Work Careers: A Developmental Perspective
Daniel C. Feldman, Editor, 2002

The Nature of Organizational Leadership: Understanding the Performance Imperatives Confronting Today's Leaders
Stephen J. Zaccaro & Richard J. Klimoski, Editors, 2001

Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions
Katherine J. Klein & Steve W. J. Kozlowski, Editors, 2000

Compensation in Organizations: Current Research and Practice
Sara L. Rynes & Barry Gerhart, Editors, 2000

The Changing Nature of Performance: Implications for Staffing, Motivation, and Development
Daniel R. Ilgen and Elaine D. Pulakos, Editors, 1999

New Perspectives on International I-O Psychology
P. Christopher Earley, Miriam Erez, Editors, 1997

Individual Differences and Behavior in Organizations
Kevin R. Murphy, Editor, 1996

The Changing Nature of Work
Ann Howard, Editor, 1995

Team Effectiveness and Decision Making in Organizations
Richard A. Guzzo and Eduardo Salas, Editors, 1995

Personnel Selection in Organizations
Neal Schmitt and Walter C. Borman, Editors, 1993

Work, Families, and Organizations
Shelton Zedeck, Editor, 1992

Organizational Climate and Culture
Benjamin Schneider, Editor, 1990

Training & Development in Organizations
Irwin L. Goldstein, Editor, 1989

Productivity in Organizations
John P. Campbell and Richard J. Campbell, Editors, 1988

Career Development in Organizations
Douglas T. Hall, Editor, 1986

Perspectives on Organizational Fit

EDITED BY

Cheri Ostroff

Timothy A. Judge

LEA Lawrence Erlbaum Associates
Taylor & Francis Group

New York London

Cover design by Tomai Maridou

Lawrence Erlbaum Associates
Taylor & Francis Group
270 Madison Avenue
New York, NY 10016

Lawrence Erlbaum Associates
Taylor & Francis Group
2 Park Square
Milton Park, Abingdon
Oxon OX14 4RN

Lawrence Erlbaum Associates is an imprint of Taylor & Francis Group, an Informa business

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-0-8058-5195-3—0-8058-5195-X (case)
978-1-4106-1778-1—1-4106-1778-5 (e book)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the LEA Web site at
http://www.erlbaum.com

Contents

Series Foreword

This is the 25th book in the Organizational Frontiers Series of books initiated by the Society for Industrial and Organizational Psychology (SIOP). The overall purpose of the Series volumes is to promote the scientific status of the field. Ray Katzell first edited the Series. He was followed by Irwin Goldstein, Sheldon Zedeck, and Neal Schmitt. The topics of the volumes and the volume editors are chosen by the editorial board, or individuals propose volumes to the editorial board. The series editor and the editorial board then work with the volume editor(s) in planning the volume.

The success of the series is evident in the high number of sales (now over 50,000). Volumes have also received excellent reviews, and individual chapters as well as volumes have been cited very frequently.

This volume, edited by Cheri Ostroff and Timothy A. Judge, reflects thinking and research on issues concerning the fit between individuals and their organizations. The idea of fit is an important topic receiving greater and greater attention. However, there has been no systematic integration of this material. There have been many different approaches to fit, and this volume identifies them and uses a strong conceptual focus to integrate these different points of view.

There are several other strengths of this volume. Almost by definition, fit reflects a multilevel perspective ranging from micro, individual-level, variables to macro, organizational, variables. This volume comprehensively addresses this multilevel issue. It also integrates theory and research done in different domains such as selection, recruitment, diversity, and leadership. The volume, especially the concluding chapter, specifically addresses issues of methodology in this area and needed research.

The editors and chapter authors deserve our gratitude for clearly communicating the nature, application, and implications of the theory and research described in this book. Production of a volume such as this involves the hard work and cooperative effort of many individuals. The editors, the chapter authors, and the editorial board all played important roles in this endeavor. As all royalties from the Series volumes are used to help support SIOP; none of the editors or authors received any remuneration. The editors and authors deserve our appreciation for engaging a

Sampling Issues

Studies of PE fit can be complicated by the fact that the individuals within the same unit or organization are exposed to the same environmental conditions; hence, variability on *E* is of concern. For example, in a single study whereby variability exists on *P* but *E* is the same or relatively similar for all individuals, the study will be reduced to an individual difference study, not a true fit study, because *E* is a constant across those under investigation. Similarly, based on the attraction–selection–attrition process, individuals within a unit are likely to be relatively homogeneous (e.g., Schneider et al., 1997), making variability on *P* a potential concern. Taken together, PE fit studies conducted within a single unit or organization are likely to exhibit severe restrictions on the variability of both *P* and *E* factors. This often necessitates collection of data across units and/or across organizations to provide the necessary variability on both factors.

Timing of Data Collection

As with any study, when all measures (*P* and *E* factors, antecedents, and outcomes) are collected at the same time, from the same sources, response bias is likely to be a problem. These problems can be ameliorated to some degree by careful attention to the design of the measures as well by providing time lags between collection of measures (e.g., Harrison, McLaughlin, & Coalter, 1996). Equally important is the consideration of fit as a dynamic process that occurs over time (Tinsley, 2000). The notion that individuals can change the *E* and that *E* can change individuals' attributes over time requires longitudinal and panel type of designs as opposed to concurrent measurement strategies.

In this section of the chapter, some basic measurement and design issues were highlighted. In the remaining subsections of this chapter, each author summarizes a specific statistical or analytical technique that can be used to assess fit and test fit hypotheses.

11.2 PROFILE COMPARISON METHODS FOR ASSESSING PERSON–SITUATION FIT

David F. Caldwell, Jennifer A. Chatman and Charles A. O'Reilly

Characterizing people and situations and comparing them in meaningful ways remain key challenges for person–environment (PE) fit researchers. One common approach is to identify a specific individual characteristic, such as a personality trait, and a specific situational characteristic, such as a job attribute, and to directly investigate the joint effects or interactions between the two variables on some outcome. Such an approach is effective for testing specific predictions, such as whether sales people who are more extraverted fit the customer-oriented demands of the jobs and therefore sell more than those who are less extraverted. This type of approach is appropriate when strong theoretical links between the specific *P* and *E* variables exist, and the variables are considered centrally relevant to describing both the person and situation. For example, compared with introverted people, extraverted people should be more comfortable with the high level of interpersonal interaction required to be successful in sales jobs (Volard & McCarthy, 1979).

This approach is not appropriate, however, for answering questions that require information about a person's overall fit to a situation and how that fit influences behavioral outcomes. Overall fit is important because neither people nor situations are unidimensional. Thus, one attribute of an individual could fit well with a particular situational attribute, but the person might also have other characteristics that are incompatible with important attributes also present in that situation. This becomes important even if a person or situation is characterized by two dimensions rather than by one. To extend the aforementioned example of extraversion and interpersonal abilities, for example, it is possible that only extraverted people who are also conscientious are effective as sales representatives because success in sales involves both getting along with others as well as gathering information, being prepared to answer questions, and following up with customer requests (e.g., Barrick, Mount, & Strauss, 1993). Using techniques that only assess the fit between a single person and situation attribute to predict an outcome may result in ambiguous and, potentially, misleading findings. For example, the relationship between extraversion and sales performance would be completely obscured if half of the extraverted people were highly conscientious and half were not conscientious at all. Thus, the primary threat to conducting valid person–situation fit research involves omitting important attributes that characterize persons and situations in appropriately comprehensive and relevant terms (e.g., Chatman, 1989).

Because people and situations are multidimensional, meaningful and valid tests of fit hypotheses require comprehensive descriptions of both persons and situations. One advantage of profile methods is that a profile can describe both the person and situation across a large number of dimensions. Further, a single index that captures the overall degree of fit across dimensions is created. Profiles, therefore, have the potential to be more comprehensive than, for example, approaches that are based on experimental designs examining the interaction between a person and situation variable (e.g., Chatman & Barsade, 1995) or approaches that use statistical interactions (e.g., Edwards, 1995) to study person situation fit.

A second important difference between profile techniques and other ways of assessing fit is that profiles allow for a semi-idiographic assessment of fit, yielding the benefits of both idiographic and nomothetic approaches. Most methods for assessing fit are essentially nomothetic in that they focus on between-person comparisons and assess individual differences by comparing the target person to others, usually on a single dimension. While nomothetic methods enable comparisons across people or across time (e.g., Luthans & Davis, 1982), they do not allow researchers to assess important within-person comparisons that also influence an individual's behavior or affect.

As Pelham (1993) has noted, assessing the relative importance of traits to that individual may yield very different results than when those same traits are assessed by comparing that person to other people. And, the extent to which a trait is part of an individual's self-concept influences how he or she processes information relevant to that trait (Markus, 1977). Theoretical conceptualizations of fit often assume that either the personal traits or situational characteristics being assessed are equally relevant to all individuals. This assumption is inappropriate for certain research questions such as how various traits relate to other traits within a person, how well a person's knowledge, skills, and abilities fit with an array of requisite job attributes (e.g., Caldwell & O'Reilly, 1990), or how well a person's values fit with an organization's overall culture (e.g., Chatman, 1991; O'Reilly, Chatman, & Caldwell, 1991). Thus, an advantage of profiling techniques is that they can be used semi-idiographically to understand a person's uniquely configured characteristics while still allowing for meaningful comparisons between people.

Profiles as Assessment Tools

The profile comparison process is derived from Q-methodology and, specifically, from its application in personality assessment (e.g., Block, 1978). In a Q-sort procedure, the respondent is presented with a large number of attributes or statements (typically between 50 and 100) and asked to sort those statements into a set of predetermined categories according to some criterion, usually the extent to which the statement is characteristic of the individual. Typically, the respondent is asked to sort the items into nine categories ranging from *most characteristic* to *most uncharacteristic*. The number of items to be placed in each category is specified by a 9-point unimodal, symmetrical distribution so that the largest numbers of items are placed in the middle categories and the smallest numbers in the extreme categories. For example, if a respondent was sorting 70 items, he or she might be asked to designate 3 items as *most characteristic* (Category 9) and 3 items as *most uncharacteristic* (Category 1). The middle categories would have substantially more items. In this case the middle category, *neither characteristic nor uncharacteristic* (Category 5), might contain 18 items.

