

# Functional Relations Among Constructs in the Same Content Domain at Different Levels of Analysis: A Typology of Composition Models

David Chan

Michigan State University and National University of Singapore

Composition models specify the functional relationships among phenomena or constructs at different levels of analysis (e.g., individual level, team level, organizational level) that reference essentially the same content but that are qualitatively different at different levels (M. T. Hannan, 1971; K. H. Roberts, C. L. Hulin, & D. M. Rousseau, 1978; D. M. Rousseau, 1985). Specifying adequate composition models is a critical component of good multilevel research. A typology of composition models is proposed to provide a framework for organizing, evaluating, and developing constructs and theories in multilevel research. Five basic forms of composition are described and illustrated. Implications of the typology are discussed.

Organizational phenomena have the properties of dynamic systems, with critical antecedents, processes, and outcomes conceptualized and measured at multiple levels of organizational analysis (e.g., individual, group, organization). Because more researchers are beginning to realize that the organizational phenomenon under investigation often is inherently multilevel as opposed to occurring at a single level or in a level vacuum, organizational studies increasingly are adopting a multilevel approach. Several influential theoretical frameworks for multilevel research have been proposed (e.g., House, Rousseau, & Thomas-Hunt, 1995; Klein, Dansereau, & Hall, 1994; Rousseau, 1985). Excellent discussions on important mathematical issues related to the analysis of multilevel data (e.g., Bliese & Halverson, in press; Ostroff, 1993a) and analytical models for structuring multilevel data (e.g., Bryk & Raudenbush, 1987; McArdle & Epstein, 1987; Meredith & Tisak, 1990; Muthen, 1994; Willett & Sayer, 1994) also are available.

However, despite the existence of broad theoretical

frameworks and methodological advances, the fundamental substantive issue of construct validation in multilevel research has not been addressed adequately. Accompanying the increased interest in multilevel research is an increased proliferation of new constructs at multiple levels. Unless we have explicit *composition models* to guide the development and validation of newly proposed constructs in multilevel research, there is a danger of violating the scientific principle of parsimony. Organizational researchers could easily end up with a multitude of labels, all of which purportedly refer to scientific constructs but in reality have no incremental explanatory value. Composition models specify the functional relationships among phenomena or constructs at different levels of analysis (e.g., individual level, team level, organizational level) that reference essentially the same content but that are qualitatively different at different levels (Hannan, 1971; Roberts, Hulin, & Rousseau, 1978; Rousseau, 1985). Specifying functional relationships between constructs at different levels provides a systematic framework for mapping the transformation across levels. The explicit transformation relationships provide conceptual precision in the target construct, which in turn aids in the derivation of test implications for hypothesis testing. Unfortunately, the specification of functional relationships between constructs has not always been adequate or even explicit in multilevel research. This is partly because no systematic frameworks for specifying functional relationships exist.

An adequate typology of composition models addresses the above problems and contributes to multilevel research in at least two important ways. First, it provides an organizing framework for existing focal constructs facilitating scientific communication in multilevel research. Researchers can be more confident that they are referring to

---

David Chan, Department of Psychology, Michigan State University, and Department of Social Work and Psychology, National University of Singapore.

I thank Neal Schmitt, Daniel Ilgen, and Richard DeShon for their helpful comments on the article, and I am grateful to Steve Kozlowski for introducing me to multilevel research.

An earlier version of this article was presented at the 1997 Annual Meeting of the Academy of Management in Boston, Massachusetts, August 1997.

Correspondence concerning this article should be addressed to David Chan, Department of Psychology, 129 Psychology Research Building, Michigan State University, East Lansing, Michigan 48824-1117. Electronic mail may be sent to dchan@pilot.msu.edu.

the same construct when it is explicated according to the same form of composition. Meaningful replications and extensions of current findings then are possible. Apparent contradictory findings may be reconciled, and debates may be clarified. For example, many so-called inconsistent findings simply could be a result of confusion of terminology (i.e., comparing apples and oranges), and the confusion may become apparent when each study locates its construct in the typology corresponding to the composition model. Organizing existing constructs also aids cumulation of research findings by providing a framework for performing meaningful meta-analytic studies in multilevel research.

Second, a typology provides a conceptual framework for developing and validating new focal constructs and multilevel theories. As described later in this article, the typology of models could help compose new explanatory constructs from established ones. In addition, being cognizant of different models allows the researcher to consider alternative designs, measurements, and data analyses for testing competing hypotheses, modifying existing theories or developing new ones, or performing a more rigorous test of the original hypothesis. The purpose of this article is to propose a typology of composition models.

### A Typology of Composition Models

The proposed typology is concerned with *elemental* composition; that is, situations in which data from a lower level are used to establish the higher level construct. In other words, the higher level construct is of a collective or aggregate nature and is construed as some form of combination of the lower level units. All lower level units play some substantive role in composing the lower level construct to the higher level construct, and the value of the higher level construct is not solely determined by any single lower level unit (i.e., each unit is used in some way or another, such as for computing the mean level). Note that the use of data from the lower level to establish the higher level construct does not imply that it is necessary to begin conceptualization at a level lower than the level of the target or composed construct. The starting level of conceptualization is dependent on the research question. For example, a researcher may start at the group level<sup>1</sup> with the established construct of group norms and then move down to the individual level to collect perceptual data for subsequent aggregation to the group level to establish the construct of group norms. The focus on elemental composition is consistent with the actual constraints and practice in empirical multilevel research. As noted by several researchers (e.g., Ostroff, 1993a; Roberts et al., 1978), we often do not have global indices of the higher level (organizational or group) variables of interest and hence have to rely on aggregated data from the lower

level (individuals) to represent the higher level variable. Because the focus is on elemental composition, this article does not address the traditional issues of disaggregation and ecological fallacies (Cronbach, 1976; Hannan, 1971; Langbein & Lichtman, 1978).

Table 1 presents the typology of composition models. The typology describes the basic forms composition models can take. The forms described are ideal types. The five basic forms of composition models are: (a) *additive*, (b) *direct consensus*, (c) *referent-shift consensus*, (d) *dispersion*, and (e) *process composition*. A theory of the focal construct in a multilevel study may contain one or more of the five composition forms. The five forms of composition models in Table 1 are not presented in any specific order. The reader is cautioned against ordering the five forms in terms of the similarity of functional relationships between constructs. The simplistic notion of similarity of functional relationships is not an adequate way of representing the complexity in or guiding the development of composition forms because there is an infinite number of answers (i.e., dimensions) to the question, "Similar with respect to what?"

