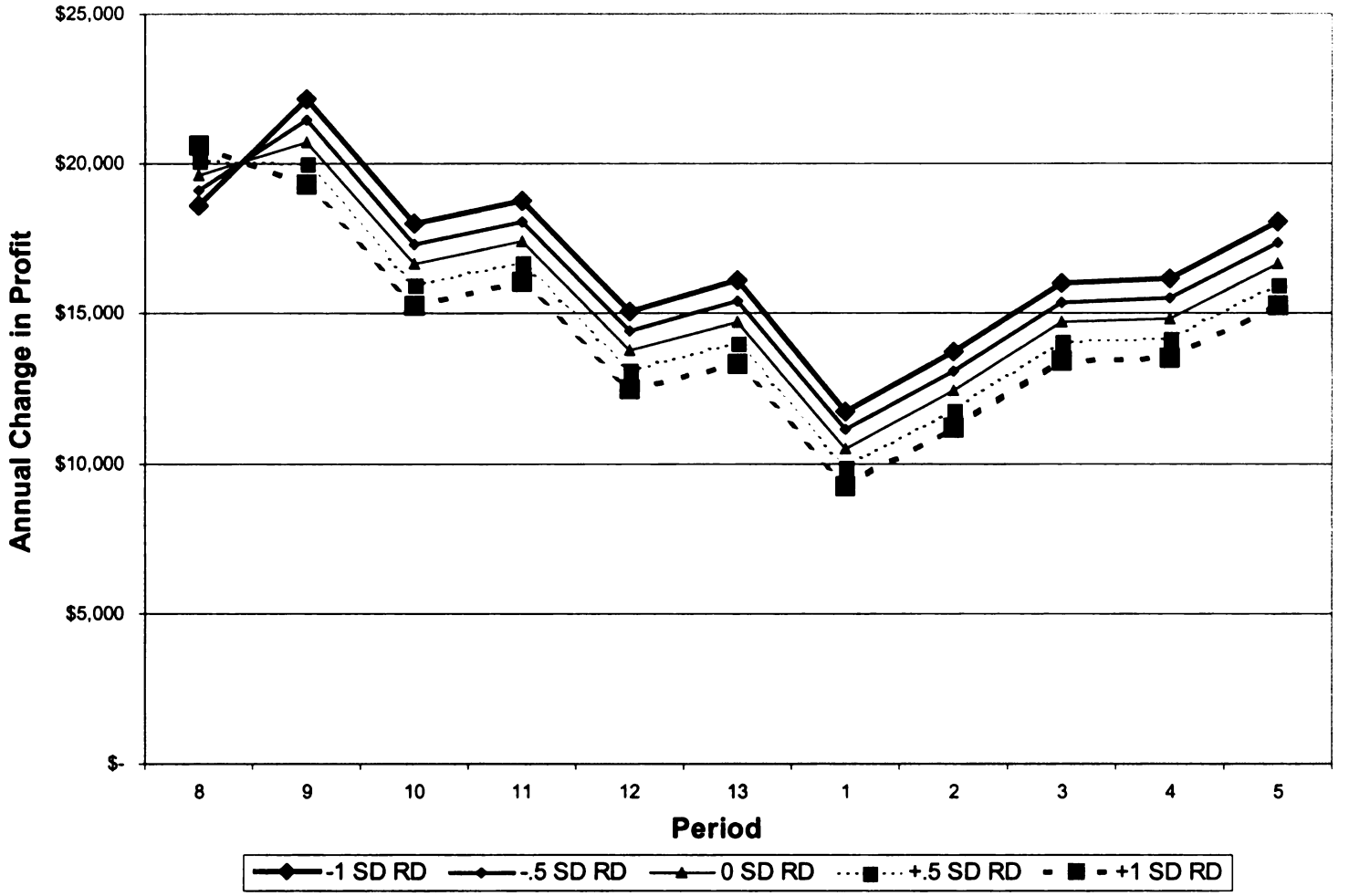

Figure 15 – Predicted trajectories for controllable profit based on several different levels of racial diversity