The forced shape of the distribution offers a number of advantages over free rating schemes. In free rating schemes, respondents are permitted to place any number of items in any category and are not required to place all items in categories, thereby making this method purely idiographic because it results in a unique configuration for each focal individual; however, the unequal numbers of attributes per category precludes comparisons across individuals (Block, 1978). In particular, the items assigned to the middle categories are relatively less important in describing an individual than are those assigned to the extreme categories, yet raters find discriminating among the items in the middle categories to be relatively difficult. Increasing the number of items that must be placed in the middle categories (as in a forced distribution method) minimizes these difficult, yet relatively unimportant, discriminations and better fits with people's cognitive capabilities to discriminate reliability among attributes.

Profile techniques are also robust in that the concourse of items individuals are asked to sort can vary, as can the definition of the categories into which the items are sorted. Although for most of the early uses of the technique individuals were asked to sort personality descriptors, we have used competencies (Caldwell & O'Reilly, 1990) and values (Chatman, 1991; O'Reilly, Chatman, & Caldwell, 1991) in addition to personality descriptors (Chatman, Caldwell, & O'Reilly, 1999). Researchers can also use standard sets of items [e.g., California Adult Q-Sort (Block, 1978) or Organization Culture Profile (O'Reilly et al., 1991) or a unique set of items constructed for specific research contexts.

Respondents are frequently asked to sort items in terms of how characteristic of them they are; however, they can also be asked to sort the items on another dimension. In a personality–situation fit study, for example, items might be sorted to create both "real" and "ideal" personality profiles (e.g., Chatman et al., 1999). To test fit hypotheses, job experts might sort competencies on the extent to which they are required in a particular job and job incumbents might sort the same items in terms of how self-descriptive they are (e.g., Caldwell & O'Reilly, 1990).

Analyzing Profile Data

The statistical techniques for analyzing profile data are straightforward. The value of each item sorted is based on the category to which it is assigned. For example, with nine categories, an item placed in the *most characteristic* category would receive a value of 9, whereas an item described as *neither characteristic or uncharacteristic* would receive a value of 5. In a typical person–situation fit study, one set of raters might array a set of situational attributes in terms of the traits that are required to be successful in a job. A set of job incumbents would be asked to sort the same set of traits in terms of how self-descriptive they are. Fit would be measured by simply correlating the vectors of values assigned to the items (e.g., for each individual, an individual's self-profile is correlated with the situational profile derived from expert raters). Fit is, thus, operationalized as a correlation coefficient.

In many cases, multiple raters might provide profiles of either the individual or the situational variables. When multiple raters act as informants, the vector of scores can be calculated by averaging across their responses. The consistency of raters' responses can be assessed by a variation of the Spearman–Brown general prophecy formula (e.g., Chatman, 1989). In the case of profile data, this coefficient is interpreted as evidence of agreement among raters rather than evidence of the underlying construct and specifically reflects the stability of the profile. An intraclass correlation coefficient, calculated as the median correlation coefficient among all pairs of raters, can also be used to assess agreement among raters on the focal individual or situation (e.g., Block, 1978; Kenny, Albright, Malloy, & Kashy, 1994).

Using Profile Data

The semi-idiographic nature of profiles allows comparison between individuals yet also provides a fine-grained picture of both people and situations. People's relative fit to a situation can be assessed by comparing the magnitude of the correlation coefficients between each individual's profile and a common profile of the situation. In this sense, correlation coefficients reflect the overall fit of an individual to a situation, relative to other individuals, across a wide range of dimensions.

The idiographic nature of the profile can provide a rich, clinical picture of the individual or the situation. In a typical profile study, individuals who are familiar with the situation, either through their tenure, experience, or expertise, would sort items about the situation. Similarly, raters who were familiar with the focal individual (either him or herself or observers) would sort the same items to describe the person. This eliminates common-response bias, and it also allows the use of true "experts" who are likely to be distinctly familiar with either the person or the situation, but not both. In addition, inspecting the rankings of the items and clustering of related items can provide an interpretative portrait of the both the person and situation. Differences between the person and situation across these clusters of items can provide insights into why a person might fit or not fit a particular situation.

Although the semi-idiographic nature of profile data offers advantages in studying fit, there is no guarantee that profile methodology will provide a more accurate test of a specific hypothesis than a completely nomothetic approach. The ipsative nature of profile data limits the use of some statistical techniques. Further, the correlation-based fit score

provides an overall measure of fit, but no statistical test of the nature and significance of mismatches between individual variables and situational variables exists and researchers have relied on judgment to determine what constitutes relative fit or misfit. For example, two individuals could have the same overall fit score yet in one case it was because the individual was higher on some individual variables than called for in the situation and the other individual was lower on those same items. Although inspecting the profiles could provide insights regarding those differences, the ipsative nature of the data make some statistical tests problematic.

The use of difference scores to interpret profile data has been criticized because of their potential unreliability (e.g., Edwards, 1995). Johns (1981), however, listed a number of ways in which difference scores can be reliably developed and assessed including using commensurate terms, which we strongly advocate, and using different raters to create the situational and person profiles. Further, although Edwards' (1995) critique raised a number of valid points, profile comparison approaches remain valuable tools for fit research because they allow researchers to answer important and different types of research questions that take into account the multidimensional nature of people and situations. Profile approaches enable researchers to derive a single index that simultaneously captures fit across multiple dimensions, thereby making it possible to test notions regarding the importance of overall fit, as well as relationships between overall fit and various behavioral and attitudinal responses in organizations.

11.3 POLYNOMIAL REGRESSION AND RESPONSE SURFACE METHODOLOGY

Jeffrey R. Edwards

Person–environment (PE) fit research often relies on methods that collapse *P* and *E* measures into a single score intended to represent PE fit. Typically, these methods involve computing the difference between person and environment measures or the similarity between profiles of measures that describe person and environment on multiple dimensions. These methods suffer from numerous methodological problems, as documented elsewhere (Cronbach, 1958, 1992; Edwards, 1994; Johns, 1981). Problems with difference scores and profile similarity indices are avoided by polynomial regression (Edwards, 1994, 2002; Edwards & Parry, 1993), which uses separate measures of *P* and *E* and examines their joint relationships with causes and consequences of PE fit. Polynomial regression is based on the premise that *P* and *E* measures represent distinct constructs and the assumptions embedded in difference scores and profile similarity indices represent hypotheses that should be tested empirically.

In this section polynomial regression and its relevance to P-E fit research are discussed. The section has three objectives: (a) to show how polynomial regression can be viewed as a generalization of difference scores and profile similarity indices; (b) to explain how results from polynomial regression analyses can be understood using response surface methodology; and (c) to emphasize that polynomial regression and response surface methodology can facilitate theory development in PE fit research. This overview focuses on fit as a cause of outcomes, and procedures for treating fit as an outcome are briefly discussed at the end of the summary.

Polynomial Regression as a Generalization of Difference Scores and Profile Similarity Indices

The basics of polynomial regression can be understood by contrasting it with difference scores and profile similarity indices. To illustrate, consider Fig. 11–1a, which depicts a positive relationship between an algebraic difference score and an outcome. This relationship can be represented by the following regression equation:

$$Z = b_0 + b_1(X - Y) + e, \tag{1}$$

where X is the environment, Y is the person, Z is the outcome, and e is a random disturbance term. The positive relationship in Fig. 11–1a corresponds to a positive value for b_1 in Equation 1. The connection between Equation 1 and polynomial regression can be seen by expanding Equation 1, which yields

$$Z = b_0 + b_1X - b_1Y + e. \tag{2}$$

Equation 2 shows that using an algebraic difference as a predictor is equivalent to using the components of the difference as predictors and constraining their coefficients to be equal in magnitude and opposite in sign. The relationship of X and Y with Z indicated by Equation 2 is illustrated by the three-dimensional surface in Fig. 11–2a. The constraint imposed by Equation 2 can be empirically tested using the following equation:

$$Z = b_0 + b_1X + b_2Y + e. \tag{3}$$

Equation 3 is a linear polynomial regression equation in which the relationships of X and Y with Z can differ in sign and magnitude. Results from Equation 3 can be used to determine whether $b_1 = -b_2$, as indicated by Equation 2, and whether b_1 is positive and b_2 is negative, as implied by Fig. 11–2a.