As shown in Table 1, each composition model is defined by a particular form of functional relationship specified between constructs at different levels. Corresponding to each form of functional relationship is a typical operational process by which the lower level construct is combined to form a higher level construct. Note that the operational combination process is the typical form as opposed to a necessary consequence of the functional relationship specified. The column in Table 1 labeled *empirical support* suggests what constitutes the forms of evidence needed to support the relevant functional relationships and to establish that appropriate combination rules are applied. To illustrate the forms of composition, examples from climate research are consistently used throughout

<sup>1</sup> In describing the typology and throughout this article, the terms *group* and *team* are used interchangeably to refer to the level of the collection of individuals immediately higher than the individual level. Several researchers have distinguished groups and teams as the polar ends on a continuum of task interdependence (e.g., Salas, Dickinson, Converse, & Tannenbaum, 1992). In addition, groups are often characterized by low role differentiation and low task differentiation, whereas teams are characterized by high role differentiation, high task differentiation, distributed expertise, and high levels of task interdependence (Sundstrom et al., 1990). The present typology is concerned with work groups or work teams in organizations. The terms *groups* and *teams* are used to refer to multiple individuals formed to perform some organizationally relevant task-functions. These individuals interact, exhibit task interdependence, possess one or more shared goals, and are embedded in a larger organizational setting (Kozlowski et al., 1994, 1996; Salas et al., 1992).

Table 1  
A Typology of Composition Models

Functional relationships	Typical operational combination	Empirical support	Example from climate research
Additive model			
Higher level unit is a summation of the lower level units regardless of the variance among these units	Summing or averaging lower level scores	Validity of additive index (e.g., mean of lower level units)	From <i>psychological climate</i> to <i>organizational climate</i> (Glick's [1985] conceptualization)
Direct consensus model			
Meaning of higher level construct is in the consensus among lower level units	Within-group agreement to index consensus and justify aggregation	Value of within-group agreement index (e.g., $r_{wg}$ ); validity of aggregated scores	From <i>psychological climate</i> to <i>organizational climate</i> (James et al.'s [1984] conceptualization)
Referent-shift consensus model			
Lower level units being composed by consensus are conceptually distinct though derived from the original individual-level units	Within-group agreement of new referent lower level units to index consensus and justify aggregation	Value of within-group agreement index (e.g., $r_{wg}$ ); validity of aggregated scores	From <i>psychological climate</i> to <i>organizational collective climate</i>
Dispersion model			
Meaning of higher level construct is in the dispersion or variance among lower level units	Within-group variance (or its derivative) as operationalization of the higher level construct	Absence of multimodality in within-group distributions of lower level scores; validity of dispersion index	From <i>psychological climate</i> to <i>climate strength</i>
Process model			
Process parameters at higher level are analogues of process parameters at lower level	No simple algorithm; ensure analogues exist for all critical parameters	Nomological validity for source and target constructs at their respective levels to distinguish shared core content from level-specific aspects	From <i>psychological climate</i> development to <i>organizational climate emergence</i>

the typology. I assume that a researcher collects climate data (i.e., individual perceptual responses on climate questionnaires) from hundreds of individuals across many organizations (or groups). For each composition model, I present examples of hypotheses and explanations of how the model can be applied to the data. Where appropriate, examples other than climate also are presented to clarify the composition form. The first two forms of composition are familiar to most researchers, and I will only describe them briefly. The remaining three forms are less familiar even though they characterize many focal constructs examined in multilevel studies. As mentioned earlier, composition models are seldom made explicit in existing multilevel research. The following sections describe each of the five composition models in the typology.

### Additive Models

Additive composition models specify a straightforward functional relationship between constructs at different lev-

els. In such models, the meaning of the higher level construct is a summation of the lower level units regardless of the variance among these units. In additive composition models, the variance of the lower level units is of no theoretical or operational concern for composing the lower level construct to the higher level construct. The typical operational combination process is a simple sum or average of the lower level scores on the lower level variable to represent the value on the higher level variable. The validity of the additive index (e.g., the mean) constitutes empirical support for the composition.

In the climate example, the researcher may be interested in relating organizational climate to organizational performance. The researcher has an established measure of organizational performance but he or she has to develop some conceptualization and measure of the construct of organizational climate. Adopting Glick's (1985) conceptualization, the researcher stipulates that all organizations have an organizational climate that can be described as high or low on various dimensions regardless of the level of

within-organization individual-level agreement. Within-organization agreement, according to this view, is an issue of measurement accuracy reflecting individual-level random error and sources of bias (see Glick, 1985, pp. 604–605). Hence, using an additive composition model, the researcher averages the climate perceptions of individuals within each organization, regardless of the within-organization variance, to represent the organizational climate variable. The organizational mean climate scores and the organizational performance variable then are correlated, and the validity coefficient (i.e., validity of the additive index; in this case, the mean) provides empirical support for the additive composition model.

Clearly, whether the relationship between organizational climate and psychological climate (i.e., individual-level climate perceptions) is additive or some other composition form depends on how the construct of organizational climate is conceptualized. For example, if the level of individual perceptual agreement within an organization is central in the substantive definition of organizational climate, then an additive composition model would be inappropriate because within-group (i.e., organization) variance among lower level units (i.e., individual perceptions) becomes relevant in composing the lower level construct (i.e., psychological climate) to the higher level construct (i.e., organizational climate). In this case, a direct consensus composition model, described in the next section, is appropriate.