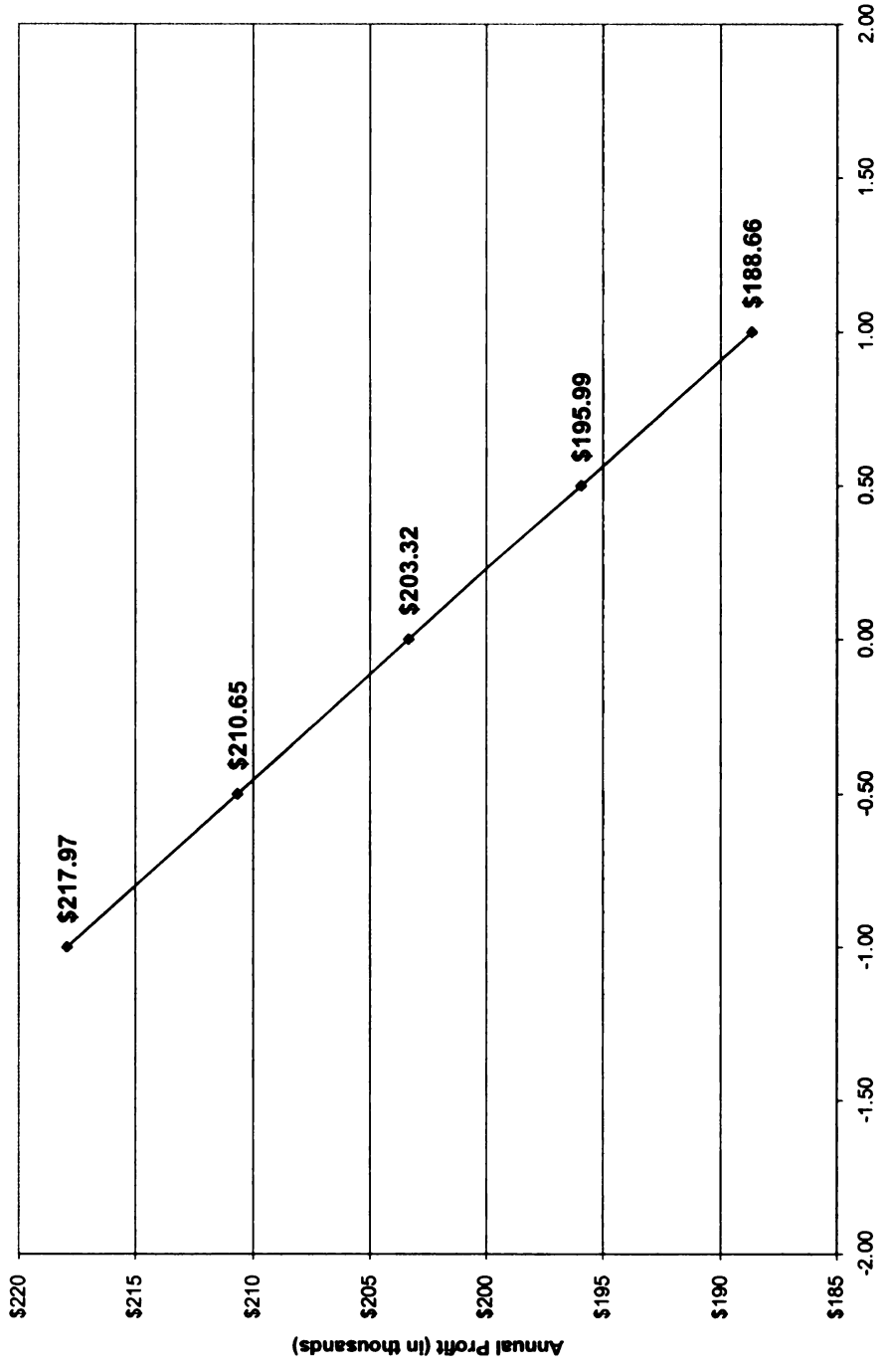

Figure 16 – Predicted yearly total impact of racial diversity on controllable profit

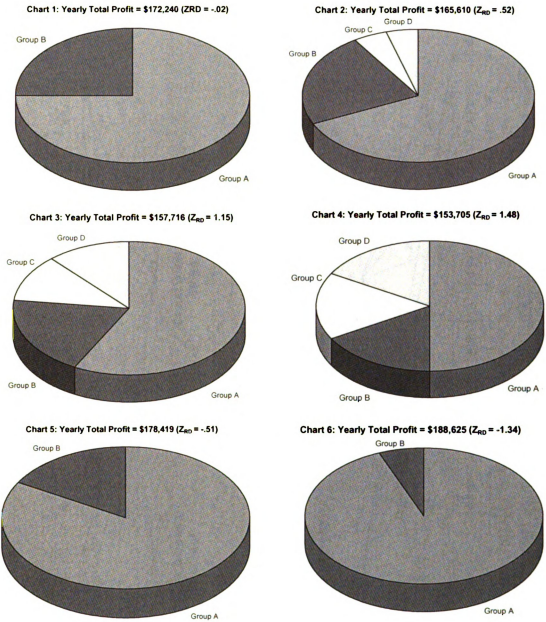

Figure 17 – Predicted yearly controllable profit for several hypothetical racial compositions.

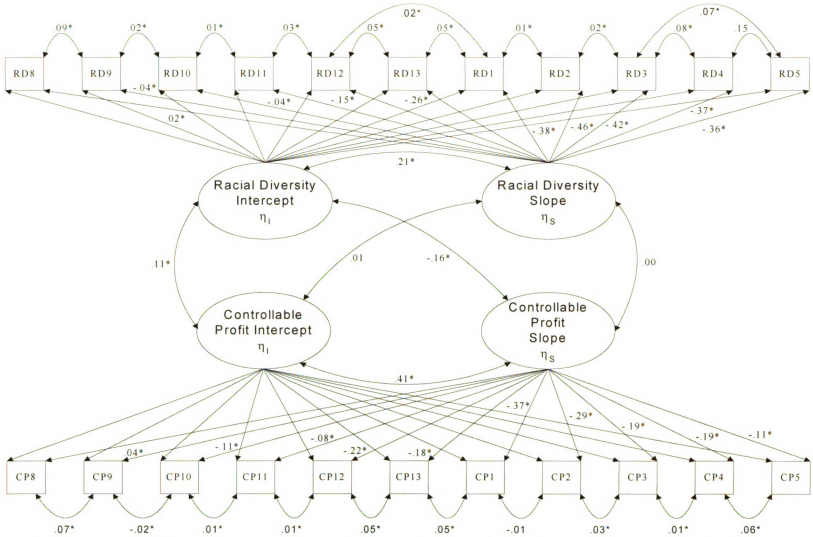

Figure 18 - Combined latent growth curve model for age diversity and controllable profit. N = 2388 restaurants.