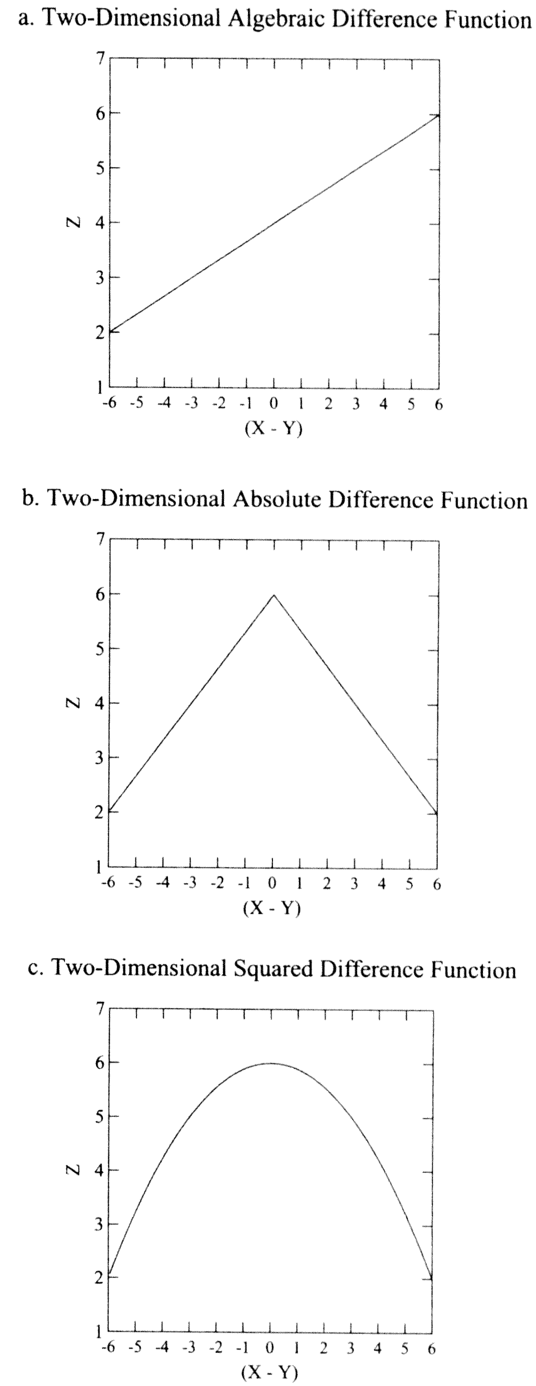

FIGURE 11–1. Two-dimensional difference score functions.

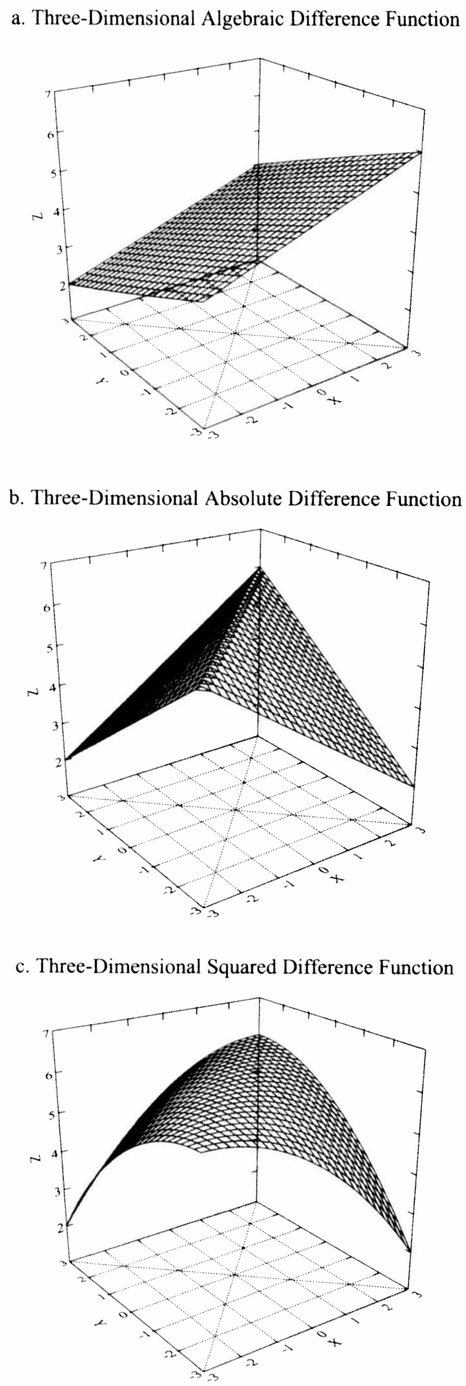

FIGURE 11–2. Three-dimensional difference score surfaces.

Figure 11–1b shows an inverted-V relationship, such that the outcome is maximized when X and Y are equal. The relationship in Fig. 11–1b is captured by the following regression equation:

$$Z = b_0 + b_1|X - Y| + e. \tag{4}$$

Equation 4 uses the absolute value of the difference between X and Y as a predictor of Z. A negative value for b_1 would correspond to the inverted-V relationship in Fig. 11–1b. As written, Equation 4 cannot be algebraically expanded, because an absolute difference is a logical rather than an algebraic transformation. This dilemma is overcome by replacing Equation 4 with the following equation (Edwards, 1994):

$$Z = b_0 + b_1(1 - 2W)(X - Y) + e. \tag{5}$$

In Equation 5, W is a dummy variable that equals 0 when $X - Y$ is positive, equals 1 when $X - Y$ is negative, and is randomly set to 0 or 1 when $X = Y$. As a result, when $X - Y$ is positive, $1 - 2W$ equals 1, and the compound term $(1 - 2W)(X - Y)$ reduces to $(X - Y)$. In contrast, when $X - Y$ is negative, $1 - 2W$ equals -1, and the term $(1 - 2W)(X - Y)$ becomes $-(X - Y)$. Hence, when $X - Y$ is positive, its sign is unaltered, and when $X - Y$ is negative, its sign is reversed, producing the same result as the absolute value transformation. When $X = Y$, $(1 - 2W)(X - Y)$ equals 0 regardless of the value of W. Expanding Equation 5 yields

$$Z = b_0 + b_1X - b_1Y - 2b_1WX + 2b_1WY + e. \tag{6}$$

Equation 6 shows that using an absolute difference as a predictor effectively constrains the coefficients on X and Y to be equal in magnitude but opposite in sign, constrains the coefficients on WX and WY to be equal in magnitude but opposite in sign, and constrains the coefficient on WY to be twice as large as the coefficient on X. Equation 6 also indicates that the coefficient on W is constrained to 0, given that it is excluded from the equation. Fig. 11–2b shows a three-dimensional surface that corresponds to Equation 6. The constraints imposed by Equation 6 can be tested with the following equation:

$$Z = b_0 + b_1X + b_2Y + b_3W + b_4WX + b_5WY + e. \tag{7}$$

Equation 7 is a piecewise polynomial regression equation. Coefficient estimates from this equation can be used to assess the constraints imposed by Equation 6 by testing whether (a) $b_1 = -b_2$, (b) $b_4 = -b_5$, (c) $b_5 = 2b_1$, and (d) $b_3 = 0$. The direction of the inverted-V relationship in Fig. 11–2b further stipulates that b_1 and b_5 are positive and that b_2 and b_4 are negative.

Figure 11–1c shows an inverted-U relationship. Like Fig. 11–1b, Fig. 11–1c indicates that the outcome is maximized when X and Y are equal. However, as the difference between X and Y increases, the outcome decreases linearly in Fig. 11–1b, as opposed to the curvilinear decrease in Fig. 11–1c. The relationships shown in Fig. 11–1b and c are usually treated the same in PE fit research. This relationship in Fig. 11–1c corresponds to the following regression equation:

$$Z = b_0 + b_1(X - Y)^2 + e. \tag{8}$$

Equation 8 uses the squared difference between X and Y as a predictor of Z. For the relationship in Fig. 11–1c, the sign of b_1 in Equation 8 would be negative. Expanding Equation 8 yields:

$$Z = b_0 + b_1X^2 - 2b_1XY + b_1Y^2 + e. \tag{9}$$

Equation 9 indicates that using a squared difference as a predictor constrains the coefficients on X^2 and Y^2 to be equal and the coefficient on XY to be twice as large in magnitude and opposite in sign of the coefficient on X^2 or Y^2. Equation 9 also implicitly constrains the coefficients on X and Y to be 0, given that both of these variables are excluded from Equation 9. A three-dimensional surface corresponding to Equation 9 is shown in Fig. 11–2c, and the four constraints imposed by Equation 9 can be tested with the following equation:

$$Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 + e. \tag{10}$$

Equation 10 is a quadratic polynomial regression equation. Coefficient estimates from Equation 10 can be used to evaluate the constraints imposed by Equation 9 by testing whether (a) $b_1 = 0$; (b) $b_2 = 0$; (c) $b_3 = b_5$; and (d) $b_4 = -2b_3$. The direction of the inverted-U relationship in Fig. 11–2b further indicates that b_3 and b_5 are negative and b_4 is positive.

The logic used to translate difference scores into polynomial regression equations can be applied to profile similarity indices (Cronbach & Gleser, 1953; Edwards, 1993, 1994). For example, profile similarity indices that represent sums of algebraic, absolute, or squared differences can be written by extending Equations 2, 6, and 9, respectively, to include multiple pairs of X and Y measures in which each pair represents a dimension on which the profiles are compared. The constraints imposed by these indices can be tested using extended versions of Equations 3, 7, and 10. Euclidean distance and profile correlation indices, which are commonly used in PE fit research, cannot be algebraically expanded, but the conceptual principles they are intended to capture can be examined using unconstrained regression equations for the sums of absolute or squared differences (Edwards, 1993).

Applying Response Surface Methodology to Polynomial Regression Analysis

When polynomial regression yields coefficients that satisfy the constraints associated with Equations 2, 6, and 9, results are easily interpreted because they conform to the idealized surfaces shown in Fig. 11–2. However, these constraints are usually rejected, which complicates the interpretation of results. Furthermore, the surfaces in Fig. 11–2 comprise a narrow subset of hypotheses that could be developed regarding the joint effects of the person and environment on outcomes. For example, outcomes produced by PE misfit may differ depending on whether the environment is greater than or less than the person (Edwards, Caplan, & Harrison, 1998; Naylor, Pritchard, & Ilgen, 1980; Rice, McFarlin, Hunt, & Near, 1985). In addition, the effects of PE fit may depend on whether the person and environment are both low or high in an absolute sense (Edwards & Rothbard, 1999). Complexities such as these are important from a theoretical standpoint but are not captured by the surfaces in Fig. 11–2.

The foregoing issues can be addressed using response surface methodology (Edwards & Parry, 1993), which allows researchers to rigorously analyze three-dimensional surfaces relating the person and environment to outcomes. Response surface methodology facilitates substantive interpretation when constraints imposed by difference scores are rejected, as is usually the case. Perhaps more importantly, response surface methodology allows PE fit researchers to develop and test hypotheses that go far beyond the simplified surfaces shown in Fig. 11–2.