### *Direct Consensus Models*

Direct consensus composition is probably the most familiar and popular form of composition among multilevel researchers. This model uses within-group consensus of the lower level units as the functional relationship to specify how the construct conceptualized and operationalized at the lower level is functionally isomorphic to another form of the construct at the higher level. The typical operational combination process is using within-group agreement of scores to index consensus at the lower level and to justify aggregation of lower level scores to represent scores at the higher level (e.g., James, Demaree, & Wolf, 1984; Kozlowski & Hattrup, 1992; Ostroff, 1993b; Ostroff & Rothausen, 1997). This operational combination process has two components. The first component includes an operationalization (i.e., a measure) of the conceptual definition for each of the two constructs (i.e., one at each level). For example, individual-level perceptual responses on a climate measure are used to operationalize psychological climate, whereas the mean of those individual responses within an organization is used to operationalize organizational climate. The second component specifies the manner of and precondition(s) for combining the individual lower level measurements to represent the

higher level measurement. For example, within-group agreement indexes such as the  $r_{wg}$  index (James et al., 1984) may be calculated and some cutoff level of agreement is used to justify aggregation of individual responses. In this example, the aggregation procedure and preconditions, together with the conceptual definition of the higher level construct, determine the meaningfulness and validity of the operationalization of the higher level construct.

In the climate example, the researcher may follow James (1982) to construe psychological climate as an individual's perception or cognitive representation of the work environment in terms of the psychological meaning and significance to the individual. Organizational climate simply refers to the shared assignment of meanings among individuals within the organization. In this conceptualization, within-group agreement among individual climate perceptions indicates shared assignment of psychological meaning. It is this sharedness that constitutes functional equivalence between the climate constructs at the two levels. Hence, the definition of *organizational climate* is essentially the same as psychological climate, except that the former refers to the *shared* perceptions among the individuals. The conceptual relationship between the two forms of the construct at different levels then drives the manner in which the lower level construct composes to the higher level construct. So the researcher would proceed to check within-group agreement (*group* here refers to the organization) of individual climate responses using some agreement index (e.g.,  $r_{wg}$ ). High within-group agreement indicates consensus and justifies aggregation of individual climate responses to represent scores on the organizational climate variable.

In direct consensus composition, the lower level attributes need not be restricted to individual perceptions. Consensus, as indexed by within-group agreement, can apply to individual-level attributes such as cognitive ability and styles, personality, mental representation, and behavioral variables. The critical issue is the specification of an adequate direct consensus composition theory to present a substantive meaning for the mean or some other central tendency index of the within-group consensual attributes.

Finally, George and James (1993) argued that the use of within-group agreement as a valid justification for aggregation is independent of variability between the higher level units. The authors are correct that within-group agreement is the appropriate criterion for aggregation in a direct consensus composition because the model specifies that the meaning of the higher level construct is in the consensus among the lower level units. However, between-group variability may be more important in construct validation of the aggregated data than George and James appear willing to grant. Assuming adequate sampling, the

absence of between-group variability indicates that hypothesized differences at the group level do not exist. But more important, it calls into question the validity of the hypothesized group-level construct. This is especially relevant in exploratory situations in which the researcher has little clue as to which level is the appropriate level to aggregate the data to for composing a higher level construct. In the climate example, assuming adequate sampling, the absence of between-department (within an organization) variability in aggregated scores (from individual to department level) and the presence of between-organization variability in aggregated scores (from individual to organizational level) suggest that it is more appropriate to aggregate the individual-level data to the organizational than departmental level, and it provides construct validity evidence for organizational climate but not departmental climate. The role of between-group variability in the interpretation of aggregated scores remains a debatable point in the multilevel literature (see George, 1990; George & James, 1993; Yammarino & Markham, 1992).

### *Referent-Shift Consensus Models*

Referent-shift consensus composition is similar to direct consensus composition in that within-group consensus, as indexed by agreement of lower level attributes, is used to compose the lower level construct to the higher level construct. The critical difference between the two forms of composition is that in referent-shift consensus composition, the lower level attributes being assessed for consensus are conceptually distinct though derived from the original individual-level construct. That is, there is a shift in the referent prior to consensus assessment, and it is the new referent that is actually being combined to represent the higher level construct.

In the case of referent-shift consensus, the composition proceeds as follows: First, the researcher begins with a conceptual definition and operationalization of the focal construct at the lower level. While maintaining the basic content of the construct, the researcher then derives a new form of the construct at the same level by shifting the referent of the basic content. That is, the referent for the conceptual definition and operationalization is now changed. The new form of the construct then is aggregated to the higher level construct based on within-group consensus.

Consider our climate example. Rather than an individual's own climate perceptions (i.e., psychological climate) or the aggregation of individuals' perceptions (i.e., organizational climate), the researcher now is interested in how an individual believes others in the organization perceive the climate and whether there is within-organization consensus in such beliefs. So now there are two new variables: One, *psychological collective climate*, is de-

fined as the individual's description of other organizational members' perceptions of the climate. Note that the basic content of climate perception in the original construct (psychological climate) remains unchanged in the new variable but the referent of the content has changed (from self to others). Psychological collective climate is still at the original individual level of conceptualization. Within-group consensus (as indexed by within-group agreement of individual description scores) then is used to justify the aggregation of individuals' psychological collective climate perceptions to represent the value of the higher level (i.e., organizational level) construct, which is the second new variable, called *organizational collective climate*.

Examples of referent-shift consensus exist in multilevel research, but unfortunately the composition often is not made explicit. In research on work teams, self-efficacy at the team level is often a case of referent-shift consensus composition (e.g., Guzzo, Yost, Campbell, & Shea, 1993; Kozlowski et al., 1994). The composition starts with the individual-level construct of *self-efficacy*. Self-efficacy is defined as an individual's belief and confidence in mobilizing his or her resources for successful task performance (Bandura, 1977). An example item is "I am confident that I can perform this task." A new form of the construct at the same level (i.e., individual level) then is derived by shifting the referent in the efficacy perception from the self to the team as a whole. An example item for the new form of the construct is "I am confident that my team can perform this task." That is, there is a new form of the construct, namely, *collective efficacy*, defined as the individual team member's belief and confidence that the team can mobilize its resources for successful task performance. The basic content of efficacy perception in the original construct (self-efficacy) remains unchanged in the new form (collective efficacy), but the referent of the content has changed (from self to team). Collective efficacy is still at the original individual level of conceptualization (Guzzo et al., 1993). Within-group consensus (as indexed by within-group agreement of individuals' perceptual scores) then is used to justify the aggregation of individuals' collective efficacy perceptions to represent the value of the higher level (i.e., group-level) construct called *team efficacy*.<sup>2</sup>

<sup>2</sup> *Team-efficacy* or *group-efficacy* should be distinguished from Guzzo et al.'s (1993) concept of *group potency*. Efficacy beliefs (self, collective, or team) are task-specific in that they refer to expectations about performance on particular tasks, whereas potency refers to a more generalized belief about general effectiveness across multiple tasks. The distinction does not affect the logic of referent-shift consensus composition. One can always specify a referent-shift consensus composition model from the construct of self-potency to group potency by means of collective potency.