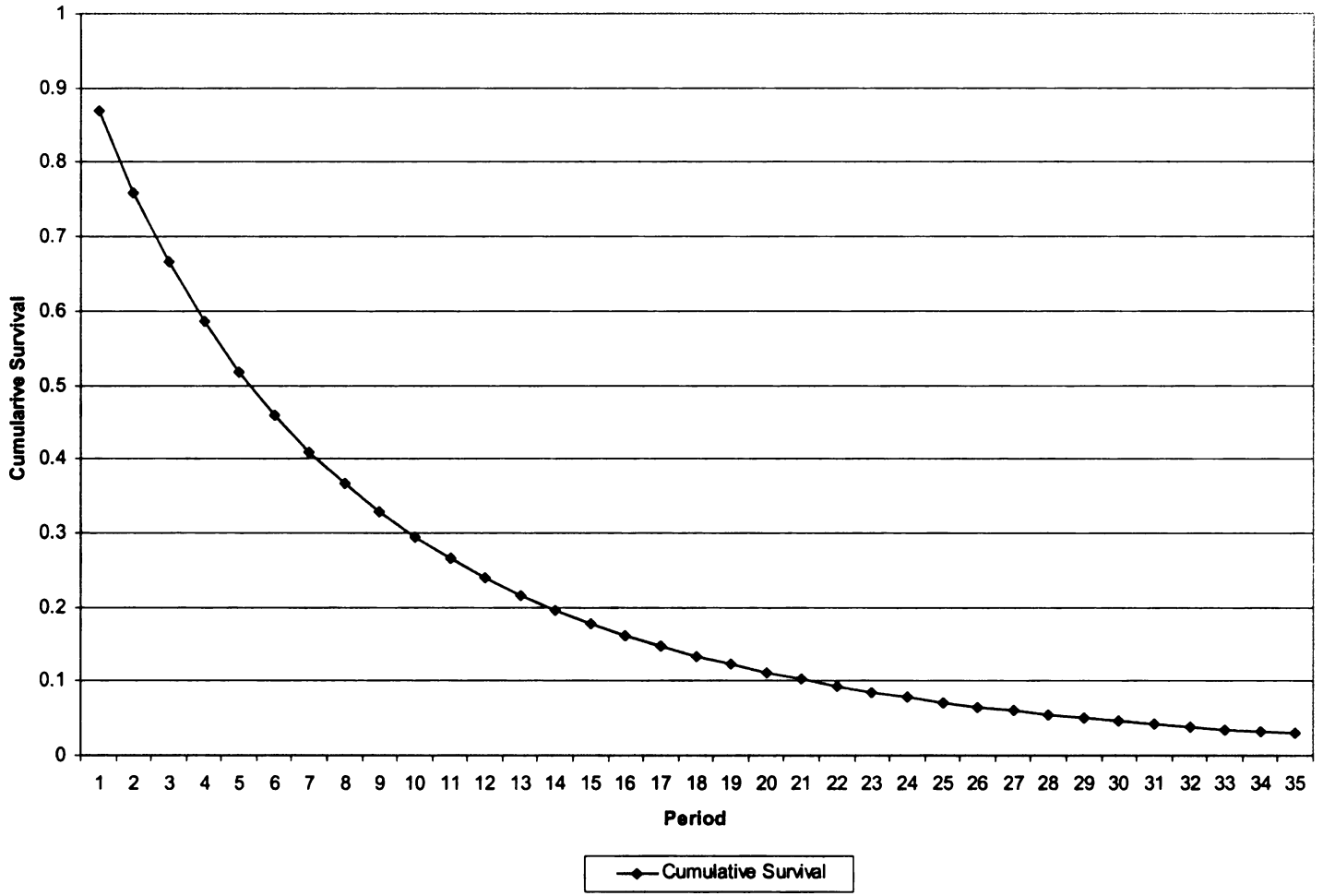

Figure 19 – Predicted cumulative survival function based on logistic regression analysis of person-period dataset with linear and quadratic time factors.

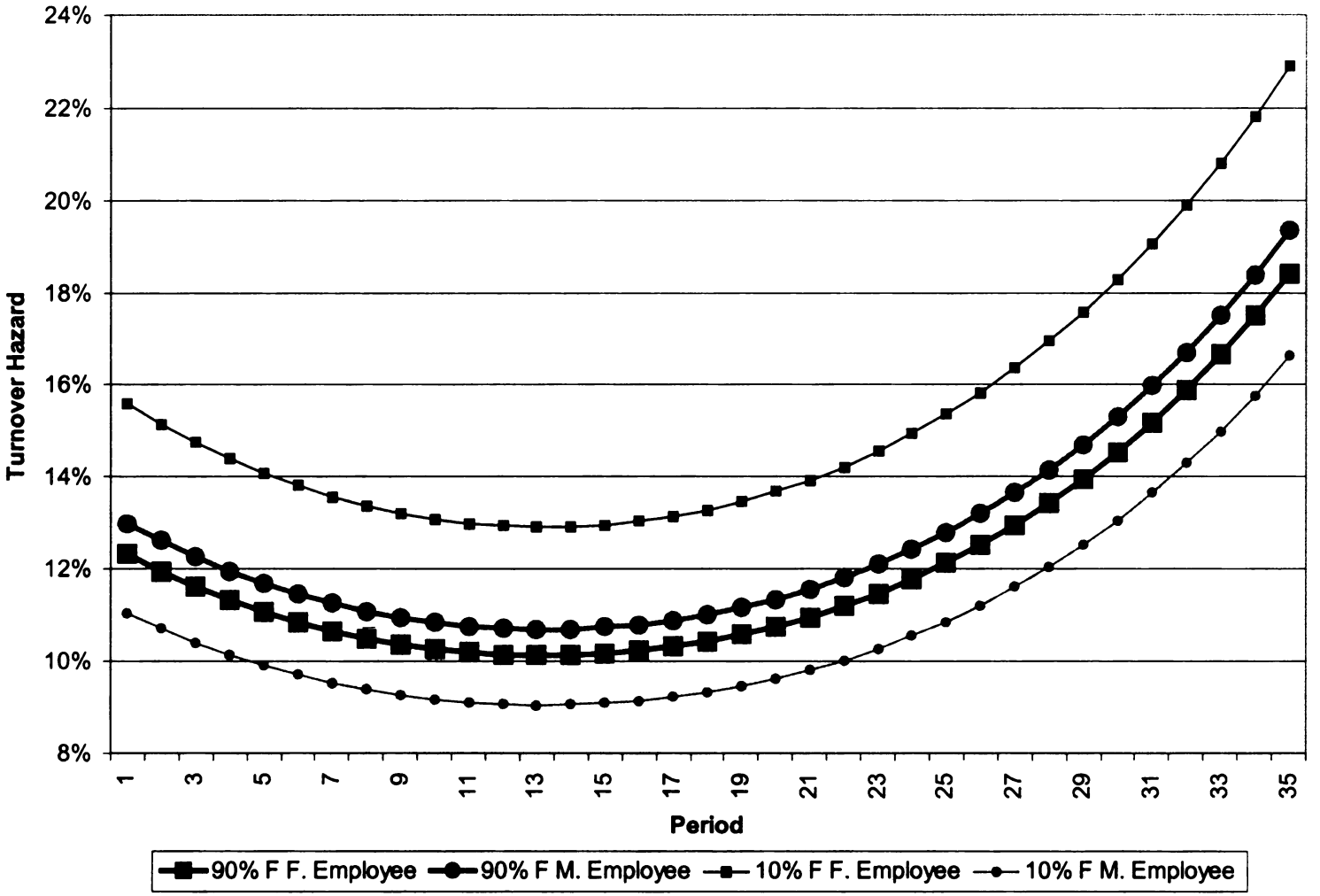

Figure 20 – Predicted turnover hazard as a function of sex and two different store sex compositions.