Response surface methodology involves analyzing features of surfaces that correspond to polynomial regression equations. The quadratic equation in Equation 10 captures a wide range of hypotheses relevant to PE fit research and is therefore the focus of the present discussion. A quadratic equation reflects one of three types of surfaces: (a) *concave*, which is dome-shaped; (b) *convex*, which is bowl-shaped; and (c) *saddle*, which is shaped like a saddle. For each surface, response surface methodology involves the analysis of three basic features.

The first feature is the stationary point, which is the point at which the surface is flat. For a concave surface, the stationary point is the overall maximum of the surface. For a convex surface, the stationary point is the overall minimum of the surface. For a saddle surface, the stationary point is the intersection of the lines along which the upward and downward curvatures of the surface are greatest. The location of the stationary point can be computed by inserting the estimated coefficient values from a quadratic regression equation into the following formulas:

$$X_0 = \frac{b_2b_4 - 2b_1b_5}{4b_3b_5 - b_4^2} \tag{11}$$

$$Y_0 = \frac{b_1b_4 - 2b_2b_3}{4b_3b_5 - b_4^2} \tag{12}$$

where X_0 and Y_0 are the coordinates of the stationary points in the X,Y plane.

The second feature involves the *principal axes,* which describe the orientation of the surface in the X,Y plane. The principal axes run perpendicular to one another and intersect at the stationary point. For a concave surface, the first principal axis is the line of minimum downward curvature, and the second principal axis is the line of maximum downward curvature. For a convex surface, the first principal axis is the line of maximum upward curvature, and the second principal axis is the line of minimum upward curvature. Finally, for a saddle surface, the first principal axis is the line of maximum upward curvature, and the second principal axis is the line of maximum downward curvature.

The principal axes can be written as equations that describe lines in the X,Y plane. An equation for the first principal axis is

$$Y = p_{10} + p_{11}X. \tag{13}$$

The slope of the first principal axis (i.e., p_{11}) is computed as follows:

$$p_{11} = \frac{b_5 - b_3 + \sqrt{(b_3 - b_5)^2 + b_4^2}}{b_4} \tag{14}$$

The intercept of the first principal axis (i.e., p_{10}) is computed as follows:

$$p_{10} = Y_0 - p_{11}X_0. \tag{15}$$

Likewise, an equation for the second principal axis is

$$Y = p_{20} + p_{21}X. \tag{16}$$

The slope of the second principal axis (i.e., p_{21}) is computed using the following formula:

$$p_{21} = \frac{b_5 - b_3 - \sqrt{(b_3 - b_5)^2 + b_4^2}}{b_4} \tag{17}$$

The intercept of the second principal axis (i.e., p_{20}) is computed as follows:

$$p_{20} = Y_0 - p_{21}X_0. \tag{18}$$

The third feature entails the shape of the surface along relevant lines in the X,Y plane, which can be computed by substituting the equation for the line into Equation 10. To illustrate, the $Y = X$ line is meaningful to PE fit research because it represents values for which the person and environment are equal. For the surface in Fig. 11–2c, this line runs diagonally across the floor of the graph from the near corner to the far corner. Substituting $Y = X$ into Equation 10 yields

$$\begin{aligned} Z &= b_0 + b_1X + b_2X + b_3X^2 + b_4X^2 + b_5X^2 + e \\ &= b_0 + (b_1 + b_2)X + (b_3 + b_4 + b_5)X^2 + e. \end{aligned} \tag{19}$$

As Equation 19 shows, the curvature of the surface along the $Y = X$ line is represented by the sum $b_3 + b_4 + b_5$, and the slope of the surface at the point $X = 0$ (and $Y = 0$, given that $Y = X$) is $b_1 + b_2$. If these sums equal zero, then the surface is flat along the $Y = X$ line, consistent with the surface in Fig. 11–2c.

Another line of interest in PE fit research is the $Y = -X$ line. When X and Y measures are centered at the midpoint of their scales, as recommended for polynomial regression analysis (Edwards, 1994; Edwards & Parry, 1993), the $Y = -X$ line runs diagonally across the X,Y plane and represents varying degrees of PE misfit. In Fig. 11–2c, the $Y = -X$ line extends from the left corner to the right corner of the floor of the graph. The shape of the surface along this line represents the effect of PE misfit. Substituting $Y = -X$ into Equation 10 yields

$$\begin{aligned} Z &= b_0 + b_1X - b_2X + b_3X^2 - b_4X^2 + b_5X^2 + e \\ &= b_0 + (b_1 - b_2)X + (b_3 - b_4 + b_5)X^2 + e. \end{aligned} \tag{20}$$

The curvature of the surface along the $Y = -X$ line equals $b_3 - b_4 + b_5$, and the slope of the surface when $X = 0$ (and $Y = 0$, given than $Y = -X$) equals $b_1 - b_2$. For the surface in Fig. 11–2c, $b_3 - b_4 + b_5$ is negative and $b_1 - b_2$ equals 0, meaning that the surface has a downward curvature along the $Y = -X$ line and is flat at $X = 0$, $Y = 0$.

The shape of the surface along other lines can be determined in a similar manner. For instance, the shape of the surface along the first principal axis is found by substituting Equation 13 into Equation 10, which yields

$$\begin{aligned} Z &= b_0 + b_1X + b_2(p_{10} + p_{11}X) + b_3X^2 \\ &\quad + b_4X(p_{10} + p_{11}X) + b_5(p_{10} + p_{11}X)^2 + e \\ &= b_0 + b_2p_{10} + b_5p_{10}^2 + (b_1 + b_2p_{11} + b_4p_{10} \\ &\quad + 2b_5p_{10}p_{11})X + (b_3 + b_4p_{11} + b_5p_{11}^2)X^2 + e. \end{aligned} \tag{21}$$

Likewise, the shape of the surface along the second principal axis is

$$\begin{aligned} Z &= b_0 + b_1X + b_2(p_{20} + p_{21}X) + b_3X^2 \\ &\quad + b_4X(p_{20} + p_{21}X) + b_5(p_{10} + p_{11}X)^2 + e \\ &= b_0 + b_2p_{20} + b_5p_{20}^2 + (b_1 + b_2p_{21} + b_4p_{20} \\ &\quad + 2b_5p_{20}p_{21})X + (b_3 + b_4p_{21} + b_5p_{21}^2)X^2 + e. \end{aligned} \tag{22}$$

This procedure can be extended to other lines of theoretical interest.

Empirical Example

Figure 11–3 depicts a response surface based on data from 358 job seekers who reported the actual variety, desired variety, and satisfaction associated with jobs for which they had recently interviewed. The estimated polynomial regression was

$$Z = 5.628 + 0.314X - 0.118Y - 0.145X^2 + 0.299XY - 0.102Y^2 + e. \tag{23}$$

The R^2 for the equation was .162, and coefficients for all variables except Y and Y^2 were statistically significant at $p < .05$. The corresponding surface is shown in Fig. 11–3. The stationary point was located at $X = -0.951$, $Y = -1.973$. The first principal axis had an intercept and slope of –0.875 and 1.154, respectively, and is represented by the solid line crossing the X,Y plane. The second principal axis had an intercept and slope of –2.797 and –0.866, respectively, and is depicted by the heavy dashed line in the X,Y plane. Along the $Y = X$ line, the surface had a curvature of 0.052 and a slope of 0.196 at the point $X = 0$, $Y = 0$, indicating that satisfaction increased at an increasing rate. Along the $Y = -X$ line, the surface had a curvature of –0.546 and a slope of 0.432 at the point $X = 0$, $Y = 0$. These results indicate that satisfaction increased as actual variety increased toward desired variety, continued to increase at the point where actual and desired variety were equal (i.e., $X = 0$, $Y = 0$), and began to decrease when actual variety exceeded desired variety by about 0.5 units, as indicated by the point where the $Y = -X$ line crossed the first principal axis.

Comparing the surface in Fig. 11–3 to Fig. 11–2c highlights two key findings revealed by polynomial regression and response surface methodology that would have been missed by using a squared difference score. First, Fig. 11–3 shows that, along the line of PE fit, satisfaction is higher when actual and desired variety are both high rather than low, whereas in Fig. 11–2c, satisfaction is forced to remain constant along the PE fit line. The increase in satisfaction as actual and desired variety increase makes sense from a conceptual standpoint, given that jobs with higher variety

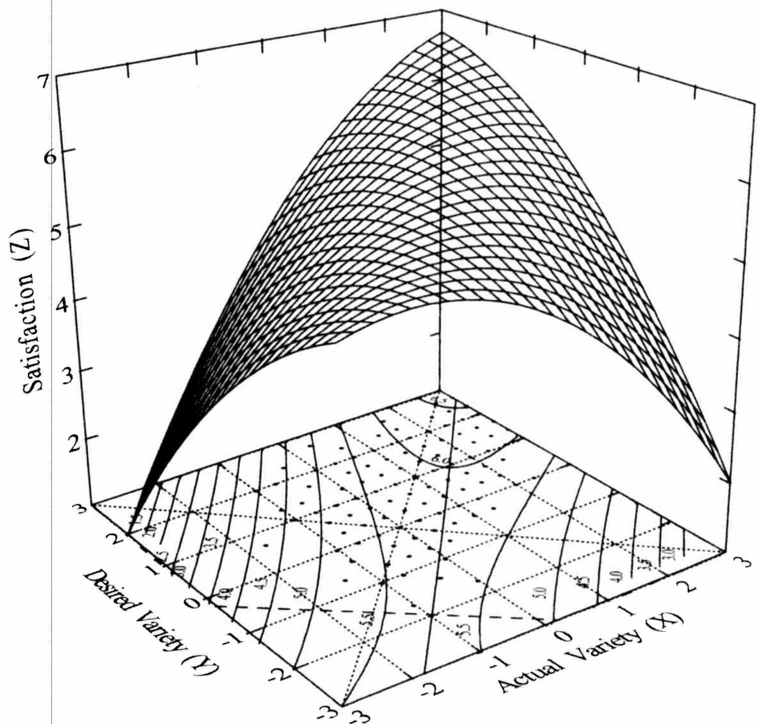

FIGURE 11–3. Response surface analysis relating actual and desired variety to satisfaction.

often bring additional rewards such as autonomy and challenge, and people who value variety are also likely to value these rewards. Second, Fig. 11–3 shows that, along the line of PE misfit, satisfaction is greatest when actual variety exceeds desired variety, whereas in Fig. 11–2c, satisfaction is constrained to reach its maximum when actual and desired variety are equal. Again, the fact that satisfaction is greatest when actual variety exceeds desired variety makes sense, because a moderate excess of variety brings opportunities for challenge and self-development, which can increase overall satisfaction with the job.