Referent-shift consensus composition is important because the change in referent results in a new form of the original focal construct that is conceptually distinct from the original form. In the climate example, an individual with a favorable psychological climate perception can have either favorable or unfavorable psychological collective climate. Similarly, a team member with high self-efficacy can have either low or high collective efficacy. The distinction between referent-shift consensus composition and direct consensus composition also clarifies some of the confusion in terminology in multilevel research. For example, according to the composition model just described, collective efficacy and team efficacy are isomorphic forms of the same construct at different levels, the latter being directly constituted by the former, provided within-group consensus exists. Yet, both collective efficacy and team efficacy are, in different ways, derived from the original focal construct of self-efficacy. Hence, one can speak of composing self-efficacy to team efficacy, bearing in mind that the form of composition is referent-shift consensus as opposed to direct consensus. Applying the direct consensus model for the composition of self-efficacy to team efficacy probably would be misleading. As noted by Guzzo et al. (1993), computing the mean of team members' self-efficacy scores as an indicator of the team-level construct of team efficacy often is wrong. The mean scores still would be an indicator of individual members' perceptions about themselves as individuals, not about the team as a whole. Bandura (1982) was probably the first researcher who suggested the concept of collective efficacy, although he did not elaborate on the concept. To date, there has been little consensus on the nature of the group-level notion of efficacy or its measurement. Gist (1987) suggested different ways of measuring group-level efficacy, but few researchers appreciated the differences in these measurements. A primary reason for the relatively slow progress in understanding group-level efficacy is the lack of explicit composition models that specify the nature of the focal construct and provide the conceptual basis for operationalization and measurement development.

### *Dispersion Models*

In both direct consensus composition and referent-shift consensus composition, within-group agreement of scores from the lower-level units or attributes is used to index consensus. The researcher hopes to achieve a high agreement at the lower level in order to justify aggregation to represent variables at the higher level. In these models, consensus is a necessary condition for construct validity at the higher level, and high within-group agreement constitutes an empirical or statistical precondition to be fulfilled for the operational combination process to be legitimate.

Given an adequate composition theory, the degree of within-group agreement of scores from the lower level units or attributes potentially could be conceptualized as a focal construct as opposed to merely a statistical prerequisite for aggregation. That is, instead of treating within-group variance as error variance (which is what consensus models do), within-group variance (i.e., the within-group dispersion of scores) could serve as an operationalization of a focal construct. The idea of treating within-group consensus or dispersion of scores as a theoretically significant phenomenon in its own right can be traced back to the works by James et al. (1984) more than a decade ago but, with two recent exceptions, the idea has not been followed up by multilevel researchers. Lindell and Brandt (1997) mentioned the possibility of using within-group agreement as a focal variable when discussing the  $r_{wg}$  index. Brown, Kozlowski, and Hattrup (1996) developed the same idea and argued for a reconceptualization of within-group agreement as a focal construct and presented several initial guidelines for construct development as such.

Based on this notion of within-group consensus or dispersion of scores as a construct, I propose dispersion composition models as constituting another ideal type in the typology. Brown et al. (1996) focused on borrowing dispersion theories from various streams of psychological research as the basis for conceptualizing agreement. The present discussion focuses on the use of within-group dispersion (i.e., variance or agreement) to specify the functional relationship in composition of a dispersion construct and on the conceptual and methodological considerations in the development of an adequate dispersion composition model. In addition, the term dispersion is used here to refer to variance (or homogeneity) of scores on any lower level units or attributes (e.g., individual cognitive ability, individual climate perceptions); hence, it is more general than Brown et al.'s focus on individual perceptions.

Dispersion is by definition a group-level characteristic (but not necessarily a group-level construct) because it refers to the variability within a group and a variance statistic is indexing an attribute of a group as opposed to an attribute of any individual-level response (Roberts et al., 1978). In dispersion composition, within-group variance (or some derivative) is used as the operationalization of the purported group-level construct. However, an adequate dispersion model always must give primacy to the construct as opposed to the variance index. Statistically, within-group dispersion is simply a result of individual differences within the group. Interindividual variability is ubiquitous for the kinds of individual-level data collected by organizational researchers. This variability could result from true differences on some construct, random error, or both. Without a conceptual definition of the group-level

construct (purportedly indexed by within-group dispersion) and a theory of its substantive meaning (e.g., how it relates to other related established constructs in a nomological network), the researcher may not be measuring what he or she intends to measure. By proceeding in a totally empirical fashion (i.e., atheoretical), one is not likely to replicate results. In short, the essence of dispersion composition is in specifying the nature of the higher level construct represented by dispersion along some lower-level variable.

In our climate example, the researcher may propose the construct of climate strength conceptualized as the degree of within-group consensus of climate perceptions and index the construct using within-group variance or some dispersion measure of individual climate responses. The dispersion measure then is correlated with the measure of organizational performance to test the researcher's hypothesis that organizational climate strength is associated positively with organizational performance.

In consensus composition, high within-group agreement is an empirical prerequisite for aggregating the lower level construct to the higher level construct. In dispersion composition, there exists a very different empirical prerequisite for composition. This prerequisite forms a critical component of the manner in which units are combined in dispersion composition. The prerequisite is the absence of multimodality in the within-group distributions of lower level scores. Multimodality in the distribution of scores within a group indicates that substantively meaningful subgroups may exist within the group, with low individual differences within each subgroup (i.e., high within-subgroup agreement) and high individual differences across subgroups (i.e., low intersubgroup agreement). When there is multimodality, it is possible that the variance or dispersion along the original grouping variable does not represent a meaningful dispersion construct. One may have to move downward from the group level to the subgroup level to identify any potentially meaningful subgrouping variable corresponding to the multimodal responses. In exploratory situations, simply graphing the group distribution could help identify the appropriate subgrouping level by matching distributional modality to potential grouping boundaries. Of course, matching modality to grouping boundaries is not sufficient evidence for a dispersion construct. A theory of the dispersion construct should be formulated, and further construct validity evidence, including establishing the validity of the dispersion variable, is required. For example, when exploring the possibility of a dispersion construct of climate strength, multimodality at the division level suggests that the construct of climate strength is probably inappropriate at that level. The modality may correspond to a subgroup level, such as the level of the department. The individual-level data then could be regrouped and a dispersion measure

of departmental climate strength could be validated by correlating it with external criterion variables.