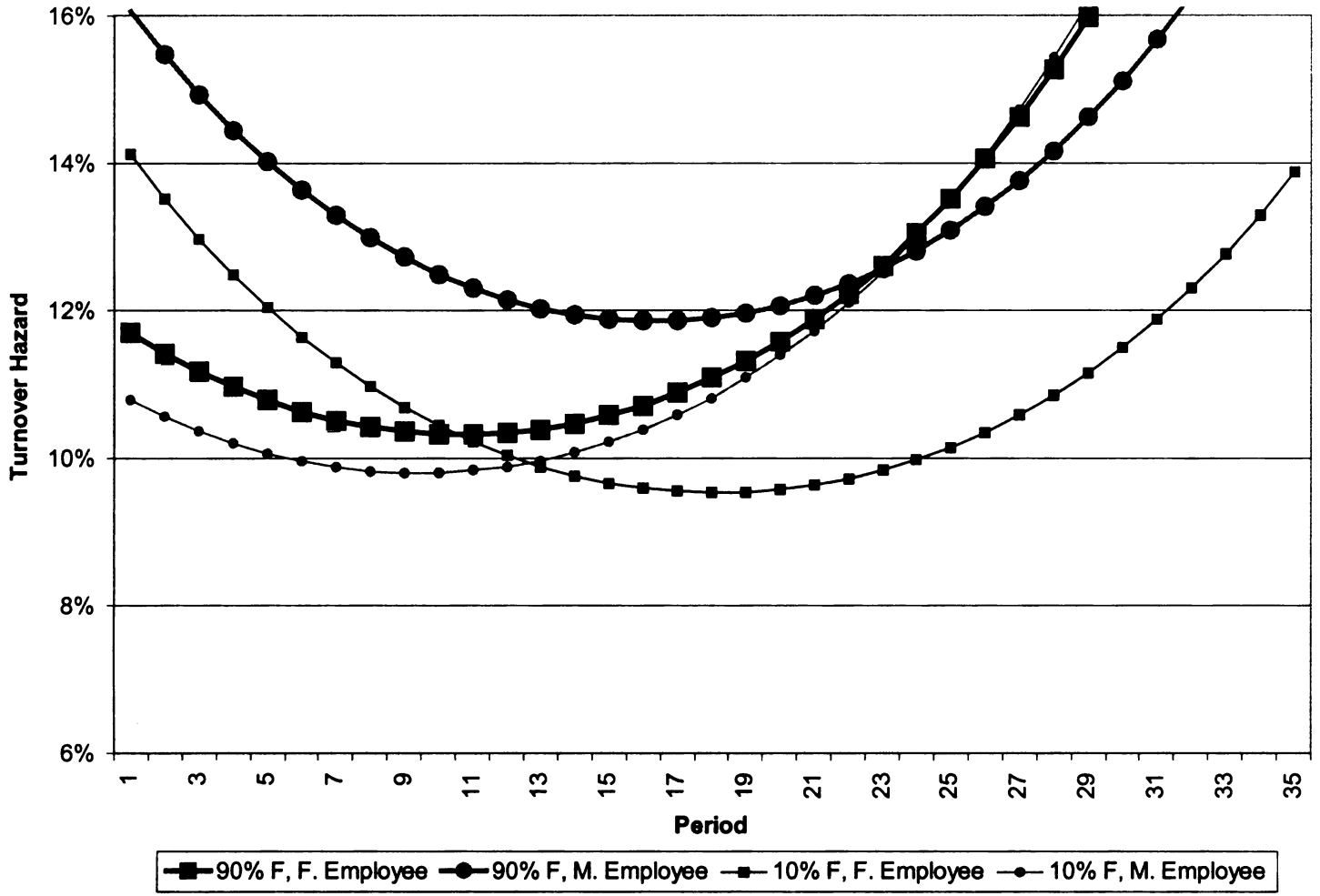

Figure 21 – Predicted turnover hazard as a function of sex and two different store compositions and time.