The Theoretical Value of Polynomial Regression and Response Surface Methodology

The incremental insights yielded by polynomial regression and response surface methodology are important for interpreting results as well as for developing theory. By conceptualizing the effects of PE fit as a three-dimensional surface rather than a two-dimensional function, an enormous range of hypotheses can be pursued. These hypotheses can address asymmetries in the effects of PE misfit, variation in outcomes along the line of PE fit, surface rotations indicating that the optimal combination of P and E depends on whether both are high or low, and so forth (Edwards, 1996; Edwards & Rothbard, 1999). Traces of hypotheses such as these are evident in the PE fit literature (Edwards et al., 1998; Rice et al., 1985), but efforts to develop and test them have been hampered by the use of difference scores. Polynomial regression and response surface methodology encourage PE fit researchers to adopt an expanded conceptualization of PE fit that opens new avenues for theory and research.

11.4 HETEROGENEITY AS MISFIT

David A. Harrison and Hock-Peng Sin

Just as similarity to the salient features of a person's job (person–job [PJ] fit) or organization (person–organization [PO] fit) can be regarded as supplementary fit, so can similarity to the salient features of others in a person's workgroup (person–group (PG) fit) or dyad (person–person (PP) fit). Many theories predict that fit, interpersonal similarity, or homogeneity, has positive attitudinal and behavioral consequences (Byrne, 1971; Newcomb, 1961; Schneider, 1987). In the obverse, interpersonal differences or heterogeneity can be construed as a form of misfit, which is expected to have negative outcomes for teams (Harrison, Price, Gavin, & Florey, 2002) or for a focal person, P, within them (Chatman, 1991). To the extent that team members are dissimilar from one another, they may feel less psychologically connected or attached to a collective. Thus, heterogeneity as misfit is an unipolar construct, in that it ranges from zero (perfect fit or perfect similarity—P is a clone of another P, or of the rest of the workgroup G) to some large amount (strong PP or PG differences). As such, misfit can be conceptualized and operationalized at both the team level (aggregation of all possible PP pairs of misfit within a team) and at the individual level (a measure of PG fit).

Misfit at the Team Level

Investigations of misfit at the team level are best represented in the burgeoning domain of research on workgroup diversity, although the diversity terminology tends not to overlap with fit research. In the diversity paradigm, researchers first assess the individual-level attribute in question and then compute one of several indices of the *dispersion* of those individual attributes within the group (see Table 11–1 and Harrison & Sin, 2006, for a detailed review). Investigators of demographic diversity often assess members' attributes such as sex, race, age, tenure, and so on, and then compute a team level diversity value depending on the scale of

TABLE 11-1.
Common Diversity Indices Useful for Operationalizing Interpersonal Misfit

Index	*Formula*	*Theoretical Min/Max*	*Operational Min/Max*	*Assumed Scale Level of X*
Blau[a]	$-\Sigma[\text{prop}_k]2$	0 to 1	0 to $(k-1)/k$	Categorical/ordinal
Teachman[a]	$-\Sigma[\text{prop}_k] \times \ln[\text{prop}_k]$	0 to $+\infty$	0 to $-1 \times \ln(1/k)$	Categorical/ordinal
Coefficient of variation[a]	$\text{sqrt}[\Sigma(X_i - X_{\text{mean}})^2/n]/X_{\text{mean}}$	0 to $+\infty$	0 to $\text{sqrt}(n-1)$	Ratio
Gini[a]	$(\Sigma\lvert X_i - X_j\rvert)/2 \times n^2 \times X_{\text{mean}}$	0 to +1	0 to $1-(1/n)$	Ratio
Standard deviation[a]	$\text{sqrt}[\Sigma(X_i - X_{\text{mean}})^2/n]$	0 to $+\infty$	0 to $[(U-L)/2]$	Interval
Euclidean distance[b]	$\text{sqrt}[\Sigma_j(X_i - X_j)^2/n]$	0 to +1 0 to $+\infty$	0 to $\text{sqrt}[(n-1)/n]$ 0 to $\text{sqrt}[(n-1)(U-L)^2/n]$	Categorical/ordinal Interval

Note. n is the number of members in group or team; k is the number of categories in a categorical variable, or ranks in an ordinal variable; prop_k is the proportion of team members in the kth category; i and j are the ith and jth persons in a team; U is the upper endpoint of a finite measurement scale; and L is the lower endpoint (e.g., $U = 5$ and $L = 1$ for a Likert-type scale).
[a]Team level.
[b]Individual level.

measurement for the attribute in question. For categorical attributes such as sex or ethnicity, Blau's and Teachman's indices are commonly used (e.g., Bunderson & Sutcliffe, 2002). For interval or ratio variables, the standard deviation (SD), Gini index, or coefficient of variation (CV) is more appropriate (e.g., Klein, Conn, Smith, & Sorra, 2001). Those who study psychological diversity derive team-level measures based on members' attitudes (e.g., Harrison, Price, & Bell, 1998), personality dimensions (e.g., Barrick, Stewart, Neubert, & Mount, 1998), or values (Barsade, Ward, Turner, & Sonnenfeld, 2000), following the same match of measurement scale to diversity (misfit) index. All such indices, whether for PP or PG fit, are minimized at zero, the point at which everyone in the group is identical on the attribute in question.

Maximum values are a different story. For categorical variables, values of diversity or heterogeneity indices (e.g., Blau's and Teachman's) are maximized when equal portions of a team make up each of the k possible classifications in the category system. A little-recognized property of the categorical indices is that they reach different maxima under different values of k, and hence, cannot be compared when different attributes have different numbers of categories. For instance, Blau's index for a team with maximum sex diversity (two categories and 50% male, 50% female) is .50. If members are instead, evenly spread across five possible ethnic categories, Blau's index is maximized at .80. If homogeneity is fit and heterogeneity is misfit, then the maximum value for misfit among team members does not approach the same number with different categorical variables.

In a similar vein, diversity or misfit indices for interval and ratio attributes take maximum values under quite different within-team distribution shapes. A sharply positively skewed distribution within a team will yield a larger value of CV and the Gini index. Indeed, both of these indices, when used as measures of misfit, are maximized when a single person in a team is at the highest point, and everyone else shares the lowest point (Allison, 1978). For example, if one team member had 10 years of tenure, and her three teammates were brand new, CV would be maximized at 1.73). On the other hand, indices such as SD will yield values that are increasingly large as the positions of team members approach a bimodal distribution (Harrison & Sin, 2006; note that here and elsewhere, researchers are not estimating a population parameter and therefore should use n the total number of team members in the denominator of the SD formula, rather than the conventional $n - 1$). In the aforementioned skewed example, SD is 4.33. But, if two team members had 10 years of tenure each and the other two members were brand new, SD would be 5.00. Note that in this latter case, CV is only 1.00, less than the original, positively skewed distribution. Hence, CV (and its variations) versus SD tell us substantially different stories about misfit at the team level.

CV and the Gini index also have other somewhat unusual properties. Their maximum values depend on (the square root of) team size and, hence, are not comparable across teams with different ns. If, in the aforementioned example, the team had seven brand new members and one member with 10 years of tenure, the CV would be 2.65 (SD = 3.31). In the bimodal case for an eight-person team, however, the CV remains at 1.00 (and the SD remains at 5.00), identical to the values for the four-person team. Reversing the scores within the team will produce substantially different CV and Gini values. With seven veteran team members and only one new member, the CV is 0.38. On the other hand, SD is symmetric; it remains at 3.31 in the reversed case. In our estimation, these latter properties work against the use of Gini or CV as misfit indices, unless a theory of fit clearly specifies that maximum interpersonal misfit occurs when one member has an excess of an attribute relative to all other members in a group.

Misfit at the Individual Level

Misfit can also be assessed at the individual level, by examining how different a person is relative to one (e.g., a supervisor) or all of the other persons within a team, and treating that difference as a distance. A typical operationalization is the Euclidean distance formula between P and another person (PP, dyadic distance) or the average Euclidean distance between P and his or her counterparts (mean PG distance; see Tsui, Egan, & O'Reilly, 1992; Tsui & O'Reilly, 1989). This index can be used for both categorical and continuous variables. For categorical variables, (a) the Euclidean distance between any two different categories is considered to be 1, (b) the distance between those who share a category is 0, and (c) the values of the index can theoretically range from 0 to +1, although the upper bound depends operationally on team size (see Table 11–1). Its maximum value occurs when the individual is different from everyone else, as would be the case if someone was the only female in a four-person team. Her Euclidean distance would be the square root of $[(1 - 0)^2 + (1 - 0)^2 + (1 - 0)^2]/n$ or sqrt(.75) = .87. For continuous variables, the values of a Euclidean distance index can theoretically range from 0 to $+\infty$, although the upper bound depends operationally on both team size and the endpoints of the instrument used to measure the attribute (Harrison & Sin, 2006). Using the aforementioned example, if the female member was the 10-year veteran in the group, her Euclidean distance from the other members on tenure would be $[(10 - 0)^2 + (10 - 0)^2 + (10 - 0)^2]/n$ or sqrt(300/4) = 8.67. On the other hand, each of the rookies is identical to all of the other rookies, but different from her. Each rookie's Euclidean distance would be $[(0 - 0)^2 + (0 - 0)^2 + (10 - 0)^2]/n$ or sqrt(100/4) = 5.00.