The failure to consider the modality of within-group distributions is probably the primary source of the mistaken assumption that at the group level, low agreement of individual responses is the same as high disagreement. Low agreement indicates lack of consensus. High disagreement indicates existence of subgroups (within which there is high consensus) rather than lack of consensus. Low within-group agreement is indicated by a unimodal platykurtic distribution (as opposed to a unimodal leptokurtic distribution, which indicates high agreement), whereas high within-group disagreement is indicated by a multimodal distribution. Kurtosis and modality are conceptually (and mathematically) distinct properties of a distribution of scores. This mathematical fact is the underlying logic implicit in Brown et al.'s (1996) discussion on theories of subcultures and coalitions within groups. The absence of multimodality and the presence of validity evidence of the dispersion index provide empirical support for the dispersion model.

Some important conceptual and methodological issues in change analysis are especially relevant to dispersion composition. When groups differ in variances, the heteroskedasticity of variances does not necessarily reflect direct absolute differences on the dispersion construct purportedly measured by the grouping variable. *Beta* or *gamma changes* in individual responses across groups (Golembiewski, Billingsley, & Yeager, 1976) also could result in differences in within-group variances. A beta change occurs when there is a change in the subjective metric, resulting in a recalibration of the measuring instrument given a constant conceptual domain. For example, given the same climate item, individuals from two different cultures may differ in the psychological metric that the rating scale represents to them. The difference in the subjective metric will result in variance differences. That is, when there is beta change, some or all of the variance difference across the two cultures *no longer reflects true differences* in climate strength but rather differences in calibration of the measuring instrument. A gamma change occurs when there is a shift in the meaning or conceptualization of the construct being measured. For example, the same climate item may in fact be measuring the climate for safety in one group but measuring the respect for authority in another. When there is gamma change, comparisons of variances across groups are no longer meaningful. Thus, before using within-group variances as values on the dispersion construct, the researcher should, whenever possible, test for scalar and factorial invariance of individual responses across groups. Methods for testing invariance are provided in Schmitt (1982); Schmitt, Pulakos, and Lieblein (1984); Drasgow (1984); Reise, Widaman, and Pugh (1993); and Chan and Schmitt (1997).

### Process Models

The preceding four composition models are concerned with *static* core attributes of focal constructs (e.g., climate perceptions, efficacy perceptions), which describe some stable units or state of affairs at the individual or higher level. These focal constructs are certainly applicable to much of multilevel research, and, in some cases, substantial empirical evidence for construct validity has been accumulated (e.g., organizational climate). However, researchers are often interested in episodes of or changes in behaviors exhibited by an individual or by a team rather than the specific behavioral acts or perceptions. That is, there is interest in the process as opposed to some stable attributes, outcomes, or state of affairs.

Process composition models are concerned with composing some process or mechanism from the lower level of conceptualization to the higher level. In these models, a process or mechanism is first specified at the lower level explicating the essential or critical parameters and their interrelationships. The process then is composed to the higher level by identifying critical higher level parameters, which are higher level analogues of the lower level parameters, and describing interrelationships among higher level parameters, which are homologous to the lower level parameter relationships.

In process composition, functional relationships for parameters at different levels are analogous relationships, and functional relationships for parameter interrelationships at different levels are homologous relationships. Thus, there is no simple algorithm (e.g., within-group agreement) to compose the lower level process to the higher level. The operational combination is best described as ensuring that all critical parameters and parameter interrelationships are adequately operationalized at both the lower and higher levels and that each parameter and parameter interrelationship at the lower level has a counterpart at the higher level.

In the climate example, assume that the researcher is examining the climate for safety (e.g., Dedobbeleer & BeLand, 1991; Zohar, 1980), and is interested in describing the process in which the organization moves from the state of lack of within-group agreement of individual-level climate perceptions to the state of high within-group agreement. That is, the researcher wants to compose an organizational-level process of organizational safety climate emergence. To do so, the researcher first specifies an individual-level process describing how an individual develops psychological safety climate perceptions. For simplicity, assume the researcher has a rudimentary theory that development of psychological climate for safety is an integration process, moving from an initial state in which distinct beliefs about various safety practices are unrelated or, at best, loosely interrelated through progressive

states in which these separate beliefs become increasingly interrelated to the eventual state in which they become integrated into a single global belief. This integration process then is composed to the higher level to specify the process of organizational safety climate emergence. Accordingly, the researcher could specify organizational safety climate emergence as an integration process, moving from an initial state in which there is little agreement among individuals' psychological safety climate perceptions, through progressive states in which the level of agreement gradually increases, to the eventual state in which high agreement is achieved. Note that within-group agreement is a higher level analogue of intraindividual correlation of safety beliefs. Similarly, the notion of increasing levels of within-group agreement as an organization progresses over time is analogous to the notion of increasing intercorrelations among safety beliefs as an individual progresses over time. The initial and final states of the organization also are analogous to those of the individual. In short, critical parameters of the integration process at the individual level have higher level analogues that constitute the critical process parameters at the organizational level.

In this climate emergence example, the within-group agreement index is the higher level operational analogue of the correlation coefficient. Note that the within-group agreement index is used here as a dispersion measure (assessing climate strength at multiple points in time) as opposed to a statistical criterion for aggregation. The integration process in organizational climate emergence can be construed as changes in organizational climate strength. Hence, when moving from the source construct of psychological climate to the higher level process of organizational climate emergence, a dispersion composition (from psychological climate to climate strength) precedes the process composition. As mentioned earlier, a theory of the focal construct in a multilevel study may contain one or more composition forms.