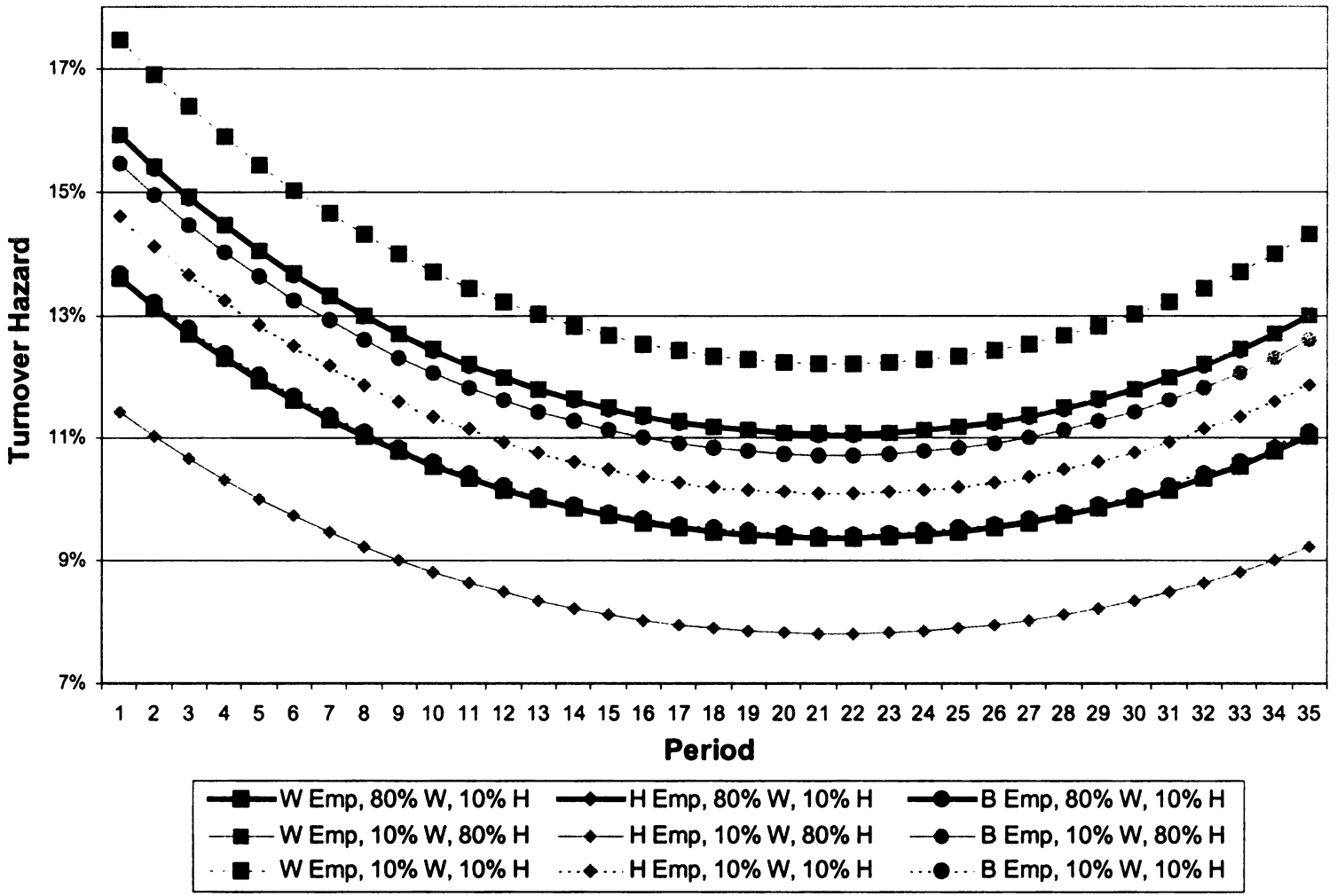

Figure 22 – Predicted turnover hazard as a function of race and three different restaurant race compositions.

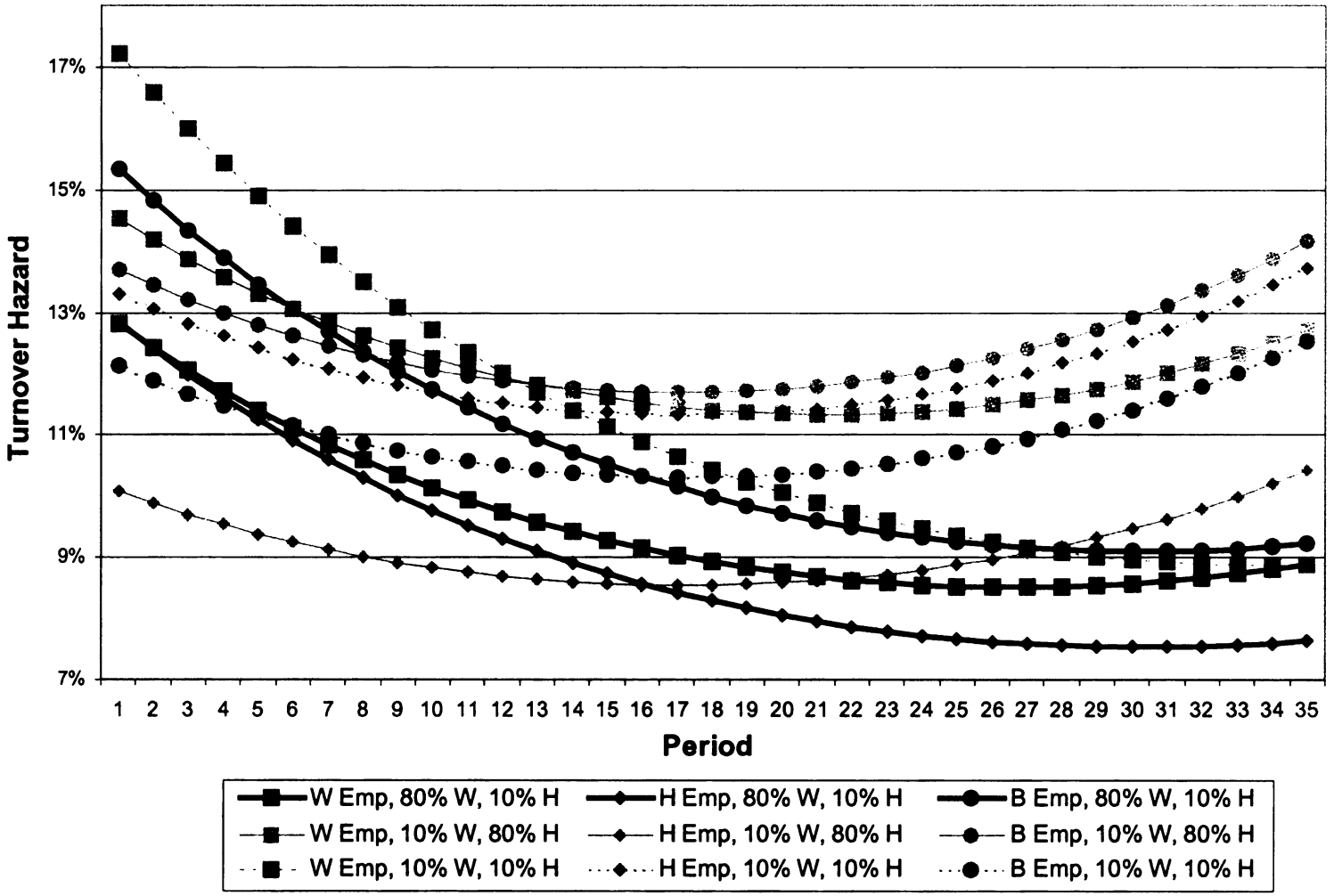

Figure 23 – Predicted turnover hazard as a function of race, three different restaurant race compositions and time.