An interesting feature of Tsui et al.'s (1992) formula, as shown in Table 11–1, is that their computation of Euclidean distance depends on team size. To gauge one's distance or misfit from others, Tsui et al. (1992) recommend dividing by n (all team members) instead of $n - 1$ (all other team members) to "derive a metric that captures both the size and the compositional effects" (p. 562). For example, being the only female in a three-, five-, seven-, or nine-person team will yield Euclidean distances or PG misfit indices of .82, .89, .93, and .94, respectively. The rationale behind this computational subtlety, that the psychological impact of misfit/dissimilarity varies nonlinearly as a function of team size, is an empirical question that may warrant investigation in itself. That is, to address the assumption embedded in the individual-level misfit index, one only needs to include team size as a statistical control variable or moderator.

Signed versus Unsigned Distances in Person–Team Misfit

Note also that Euclidean distances are symmetric. This would be appropriate for variables that are "lateral," "horizontal," or status-free, such as attitudes or personality. If there is psychological meaning in person P having higher or lower levels of resources or status than the other team members, then misfit is directional. For example, a team member who is *superior* (has excess relative to others) in cognitive ability might receive a greater amount of recognition and reward, whereas a team member who is *inferior* (is deficient relative to others) might be ignored. A similar degree of absolute PG misfit could lead to very different outcomes. Fit researchers can construct a per-individual mean of the signed Euclidean distances to capture such a phenomenon.

Cautionary Note on Dimensionality

A final caveat deals with the multidimensional nature of fit. Can a single number reflect a team's or person's *overall* misfit? We believe that it would be at best imprudent and at worst misleading. The logic underlying such aggregation assumes either a latent or emergent construct based on positive correlations among the constituent elements of fit (Harrison, 2002). For instance, if one wanted to create a "total demographic misfit" index, measurement logic would require that diversity in ethnicity, age, sex, and so on would be positively related to one another. The composite would then represent the shared or overlapping components of the diversity variables. However, this is rarely the case. More importantly, there is often no conceptual basis for expecting different dimensions of interpersonal misfit to be associated with each other (especially for demographic attributes). Perhaps a better approach would be to treat them as a multivariate set rather than adding them together to reach a less meaningful total score (Lau & Murnighan, 1998, treat them interactively). We contend that interpersonal misfit is most meaningful when it is more narrowly defined or dimensionalized.

Conclusion

Overall, the empirical treatment of misfit can appropriate some of the ideas from research in the area of workgroup diversity to operationalize (collective) PP and PG misfit. Likewise, some of the commonly used indices in those areas, relational demography, and supervisor–subordinate differences can be adopted for the study of misfit. Specifically, we recommend Blau's and Teachman's indices for team-level misfit on categorical variables, SD for team-level misfit on continuous variables, and Euclidean distances for PP and PG misfit at the individual level. When the direction of misfit is important (when interpersonal superiority or inferiority on an attribute is meaningful), signed distances are reasonable. Finally, we note that the team indices discussed here could also be used at higher levels of analysis (e.g., division or organizational level) to capture the extent of homogeneity (fit) or heterogeneity (misfit) on demographic and psychological variables at these higher levels.

11.5 CLUSTER ANALYSIS

Michael D. Mumford and Jazmine Espejo

The term *cluster analysis* refers to a set of statistical techniques, really decision rules, whereby entities, people, objects, or environments are grouped into a smaller set of "types." In cluster analytic studies, inferences about fit are made with respect to the group the individual is similar to, or a member of, on the basis of cross-group differences in relation to performance on measures held to reflect fit effects. This approach has proven attractive in studies of person (P)–environment (E) fit for two reasons (Mumford, Stokes, & Owens, 1990; Owens & Schoenfeldt, 1979). First, it permits a wide range of different types of variables to be considered simultaneously in fit assessments. Second, clustering procedures allow for qualitative differences and nonlinear effects.

Cluster analysis has been used in studies of PE fit for almost half a century (cf., Owens, 1968). In essence, cluster analysis captures the particular profile or pattern across attributes. For example, individuals could be "typed" or grouped into categories based on their patterns or profiles across measures of their ability and desires (e.g., one profile might reflect high ability, a strong desire for autonomy, and a weak desire for job variety, whereas another pattern could reflect high ability, strong autonomy, and strong job variety desires). Cluster analysis has been used to classify persons based on the pattern of their different P attributes or the work context based on different E characteristics, and the different types derived from the cluster analysis have been linked to various outcomes. In other studies (e.g., Gustafson & Mumford, 1995), independent P and E patterns have been derived and then have been considered simultaneously in relation to outcomes.

One-way cluster analysis is used in studies of PE fit is to identify types, or subsets, of people who display similarity on measures of personality, life history, or values and interests and then to examine cross-type differences with respect to performance in a given work environment. For example, Mumford, Zaccaro, Johnson, Diana, and Threlfall (2000) identified seven types of leaders commonly found on army jobs with only three of these types advancing to and performing well in more senior leadership positions. Another way cluster analysis has been applied in studies of PE fit is to identify the major kinds of psychosocial environments to which people must adapt. One example of this application may be found in Hofstede (1998), who used climate inventory scores to identify the major types of work environments operating within a particular firm. Still another way cluster analysis is used in studies of PE fit is to assess how different types of people differ in their reactions to, or performance in, different types of work environments. In one study along these lines, Gustafson and Mumford (1995) found that some types of people performed well in multiple types of work environments, whereas other types of people performed well in just one type of work environment.

Cluster Analysis Methods

Clustering techniques have not been widely applied in PE fit due to the complexity of the methods involved and certain vagaries surrounding requisite methodological operations—an issue we hope to help resolve in the present section. To begin, cluster analysis is not a single method but rather a family of methodological techniques subsuming single linkage analysis, centroid analysis, and multidimensional scaling among other techniques (see Anderberg, 1973; Everitt, 1979, for a more complete description of basic methods).

Despite these procedural differences, all cluster analysis techniques apply a similar general strategy to grouping entities into types. Essentially, entities (people, environments, and others) are measured on a set of variables and the profile of variable scores is obtained. Then, the similarity of each entities' profile to all other profiles is assessed using a similarity metric—typically a distance metric or correlation coefficient. Once similarity has been assessed, decision rules are applied to metric scores to determine the entities, or entity clusters, that can be merged together at any given point in an iterative sequence. Entities with similar profiles are merged together into a type that can be distinguished from other types. Single linkage methods cluster entities based on the distance between the nearest neighbors in relevant clusters, whereas centroid methods, such as the Ward and Hook (1963) procedure, cluster entities to minimize within-group variation around the cluster centroid or mean profile.

In hierarchical methods, the most common kind of cluster analysis, the sequence of combinations begins by treating each entity as a cluster in and of itself and continues until all entities have been merged into a single cluster. The number of clusters to be retained is determined by identifying the point in this iterative sequence at which further combinations result in sharp increases in within-cluster differences. Hierarchical procedures are used to identify the number and nature of the clusters to be retained in a population when no prior knowledge exists about likely subpopulations.

Nonhierarchical procedures, such as *k*-means, begin with a priori seed points. As a result, the accuracy of nonhierarchical analyses depends on initial specification of these seed points. Frequently in clustering studies, a nonhierarchical analysis follows a hierarchical analysis in which the results of the hierarchical analysis provide the seed points used in the nonhierarchical analysis. The stability of entity assignments across these two procedures is used to assess replicability in assignments and control for drift due to the retention of early sequence assignees within clusters. Additionally, nonhierarchical procedures allow fit to be assessed to idealized type profiles, permitting theory-based clustering or appraisals of entity fit to standards (e.g., fit of an individual's climate perception to the climate perceptions of different groups).

This thumbnail sketch of the basic procedures applied in cluster analysis not only describes the general method but also points to the key methodological considerations likely to appear in any cluster analytic study of PE fit. In the ensuing discussion, we will address these six major methodological issues: (a) sampling, (b) measurement, (c) similarity assessment, (d) choice of clustering procedure, (e) evaluation of cluster solutions, and (f) applications of clusters in PE fit studies.

Sampling

Traditionally, it has been assumed that cluster analytic investigations require rather large samples. A Monte Carlo study by Overall and Magee (1992), however, indicated that clustering procedures can adequately capture a true underlying subpopulation structure if as few as 8–10 entities from the relevant subpopulations are represented in the sample. Thus, if only 5 clusters are expected, a sample size of 50 may be adequate. Hence, cluster analysis can be applied in small samples if the number of groups under consideration is not large—conditions that often apply in environmental studies or studies of groups.

In this regard, however, it is important to bear in mind the conditions calling for larger samples. Typically, larger samples will be required when: (a) the number of expected clusters is large, (b) clusters are not well separated from each other, (c) a number of small outlier clusters might emerge, and (d) the reliability of measures is low. Additionally, a larger sample will be required if an attempt is to be made to cross-validate cluster solutions.

Measures

Of course, the success of any cluster analysis depends on the quality of the measures used to describe the similarities and differences among entities. As a result, use of more reliable measures (e.g., scales as opposed to items) improves the accuracy of clustering. Clustering, moreover, does not have the linear restraint problem characteristic of other multivariate analyses. As a result, the addition of more measures, particularly measures providing more independent information, is desirable. In fact, cluster analysis results can improve with the inclusion of measures drawn from different domains, for example, both people and environmental measures, if this is substantively appropriate (Scott, Leritz, & Mumford, in press).