Because of the complexity of process composition, I provide two additional examples. First, consider the example when a team researcher proposes that a process of self-regulation functionally similar to individual self-regulation also exists at the team level (e.g., Kozlowski, Gully, Salas, & Cannon-Bowers, 1996). At the individual level, self-regulation refers to the activities carried out by the individual to monitor and evaluate his or her own performance with respect to progress toward a goal (Hogarth, Gibbs, McKenzie, & Marquis, 1991; Kanfer & Ackerman, 1989). The critical parameters of self-regulation include understanding of the coordination of one's actions, error detection, balancing multiple tasks or workloads to stay on track toward goal achievement, and a knowledge of one's task environment (Hogarth et al., 1991; Kanfer & Ackerman, 1989; Karoly, 1993). On the



basis of these critical parameters, Kozlowski et al. (1996) composed the process of self-regulation from the individual level to the team level. The authors stated that

Teams increase their self-regulation capabilities as members develop a shared perception of the team and its environment and as they acquire teamwork skills critical to team effectiveness. . . . Team self-regulation involves an understanding of how to coordinate member actions, engage in error detection, and monitor each other's performance, so the team can balance workloads and stay on track toward stated objectives. (p. 276)

Note that the authors provided the conceptual definition of team self-regulation by explicating the team-level analogues for the critical parameters of self-regulation process at the individual level. Several researchers on team decision making also have proposed the concept of team self-regulation or similar notions, but many have not made explicit composition linkages between the individual and team levels (e.g., Hackman, 1992; Manz & Sims, 1987; Sundstrom, De Meuse, & Futrell, 1990).

Another example of process composition is composing team proceduralization from the individual-level process of proceduralization in skill acquisition. At the individual level, proceduralization occurs when new versions of old productions (condition-action rules) are built without one or more conditional elements (Anderson, 1982). Proceduralization occurs with increased experience or repeated practice on a task. Given the same task, proceduralization decreases response time by shortening the routes to arrive at goal states, which increases the efficiency of problem solving. This is because relative to old productions, new productions allow quicker testing of productions because there is less conditional information to be matched to contents of working memory. The process of proceduralization can be composed to the team level by identifying team-level analogues of the critical parameters. With increased experience or repeated practice on a team task, members acquire some shared awareness of the team situation (Kozlowski et al., 1994) and develop an understanding of the basic team strategies for dealing with established, routine, or recurrent situations (Boguslaw & Porter, 1962, as cited in Kozlowski et al., 1994). The routinization of tasks and strategies allows each team member to understand and predict other members' task behaviors or actions and hence skip the testing of certain conditional elements (i.e., information to be matched with other members' actions) as he or she performs the production-linked sequence of task behaviors. That is, team members have built new productions that make team problem solving and team task performance quicker and more efficient. Team proceduralization has occurred.

As mentioned earlier, process composition has no concrete empirical algorithm to compose the lower level process(es) to the higher level process(es). A challenge for

multilevel researchers is to develop an adequate process composition model and derive explicit hypotheses to be tested as part of the validation of the model. To complicate the matter further, processes are interrelated in a dynamic manner and do not occur in a temporal vacuum, as evident in the example of organizational climate emergence.

Finally, any researcher who has attempted to specify a process composition would have quickly realized that the processes of interest are often multifaceted or multidimensional with embedded subprocesses. For the first four composition forms in the proposed typology, the corresponding functional relationships and operational combination processes, as well as the conditions under which each model applies, are clear and consistent. However, because of its inherently dynamic and multidimensional nature, the process composition form is in need of further conceptual development. For example, is it possible to develop guidelines regarding how to distinguish shared core content from unshared aspects that are level-specific? For these guidelines to be useful, they would have to be sufficiently specific to aid the researcher in making actual decisions of what specific aspects to include or exclude and yet sufficiently general to cover a reasonably broad range of processes. The present description of process composition serves as the building block for the more comprehensive composition framework necessary for composing a complex and multidimensional process. As suggested by an anonymous reviewer of the present article, nomological nets (Roberts et al., 1978) could be specified and tested at each level, and then the hypothesized shared core content could be tested for the composition model. Note also that because multiple constructs may be specified and interrelated in the description of a process, it may be the case that an adequate process composition model in multilevel research has to be preceded by specifying one or more composition forms in the present typology for composing the relevant higher level constructs.

### Multilevel Construct Validation

Before evaluating the typology, some remarks concerning multilevel construct validation are relevant. In some types of multilevel research, both the lower level and the higher level constructs provide the components necessary for the conceptualization and operationalization of the target construct. The researcher begins with two constructs: one at the lower level (e.g., individual) and one at the higher level (e.g., organization). Both constructs may or may not be commensurate constructs. That is, the two constructs may or may not share the same core content dimensions. However, each construct must be explicitly defined. The definition of the target construct is then derived from the definitions of these two constructs. The substantive meaning (i.e., the conceptual definition) of

the target construct is some form of combination of the core elements from both the lower and higher level constructs. A theory of the target construct specifies the nature and form of the combination. This theory would provide the conceptual basis for the process that specifies how values on the lower and higher level constructs should be empirically combined to form values on the target construct.

The focal constructs in person-organization fit (P-O fit) research are prototypical examples of target constructs that are derived from a combination of constructs at different levels. In P-O fit studies, the *fit construct* is a target construct consisting of some person-level construct and some organizational-level construct. For example, in Chan's (1996) P-O fit study on problem-solving styles at work, Chan began by explicating the construct of *adaption-innovation problem-solving style*, which is an individual-level construct based on Kirton's adaption-innovation theory (Kirton, 1976). Using Kirton's theory, Chan specified the higher level construct of style demands conceptualized at the work-context level. Based on the lower level construct of problem-solving style and the higher level construct of style demands, Chan then derived a target construct he called *cognitive misfit*. Cognitive misfit refers to the degree of mismatch between an individual's cognitive style of problem solving and the style demands of the work context. Cognitive misfit was operationalized and tested in terms of a statistical interaction between problem-solving style and style demands. Chan demonstrated that whereas neither the individual-level construct nor the work context-level construct was associated with turnover probability, the cross-levels construct of cognitive misfit provided significant and substantial incremental validity in predicting actual turnover over the predictability provided by performance. A similar example is found in O'Reilly, Chatman, and Caldwell's (1991) study on the target (P-O fit) construct of *value congruence*, defined as the match between individuals' values and the organization's culture. The authors used the Q-sort methodology to develop and validate the Organizational Culture Profile (OCP), which was a measure of 54 values. The OCP was used to derive individuals' value profiles and the organization's culture. Value congruence was operationalized in terms of the Q-sort based profile correlations. O'Reilly et al. demonstrated that value congruence was a valid predictor of satisfaction, commitment, and actual turnover for a 2-year period and that person or organization variables alone were not predictive of these outcomes. There has been a surge of interest in the study of P-O fit in recent selection and organizational research (for review, see Edwards, 1991; Judge & Ferris, 1992; Kristof, 1996). P-O fit studies are inherently multilevel studies, and the focal constructs in P-O fit research are inherently multilevel in nature. P-O fit researchers should devote more

theoretical and empirical effort to specify and validate multilevel constructs in their studies.