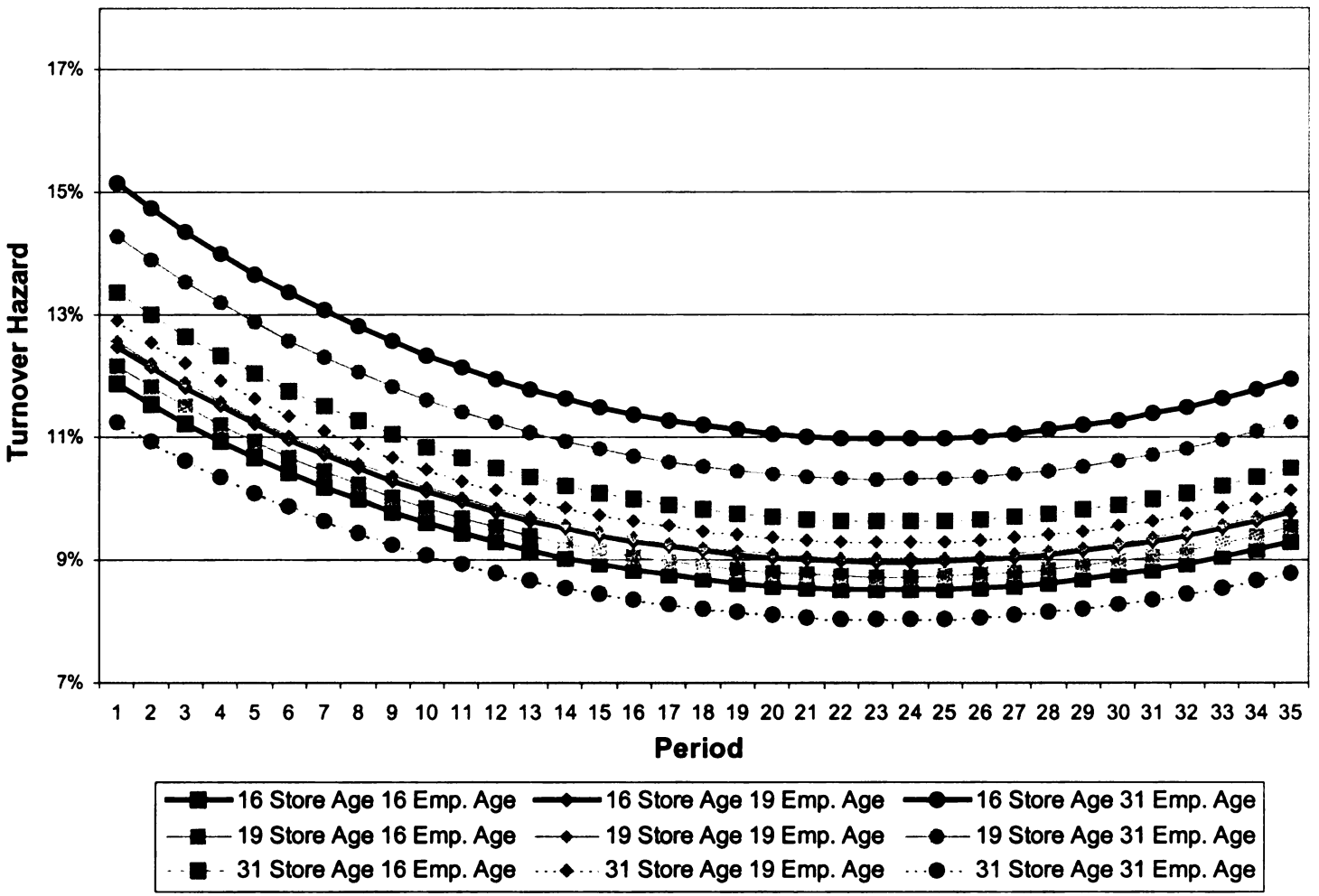

Figure 24 - Predicted turnover hazard as a function of age and three different restaurant age compositions.

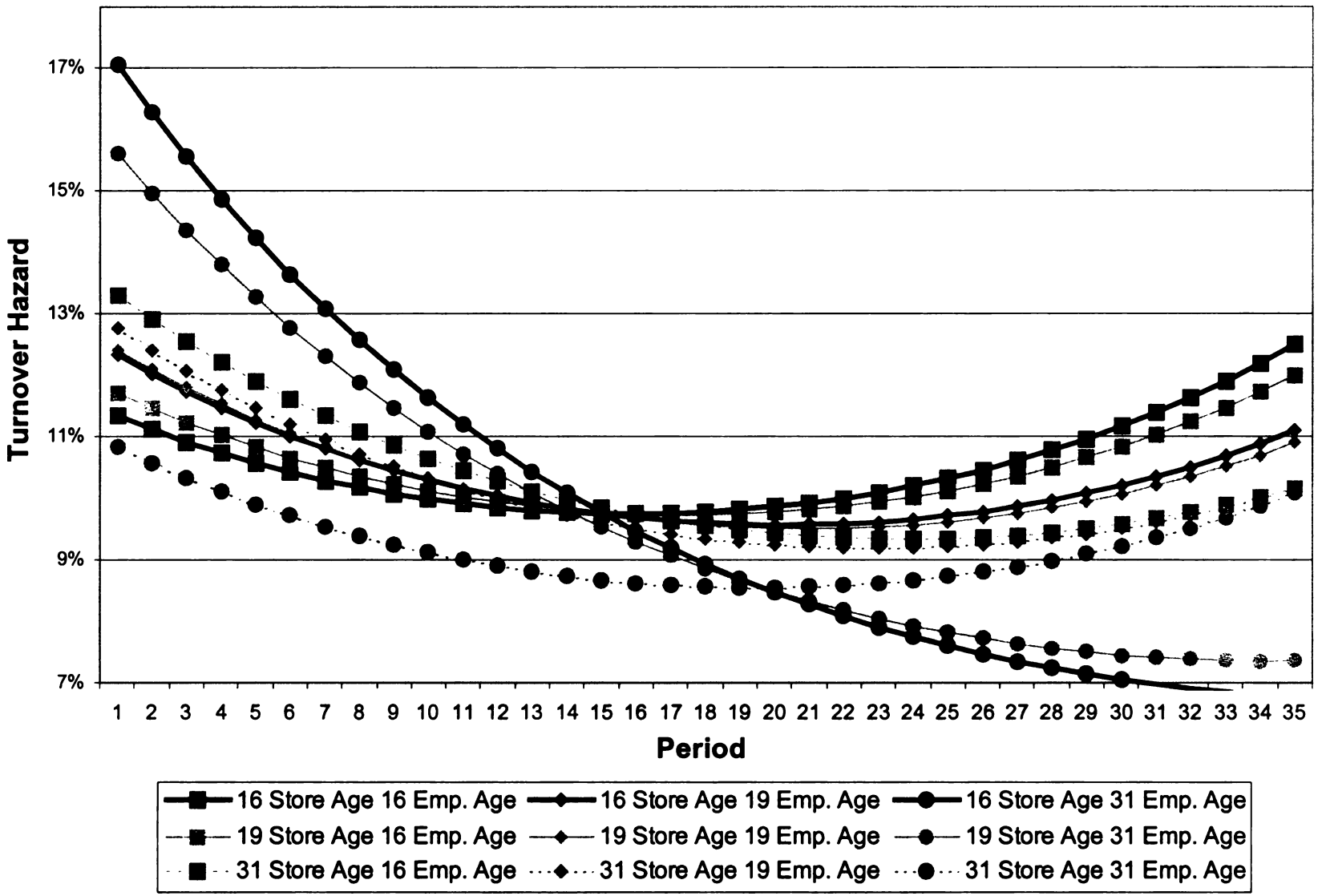

Figure 25 - Predicted turnover hazard as a function of age, three different restaurant age compositions, and time.