The proviso here, however, is that the measures under consideration must capture differences between likely subpopulations. Breckenridge (2000), in another Monte Carlo study, found that specification error, the inclusion of variables not relevant to subpopulation discrimination, led to noteworthy decrements in the performance of most major clustering procedures with respect to identification of the true number of underlying clusters. This finding is consistent with the observations of Gati, Garty, and Fassa (1996) indicating that fit effects on satisfaction are stronger when fit is appraised with respect to core occupational characteristics. Taken together, these studies suggest that substantial attention should be given to the selection and development of the measures to be applied in clustering with either theory or prior empirical work being used to identify measures likely to capture critical differences.

Similarity Assessment

A complete assessment of profile similarity takes into account the pattern, elevation, and scatter of scores on the measures under consideration. Full similarity assessments can be obtained through distance metrics—either Euclidian (independent variables) or mahalanobis (correlated variables). Correlations (or κ statistics for nonparametric data) describing similarity in score patterns are also commonly used to assess similarity in cluster analysis (Milligan, 1981). What is of note here is that these correlational indices ignore elevation and scatter resulting in a clustering based solely on pattern. Along similar lines, standardization of variable scores by setting means to 0 and standard deviations to 1 will result in a cluster analysis that considers only pattern.

These points are of some importance in studies of PE fit because there is a need to specify the elements of profile similarity that will be considered in cluster definition. For example, in studies of values in which centrality and differentiation are of concern, substantive considerations will lead to the use of distance metrics. In other studies in which the concern is identifying relative differences between people (Blashfield & Morey, 1980), a focus on pattern differences is more appropriate, thereby recommending correlational indices or the use of distance metrics based on standard scores. In this regard, however, it should be noted that if unstandardized metrics are applied, greater weight will be given to more variable measures and arbitrary cross-scale differences in range may distort cluster identification.

Choice of Clustering Procedure

Although a number of methods, really decision algorithms, have been devised to group entities together on the basis of similarity metrics. The four most commonly applied methods are (a) single linkage, (b) complete linkage, (c) average linkage, and (d) centroid (see Gnanadesikan et al., 1989, for details of these types). Single linkage clustering proceeds by joining clusters with the smallest distance between the two closest cluster members. Complete linkage proceeds by joining clusters where the two most distant members of the clusters under consideration are less distant than all other alternatives. Average linkage proceeds by joining those clusters with the smallest average distance between all pairwise linkages of cluster members. The centroid method proceeds by combining the two clusters that minimize within-group variation around the centroid, or average profile, of cluster members. The effectiveness of these alternative methods depends to some extent on the underlying structure of the relationships among subpopulations (Gnanadesikan et al., 1989). Monte Carlo studies examining how clustering procedures perform under different conditions (Breckenridge, 2000; Milligan, 1981) indicate that the centroid and average linkage methods are more accurate than other available techniques in identifying the true number of clusters and assigning entities to their true clusters—findings that recommend the use of centroid and average linkage procedures in studies of PE fit.

Evaluation of Cluster Solutions

The first question that arises in evaluating the results of a cluster analysis is how many clusters should be retained? Although some statistical procedures have been developed to answer this question (Everitt, 1979), the difficulty in formulating adequate null models has led investigators to continue to rely on scree tests. In a scree test, the plot of within-group differences, or within-group variation, is obtained at each step in the iteration. The number of clusters to be retained is determined by finding that point at which further cluster combinations result in a sharp increase in within-cluster differences. The problems that arise in applying this

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface- and deep-level diversity on group functioning. *Academy of Management Journal, 45*, 1029–1045.
Harrison, D. A., & Sin, H. P. (2006). What is diversity and how should it be measured? In A. Konrad, P. Prasad, and J. Pringle (Eds.), *Handbook of workplace diversity* (pp. 191–216). Thousand Oaks, CA: Sage.
Hollenbeck, J. R., Moon, H., Ellis, A. P. J., West, B. J., Ilgen, R. R., Sheppard, L., Porter, C. O. L. H., & Wagner, J. A., III. (2002). Structural contingency theory and individual differences: Examination of external and internal person-team fit. *Journal of Applied Psychology, 87*, 599–606.
Hofstede, G. (1998). Identifying organizational subcultures: An empirical approach. *Journal of Management Studies, 35*, 1–12.
Johns, G. (1981) Difference score measures of organizational behavior variables: A critique. *Organizational Behavior and Human Performance, 27*, 443–464.
Judge, T. A., & Cable, D. M. (1997). Applicant personality, organizational culture, and organizational attraction. *Personnel Psychology, 50*, 359–394.
Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the big five. *Psychological Bulletin, 116*, 245–258.
Klein, K. J., Bliese, P. D., Kozlowski, S. W. J., Dansereau, F., Gavin, M. B., Griffin, M. A., Hofmann, D. A., James, L. R., Yammarino, F. J., & Bligh, M. C. (2000). Multilevel analytical techniques: Commonalities, differences and continuing questions. In K. J. Klein and S. W. J. Kozlowski (Eds.), *Multilevel theory, research and methods in organizations* (pp. 512–553). San Francisco: Jossey-Bass.
Klein, K. J., Conn, A. B., Smith, D. B., & Sorra, J. S. (2001). Is everyone in agreement? An exploration of within-group agreement in employee perceptions of the work environment. *Journal of Applied Psychology, 86*, 3–16.
Kristof, A. L. (1996). Person-organization fit: an integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology, 49*, 1–49.
Lau, D. C., & Murnighan, J. K. (1998). Demographic diversity and faultlines: The compositional dynamics of organizational groups. *Academy of Management Review, 23*, 325–340.
Livingstone, L. P., Nelson, D. L., & Barr, S. H. (1997). Person-environment fit and creativity: An examination of supply-value and demand-ability versions of fit. *Journal of Management, 23*, 119–146.
Luthans, F., & Davis, T. (1982). An idiographic approach to organizational behavior research: The use of single case experimental designs and direct measures. *Academy of Management Review, 7*, 380–391.
Markus, H. R. (1977). Self-schemata and process information about the self. *Journal of Personality and Social Psychology, 34*, 63–78.
Milligan, G. W. (1981). A review of Monte Carlo tests of cluster analysis. *Multivariate Behavioral Research, 16*, 379–407.
Mumford, M. D., Snell, A. F., & Hein, M. B. (1993). Varieties of religious experience: Continuity and change in religious involvement. *Journal of Personality, 61*, 69–88.
Mumford, M. D., Stokes, G. S., & Owens, N. A. (1990). *Patterns of life adaptation: The ecology of human individuality*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Mumford, M. D., Zaccaro, S. J., Johnson, J. R., Diana, M., & Threlfall, K. V. (2000). Patterns of leader characteristics: Implications for performance and development. *The Leadership Quarterly, 11*, 115–133.
Naylor, J. C., Pritchard, R. D., & Ilgen, D. R. (1980). *A theory of behavior in organizations*. New York: Academic Press.
Newcomb, T. M. (1961). *The acquaintance process*. New York: Holt, Rinehart, and Winston.
O'Reilly, C. A., Chatman, J. A., & Caldwell, D. F. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *Academy of Management Journal, 34*, 487–516.
Ostroff, C. (1993). The effects of climate and personal influences on individual behavior and attitudes in organizations. *Organizational Behavior and Human Decision Processes, 56*, 56–90.
Ostroff, C., Kinicki, A. J., & Clark, M. A. (2002). Substantive and operational issues of response bias across levels of analysis: An example of climate-satisfaction relationships. *Journal of Applied Psychology, 87*, 355–368.
Overall, J. E., & Magee, K. N. (1992). Replication as a rule for determining the number of clusters in hierarchical cluster analysis. *Applied Psychological Measurement, 16*, 119–128.
Owens, W. A., (1968). Toward one discipline of scientific psychology. *American Psychologist, 26*, 947–997.
Owens, W. A., & Schoenfeldt, L. F., (1979). Toward a classification of persons. *Journal of Applied Psychology, 64*, 596–607.
Pelham, B. W. (1993). On the highly positive thoughts of the highly depressed. In R. F. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 183–199). Plenum Series in Social/Clinical Psychology. New York: Plenum.
Rice, R. W., McFarlin, D. B., Hunt, R. G., & Near, J. P. (1985). Organizational work and the perceived quality of life: Toward a conceptual model. *Academy of Management Review, 10*, 296–310.
Roberts, K. H., Hulin, C. L., & Rousseau, D. M. (1978). *Developing an interdisciplinary science of organization*. San Francisco: Jossey-Bass.
Schneider, B. (1987). The people make the place. *Personnel Psychology, 41*, 437–453.
Schneider, B., Kristof-Brown, A., Goldstein, H. W., & Smith D. B. (1997). What is this thing called fit? In. N. Anderson and P. Herriot (Eds.), *International Handbook of Selection and Assessment* (pp. 393–412). Chichester, England: Wiley.
Scott, G. M., Leritz, L. E., & Mumford, M. D. (in press). Types of creativity training: Approaches and their effectiveness. *Journal of Creative Behavior*.
Tinsley, H. E. A. (2000). The congruence myth: An analysis of the efficacy of the person-environment fit model. *Journal of Vocational Behavior, 56*, 147–179.
Tsui, A. S., Egan, T. D., & O'Reilly, C. (1992). Being different: Relational demography and organizational attachment. *Administrative Science Quarterly, 37*, 549–579.
Tsui, A. S., & O'Reilly, C. (1989). Beyond simple demographic effects: The importance of relational demography in supervisor-subordinate dyads. *Academy of Management Journal, 32*, 402–423.
Volard, S. V., & McCarthy, P. J. (1979). Do extraverted people make better sales representatives? *Management Decision, 17*, 467–486.
Ward, J. H., & Hook, M. E. (1963). Application of an hierarchical grouping procedure to a problem of grouping profiles. *Educational and Psychological Measurement, 1*, 69–81.