### Evaluating the Composition Typology

It will be some time before the proposed composition typology has been evaluated thoroughly as a conceptual framework for multilevel research in general and for development of useful composition models in particular. The criteria for evaluation should correspond to the two potential contributions or functions intended in proposing the typology. As discussed earlier, these are (a) providing an organizing framework for existing focal constructs and (b) providing a systematic conceptual framework for developing new focal constructs and multilevel theories. Some tentative evaluative remarks now can be made using these criteria.

#### *Organizing Framework for Existing Focal Constructs*

By distinguishing the various possible ideal types of composition, the typology provides a framework for organizing existing focal constructs in multilevel research. Instead of taking construct labels at face value, researchers now can gain substantial common understanding by locating constructs in the typology according to the form of composition model in the study. In instances in which no explicit composition models are specified, we still may be able to reconstruct the nature of the functional relationships between constructs and the operational combination process from the researcher's definitions and measurement of constructs and discussion on the research problem and interpretations of the findings. Consider the construct label collective efficacy. We now can reference the direct consensus composition form and the referent-shift consensus composition form and ask systematic composition questions to clarify the nature of the construct. Is the term used to refer to the group-level construct of team efficacy or the individual-level construct of collective efficacy as described in referent-shift consensus composition? Or is the term applied to a mere aggregation of individual self-efficacy perceptions such that the direct consensus composition rather than the referent-shift consensus composition was assumed?

The typology also helps to refine extant focal constructs in multilevel research. For example, the various dimensions in the multidimensional construct of group learning can now be teased out, with different dimensions corresponding to different forms of composition. Learning dimensions that are process oriented (e.g., proceduralization) probably would require a process composition model, whereas learning dimensions that are outcome oriented (e.g., team efficacy, team metacognitive structures)

probably would require the more static composition, such as the direct consensus or referent-shift consensus compositions.

### *Systematic Conceptual Framework for Developing New Focal Constructs and Multilevel Theories*

Specific forms of composition models in the typology have the potential for composing new explanatory constructs from established ones. For example, one could compose the construct of climate strength based on a dispersion composition model and assess the incremental explanatory value of this dispersion construct in research on organizational climate. For example, in explaining organizational performance, researchers can assess whether climate strength accounts for incremental variance over climate level. Another example of construct development is the notion of *team adaptability*. Several researchers have used this term when discussing learning and transfer in team environments (e.g., Kozlowski et al., 1994), but there has been no explicit definition of the notion in the research on teams. On the basis of present typology, we could perhaps distinguish between a static form and a dynamic form of team adaptability corresponding to an individual difference perspective and a learning perspective, respectively (Chan, 1997). In the static form, we begin with the individual-level construct of adaptability defined in terms of a stable individual difference characteristic. We then specify a direct consensus composition for composing the individual-level construct to the team level. We now could develop hypotheses concerning the new construct. For example, we could hypothesize that static team adaptability is positively associated with team morale. In the dynamic form, we begin with the individual-level construct of adaptability defined in terms of a process in which the individual suppressed proceduralized actions and developed new productions when confronted with novel task demands. We then specify a process composition for composing the individual-level construct to the team level. We now could develop hypotheses concerning the new construct. For example, we could hypothesize that team mastery orientation is positively associated with dynamic team adaptability.

The proposed typology offers the researcher a systematic way to broaden his or her conceptualization of focal constructs in a multilevel study, explore alternative interpretations of initial results, provide more rigorous tests of hypotheses, or modify theories or develop new ones. Consider the example of the researcher who is attempting to aggregate lower level units to represent a construct at a higher level (i.e., adopting a direct consensus composition model). The researcher finds high within-group agreement in some groups but not in others. Without alternative composition models, what should the researcher do? De-

pending on the proportion of groups exhibiting low within-group agreement, it is likely that the researcher either will abandon considerations of higher level constructs and focus on the original lower level units or will settle for equivocal results. However, as indicated by the proposed typology, direct consensus composition is one of several forms of composition. It may be possible to specify an appropriate dispersion model that reconceptualizes within-group agreement as the focal higher-level construct as opposed to a statistical criterion for aggregation. The original grouping variable still may be related to the focal dispersion construct. For example, in the climate example, the researcher could develop a dispersion composition model that specifies climate strength as the focal dispersion construct and could proceed to examine possible reasons for the observed variation in climate strength (indexed by within-group agreement) across organizations (i.e., using the dispersion construct as a dependent variable) or relate climate strength to organizational performance (i.e., using the dispersion construct as an independent variable). Alternatively, different composition models could be applied to the same data in a study. As noted by an anonymous reviewer, instead of replacing direct consensus with dispersion composition, the researcher also could consider combining the two models to provide a more rigorous test of the original hypothesis. For example, the dispersion form (i.e., organizational climate strength) of the original construct (i.e., organizational climate) could be controlled when examining the association between organizational climate level and organizational performance (i.e., using the dispersion construct as covariate). In this way, the researcher could examine whether climate strength is an issue while testing the original relationship of interest. Finally, besides being used as a dependent variable, independent variable, or covariate, the different composition form of the original construct could be used as a moderator. For example, with some adequate theory, the researcher may hypothesize that the effect of organizational climate level (i.e., the original construct) on organizational performance is moderated by organizational climate strength (i.e., the dispersion form). Being cognizant of alternative forms of composition models is especially helpful in the exploratory stages of conceptualization, and it reduces the probability of a premature abandonment of data or a mischaracterization of unexplained variance as error.