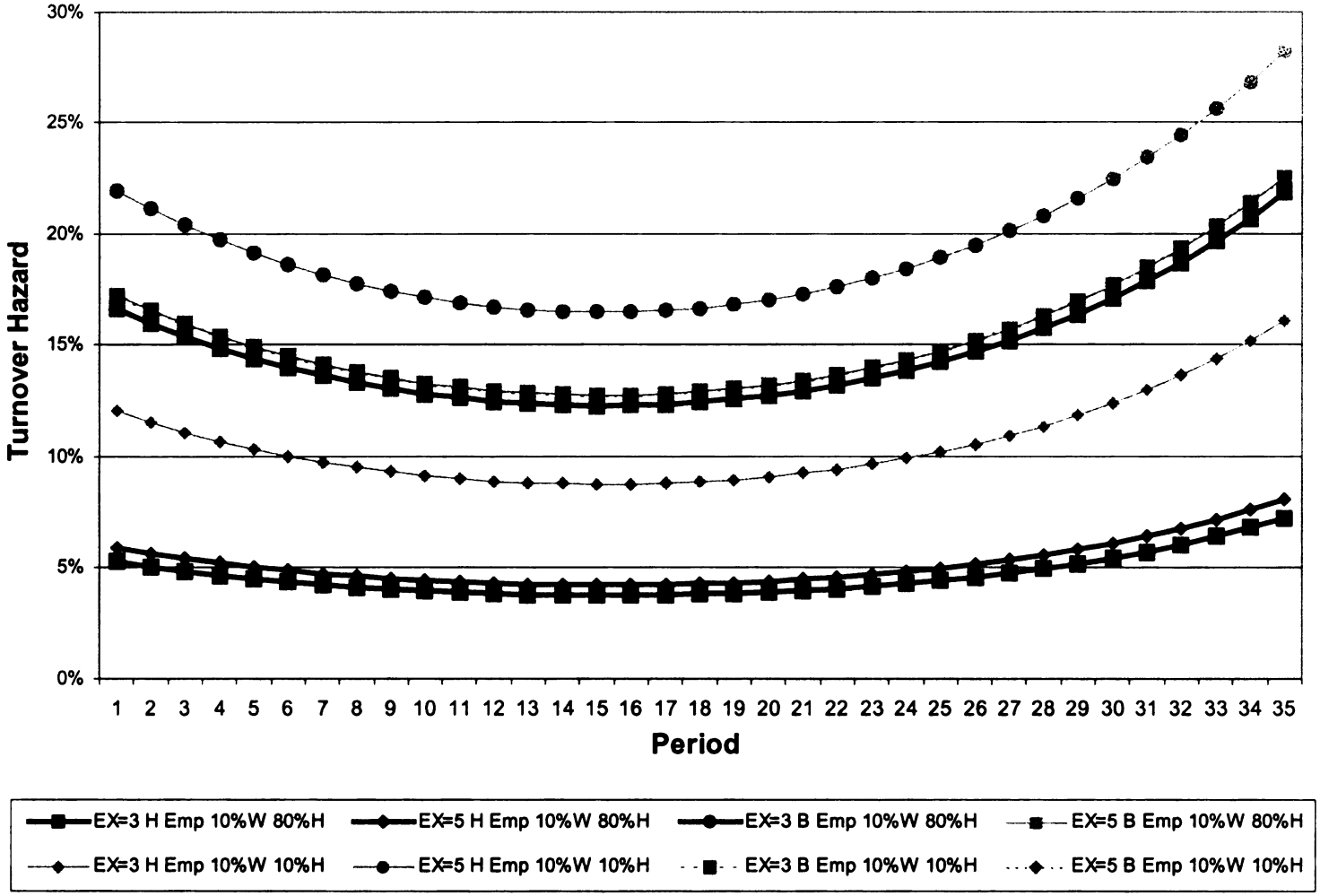

Figure 26 – Predicted turnover hazard as a function of race, two different restaurant racial compositions, and two levels of extraversion.