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.
Barsade, S. G., Ward, A. J., Turner, J. D. F., & Sonnenfeld, J. (2000). To your heart's content: A model of affective diversity in top management teams. *Administrative Science Quarterly, 45*, 802–836.
Blashfield, R., & Morey, J. (1980). A comparison of four clustering methods using MMPI Monte Carlo data. *Applied Psychological Measurement, 4*, 57–64.
Block, J. H. (1978). *The Q-sort method in personality assessment and psychiatric research.* Palo Alto, CA: Consulting Psychologist Press.
Breckenridge, J. N. (2000). Validating cluster analysis: Consistent replication and symmetry. *Multivariate Behavioral Research, 35*, 261–285.
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*, 32–54.
Bunderson, J. S., & Sutcliffe, K. M. (2002). Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects. *Academy of Management Journal, 45*, 875–893.
Byrne, D. (1971). *The attraction paradigm.* New York: Academic Press.
Cable, D. M., & Judge T. A. (1997). Person-organization fit, job choice decisions, and organizational entry. *Organization Behavior and Human Decision Processes, 67*, 294–311.
Caldwell, D. F., & O'Reilly, C. A. (1990). Measuring person-job fit using a profile comparison process. *Journal of Applied Psychology, 75*, 648–657.
Caplan, R. D. (1987). Person-environment fit theory and organizations: Commensurate dimensions, time perspectives, and mechanisms. *Journal of Vocational Behavior, 31*, 248–267.
Chatman, J. A. (1989). Improving interactional organizational research: A model of person-organization fit. *Academy of Management Review, 14*, 333–349.
Chatman, J. A. (1991). Matching people and organizations: Selection and socialization in public accounting firms. *Administrative Science Quarterly, 36*, 459–484.
Chatman, J. A., & Barsade, S. G. (1995). Personality, organizational culture, and cooperation: Evidence from a business simulation. *Administrative Science Quarterly, 40*, 423–443.
Chatman, J. A., Caldwell, D. F., & O'Reilly, C. A. (1999). Managerial personality and performance: A semi-idiographic approach. *Journal of Research in Personality, 33*, 514–545.
Cronbach, L. J. (1958). Proposals leading to analytic treatment of social perception scores. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior* (pp. 353–379). Stanford, CA: Stanford University Press.
Cronbach, L. J. (1992). Four *Psychological Bulletin* articles in perspective. *Psychological Bulletin, 112*, 389–392.
Cronbach, L. J., & Gleser, G. C. (1953). Assessing the similarity between profiles. *Psychological Bulletin, 50*, 456–473.
De Fruyt, F. (2002). A person-centered approach to P-E fit questions using a multiple-trait model. *Journal of Vocational Behavior, 60*, 73–90.
Edwards, J. R. (1991). Person-job fit: A conceptual integration, literature review, and methodological critique. *International Review of Organizational Psychology, 6*, 283–357.

Edwards, J. R. (1993). Problems with the use of profile similarity indices in the study of congruence in organizational research. *Personnel Psychology, 46*, 641–665.
Edwards, J. R. (1994). The study of congruence in organizational behavior research: Critique and a proposed alternative. *Organizational Behavior and Human Decision Processes, 58*, 51–100 [Erratum, *58*, 323–325].
Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Decision Processes, 64*, 307–324.
Edwards J. R. (1996). An examination of competing versions of the person-environment fit approach to stress. Academy of Management Journal, 39, 292–339,
Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow & N. W. Schmitt (Eds.), *Advances in measurement and data analysis* (pp. 350–400). San Francisco: Jossey-Bass.
Edwards, J. R., Caplan, R. D., & Harrison, R. V. (1998). Person-environment fit theory: Conceptual foundations, empirical evidence, and directions for future research. In C. L. Cooper (Ed.), *Theories of organizational stress* (pp. 28–67). Oxford, England: Oxford University Press.
Edwards, J. R., Klein, K. J., Shipp, A. J., & Lim, B. C. (2003, April). *The study of dispersion in organizational behavior research: An analytical framework using distributional moments.* Paper presented at the 18th annual meeting of the Society for Industrial and Organizational Psychology, Orlando, FL.
Edwards, J. R., & Parry, M. E. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *Academy of Management Journal, 36*, 1577–1613.
Edwards, J. R., & Rothbard, N. P. (1999). Work and family stress and well-being: An examination of person–environment fit in the work and family domains. *Organizational Behavior and Human Decision Processes, 77*, 85–129.
Endler, N. S., & Magnusson, D. (1976). *Interactional psychology and personality.* Washington DC: Hemisphere.
Everitt, B. S. (1979). Unresolved problems in cluster analysis. *Biometrics, 35*, 169–181.
Gati, I., Garty, Y., & Fassa, N. (1996). Using career-related aspects to assess person–environment fit. *Journal of Counseling Psychology, 43*, 196–206.
Gnanadesikan, R., Blashfield, R. K., Breiman, L., Dunn, O. J., Friedman, J. H., Fu, K., Hartigan, J. A., Kettenring, J. R., Lachenbruch, P. A., Olshen, R. A., & Rohlf, F. J. (1989). Discriminant analysis and clustering. *Statistical Science, 4*, 34–69.
Gustafson, S. B., & Mumford, M. D. (1995). Personal style and person-environment fit: A pattern approach. *Journal of Vocational Behavior, 46*, 163–188.
Harrison, D. A. (2002). Meaning and measurement of work role withdrawal: Current controversies and future fallout under changing technology. In M. Koslowsky & M. Krausz (Eds.), *Voluntary employee withdrawal and inattendance: A current perspective* (pp. 95–132). London: Plenum.
Harrison, D. A., McLaughlin, M. E., & Coalter, T. M. (1996). Context, cognition, and common method variance: Psychometric and verbal protocol evidence. *Organizational Behavior and Human Decision Processes, 68*, 246–261.
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*, 96–107.

approach are that (a) clustering procedures are biased toward retaining too many groups and (b) scree tests typically are ambiguous, indicating two or three potential solutions. Given these problems, the common strategy used to select a solution is to obtain, for each of the available alternatives, the distances between clusters, the number of entities assigned to each cluster, and the cluster members' average score profiles. The solution retained is the one for which clusters are well separated, relatively few unduly small clusters are identified, and the clusters evidence substantively meaningful score profiles.

After identification of a plausible solution, a more formal evaluation of the candidate solution occurs that typically involves some combination of the following analyses: (a) assessment of the stability of the cluster solution by replicating the clustering in the cross-validation samples, (b) assessment of assignment replication using a nonhierarchical procedure applied to entities in the validation and cross-validation samples, (c) assessment of simple structure in the assignment of entities to clusters, (d) assessment of the number of variables used in clustering evidencing significant differences across clusters, (e) assessment of the substantive coherence of the observed differences, (f) assessment of the results obtained in a discriminant analysis in which relevant measures serve as predictors of cluster membership to obtain canonical correlations, chi-square values, and an understanding of the underlying variables that distinguish clusters, and (g) assessment of the differences observed among clusters on a set of reference measures, measures not applied in the initial clustering, to provide evidence for the convergent and divergent validity of the solution with respect to observed group differences.

Application of Clusters

After identification of a cluster solution and provision of some evidence for its validity, it becomes possible to apply the cluster solution in studies of PE fit. What should be recognized here, however, is that clusters per se have no direct implications for PE fit. Instead, it is the pattern of differences evidenced by cluster members on criteria indicative of fit, such as satisfaction (Bretz & Judge, 1994), more rapid skill development (De Fruyt, 2002), or higher creativity (Livingstone, Nelson, & Barr, 1997), that permits inferences about the effects of fit. Thus, the application of clusters in PE fit studies requires careful development of criteria, ideally process as well as performance criteria, and demonstration of differences on these criteria across clusters.

In analyses along these lines, however, it should be borne in mind that because individual and environmental characteristics exert main effects, differences across clusters in criterion scores speak to PE effects only under two conditions. First, when the environment is "fixed," differences in process or performance measures across clusters can be attributed to persons', or types', differential reactions to a common environment. Second, when interactions between certain person types and environmental types emerge, under conditions where people are not exposed to a common environment, differences in criterion scores point to PE effects.

These boundary conditions on PE fit inferences should be evaluated in light of the advantageous characteristics of the cluster analysis approach. First, the clustering strategy allows fit effects to be examined across a wide range of measures, potentially allowing the development of more comprehensive models of the processes and outcomes of PE fit. Second, these processes and outcomes need not be constant across all clusters. As a result, it becomes possible to develop both person-specific and/or environment-specific models of fit processes and outcomes (Mumford, Snell, & Hein, 1993).

Conclusions

The promise of applying cluster analysis in studies of PE fit is that it may allow us to develop more sophisticated models of fit processes and outcomes tailored to the unique characteristics of certain types of people and the demands of certain types of environments. It is particularly relevant for assessing complementary fit. Models of this sort will permit us to move beyond superficial statements about general fit effects that ignore the unique ways different people adapt to different environmental demands. It is hoped that the present effort, by clarifying the methodological requirements for cluster analytic studies of PE fit, will provide an impetus for a new wave of research along these lines.

ACKNOWLEDGMENTS

The major sections of this chapter were authored as follows: General Methodological and Design Issues was authored by Cheri Ostroff; Profile Comparison Methods for Assessing Person–Situation Fit was authored by David F. Caldwell, Jennifer A. Chatman, and Charles A. O'Reilly; Polynomial Regression and Response Surface Methodology was authored by Jeffrey R. Edwards; Heterogeneity as Misfit was authored by David A. Harrison and Hock-Peng Sin; and Cluster Analysis was authored by Michael D. Mumford and Jasmine Espejo.

REFERENCES

Allison, P. D. (1978). Measures of inequality. *American Sociological Review, 43*, 865–880.

Anderberg, M. R. (1973). *Cluster analysis for applications*. New York: Academic Press.

Aronoff, J., & Wilson, J. P. (1985). *Personality in the social process*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Barrick, M. R., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives. *Journal of Applied Psychology, 78*, 715–723.