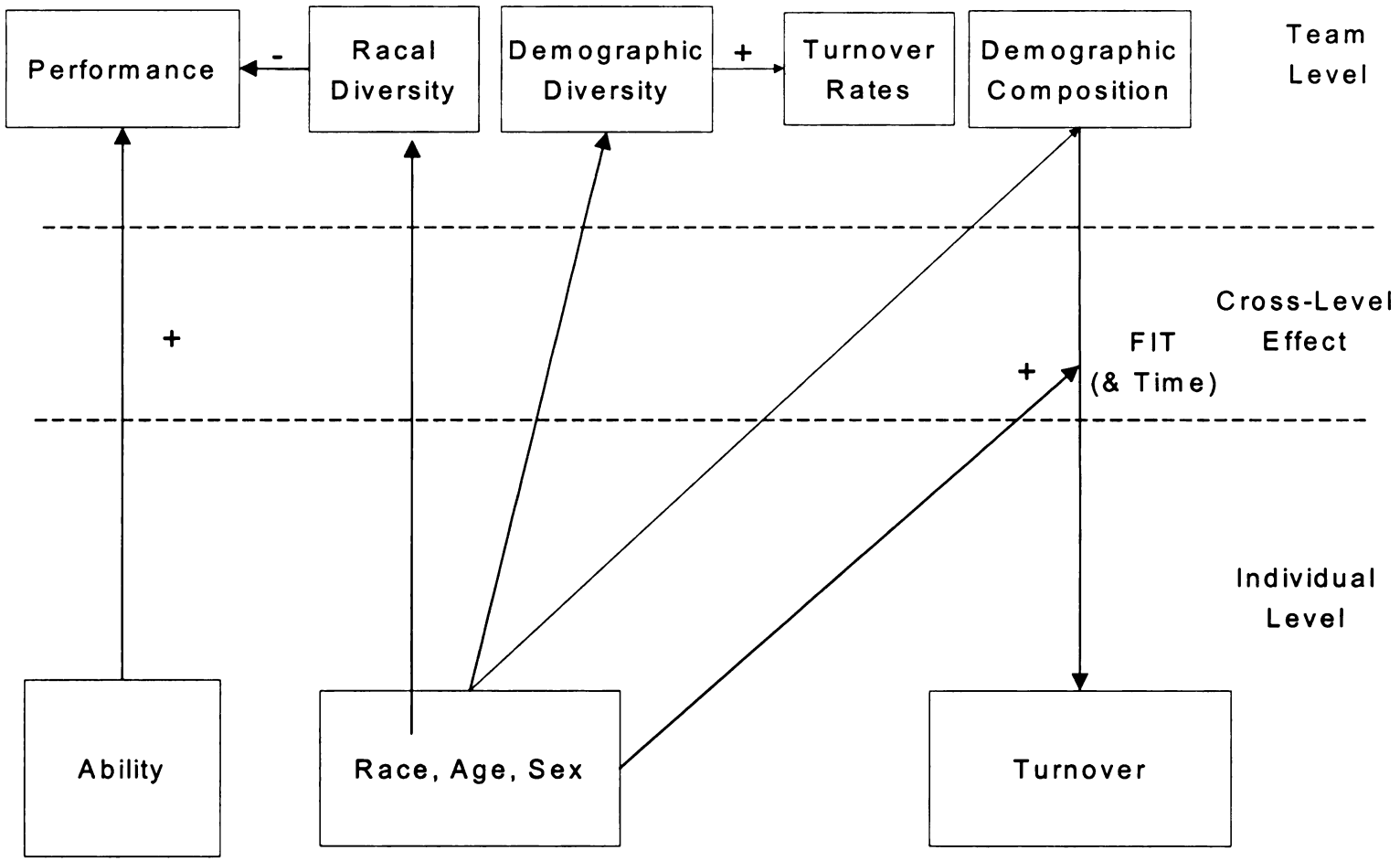

Figure 27 – Schematic summary of results

Appendices

Appendix A – IPIP items used in construct validity study.

Item	Scale	Reverse Coded?
Am the life of the party.	Extraversion	No
Feel little concern for others.	Agreeableness	Yes
Am always prepared.	Conscientiousness	No
Don't talk a lot.	Extraversion	Yes
Am interested in people.	Agreeableness	No
Leave my belongings around.	Conscientiouesness	Yes
Feel comfortable around people.	Extraversion	No
Insult people.	Agreeableness	Yes
Pay attention to details.	Conscientiouesness	No
Keep in the background.	Extraversion	Yes
Sympathize with others' feelings	Agreeableness	No
Make a mess of things.	Conscientiouesness	Yes
Start conversations.	Extraversion	No
Am not interested in other people's problems.	Agreeableness	Yes
Get chores done right away.	Conscientiouesness	No
Have little to say.	Extraversion	Yes
Have a soft heart.	Agreeableness	No
Often forget to put things back in their proper place.	Conscientiouesness	Yes
Talk to a lot of different people at parties.	Extraversion	No
Am not really interested in others.	Agreeableness	Yes
Like order.	Conscientiouesness	No
Don't like to draw attention to myself.	Extraversion	Yes
Take time out for others.	Agreeableness	No
Shirk my duties.	Conscientiouesness	Yes
Don't mind being the center of attention.	Extraversion	No
Feel others' emotions.	Agreeableness	No
Follow a schedule.	Conscientiouesness	No
Am quiet around strangers.	Extraversion	Yes
Make people feel at ease	Agreeableness	No
Am exacting in my work.	Conscientiouesness	No

Appendix B – Factor analysis results for job performance and skill area ratings

	Factor Titles and Loadings			
	Working Conscien-tiously & Taking Initiative	Team and Interper-sonal Skills	Customer Focus	Attendance, Tardiness, and Dependability
Takes initiative/is self directed	.76	.28	.05	.13
Requires Close Supervision	.73	.20	.18	.13
Completes Assigned Tasks	.68	.07	.16	.26
Goes Beyond Required Duties	.67	.32	.17	.17
Ability to Follow Procedures	.61	.47	.09	.14
Follows Policies & Procedures	.60	.11	.28	.20
Makes Careless Mistakes	.60	.08	.21	.10
Communication Skills	.19	.71	.25	.05
Flexibility	.04	.64	.09	.39
Leadership Characteristics	.48	.63	.08	.03
Team Skills	.42	.62	.16	.27
Ability to Learn Procedures	.54	.55	.03	.06
Customer Perspective	.33	.54	.49	.03
Agreeableness	.15	.51	.41	.38
Stability	.39	.40	.31	.34
Pleasant to Customers Is Rude or Impatient with Customers	.23	.16	.78	.04
Does Not Get along with Coworkers	.14	.00	.77	.13
Meets Customers 'Needs	.39	.24	.63	.05
Does not Get Along with Coworkers	.01	.19	.51	.30
Misses Work	.15	.07	.11	.81

Appendix C – Intercorrelations and internal consistencies for the performance composites

	Variable					
Performance Measure	1	2	3	4	5	6
1. Working Conscientiously and Taking Initiative	(.91)	.47	.70	.39	.77	.91
2. Customer Focus		(.82)	.48	.26	.44	.61
3. Team and Interpersonal Skills			(.86)	.39	.78	.89
4. Attendance, Tardiness, and Dependability				(.86)	.45	.57
5. Overall Job Performance Rating					--	.83
6. Performance Composite Average						(.95)

Notes: Internal consistency estimates are on the diagonol. N = 843 – 867. All correlations are significant at $p < .05$.

Appendix D – Summary of validity coefficients for CRI scales

Job Performance and Job Skill Area	Working Conscientiously & Taking Initiative	Team and Interpersonal Skills	Performance Customer Focus	Dimensions Attendance, Tardiness, & Dependability	Overall Performance	Composite Performance
CRI Total Score	.28*	.24*	.23*	.18*	.21*	.28*
CRI Subscales						
Work Orientation	.26*	.21 *	.13*	.14*	.20*	.26*
Service Orientation	.16*	.14*	18*	.05	.12*	.17*
Team Orientation	19*	19*	18*	.06*	.13*	.21*
Tenure	.21*	.20*	.12*	.17*	.19*	.23*
Fundamentals	11*	11*	.14*	.01	10*	.12

Notes: N = 764 – 774. * $p < .05$. Performance composite is a unit-weighted linear combination of the four performance dimensions. The CRI total score was calculated using regression-based weights (correcting for shrinkage reduces these total score correlations by .01 - .02).