### Concluding Remarks

Without adequate composition models, researchers have no clear statements of and linkages between conceptual definitions of constructs and operationalizations. In addition, we have no conceptual basis for using data collected at one level (e.g., individual) to operationalize a

construct at another level (e.g., group). For studies using poorly composed constructs, replications and extensions of findings would be difficult or meaningless. Theory development and testing would be seriously impeded if not impossible. It is likely that future researchers will specify new composition forms in addition to the five ideal types in the proposed typology. The typology is by no means exhaustive. The contribution of the typology is that it provides a systematic and relatively comprehensive framework for organizing existing focal constructs and developing new constructs and theories in multilevel research. It is hoped that the present typology would motivate researchers to pay more attention to composition issues and stimulate more research on composition models.

### References

- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122–147.
- Bliese, P. D., & Halverson, R. R. (in press). Group size and measures of group-level properties: An examination of eta-squared and ICC values. *Journal of Management*.
- Brown, K. G., Kozlowski, S. W. J., & Hattrup, K. (1996, August). *Theory, issues, and recommendations in conceptualizing agreement as a construct in organizational research: The search for consensus regarding consensus*. Paper presented at the annual meeting of the Academy of Management, Cincinnati, OH.
- Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, 101, 147–158.
- Chan, D. (1996). Cognitive misfit of problem-solving style at work: A facet of person–organization fit. *Organizational Behavior and Human Decision Processes*, 68, 194–207.
- Chan, D. (1997, August). *Individual difference and learning perspectives on the construct of adaptability: An integrative person–situation interaction approach*. Paper presented at the annual meeting of the Academy of Management, Boston.
- Chan, D., & Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology*, 82, 143–159.
- Cronbach, L. J. (1976). *Research on classrooms and schools: Formulation of questions, design, and analysis* (Unpublished report). Stanford, CA: Stanford University, Stanford Evaluation Consortium.
- Dedobbeleer, N., & BeLand, F. (1991). A safety climate measure for construction sites. *Journal of Safety Research*, 22, 97–103.
- Drasgow, F. (1984). Scrutinizing psychological tests: Measurement equivalence and equivalent relations with external variables are central issues. *Psychological Bulletin*, 95, 134–135.
- Edwards, J. R. (1991). Person–job fit: A conceptual integration, literature review, and methodological critique. In *International review of industrial/organizational psychology* (Vol. 9, pp. 283–357). London: Wiley.
- George, J. M. (1990). Personality, affect, and behavior in groups. *Journal of Applied Psychology*, 75, 107–116.
- George, J. M., & James, L. R. (1993). Personality, affect, and behavior in groups revisited: Comment on aggregation, levels of analysis, and a recent application of within and between analysis. *Journal of Applied Psychology*, 78, 798–804.
- Gist, M. E. (1987). Self-efficacy: Implications for organizational behavior and human resource management. *Academy of Management Review*, 12, 472–485.
- Glick, W. H. (1985). Conceptualizing and measuring organizational and psychological climate: Pitfalls in multilevel research. *Academy of Management Review*, 10, 601–616.
- Golembiewski, R. T., Billingsley, K., & Yeager, S. (1976). Measuring change and persistence in human affairs: Types of change generated by OD designs. *Journal of Applied Behavioral Science*, 12, 133–157.
- Guzzo, R. A., Yost, P. R., Campbell, R. J., & Shea, G. P. (1993). Potency in groups: Articulating a construct. *British Journal of Social Psychology*, 32, 87–106.
- Hackman, J. R. (1992). Group influences on individuals in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3, pp. 199–267). Palo Alto, CA: Consulting Psychologists Press.
- Hannan, M. T. (1971). *Aggregation and disaggregation in sociology*. Lexington, MA: Heath.
- Hogarth, R. M., Gibbs, B. J., McKenzie, C. R. M., & Marquis, M. A. (1991). Learning from feedback: Exactingness and incentives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 734–752.
- House, R., Rousseau, D. M., & Thomas-Hunt, M. (1995). The meso paradigm: A framework for the integration of micro and macro organizational behavior. *Research in Organizational Behavior*, 17, 71–114.
- James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology*, 67, 219–229.
- James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, 69, 85–98.
- Judge, T. A., & Ferris, G. R. (1992). The elusive criterion of fit in human resources staffing decisions. *Human Resource Planning*, 15, 47–67.
- Kanfer, R., & Ackerman, P. L. (1989). Motivational and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74, 657–690.
- Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of Psychology*, 44, 23–52.
- Kirton, M. J. (1976). Adaptors and Innovators: A description and a measure. *Journal of Applied Psychology*, 61, 622–629.
- Klein, K. J., Dansereau, R. G., & Hall, R. J. (1994). Levels issues in theory development, data collection, and analysis. *Academy of Management Review*, 19, 195–229.
- Kozlowski, S. W. J., Gully, S. M., Nason, E. R., Ford, J. K., Smith, E. M., Smith, M. R., & Futch, C. J. (1994, April). *A composition theory of team development: Levels, content, process, and learning outcomes*. Paper presented at the 9th

- Annual Conference of the Society for Industrial and Organizational Psychology, Nashville, TN.
- Kozlowski, S. W. J., Gully, S. M., Salas, E., & Cannon-Bowers, J. A. (1996). Team leadership and development: Theory, principles, and guidelines for training leaders and teams. *Advances in Interdisciplinary Studies of Work Teams*, 3, 253–291.
- Kozlowski, S. W. J., & Hattrup, K. (1992). A disagreement about within-group agreement: Disentangling issues of consistency versus consensus. *Journal of Applied Psychology*, 77, 161–167.
- Kristof, A. L. (1996). Person–organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, 49, 1–48.
- Langbein, L. I., & Lichtman, A. J. (1978). *Ecological inference*. Beverly Hills, CA: Sage.
- Lindell, M. K., & Brandt, C. J. (1997). Measuring interrater agreement for ratings of a single target. *Applied Psychological Measurement*, 21, 271–278.
- Manz, C. C., & Sims, H. P. (1987). Leading workers to lead themselves: The external leadership of self-managing work teams. *Administrative Science Quarterly*, 32, 106–128.
- McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development*, 58, 110–133.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55, 107–122.
- Muthen, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods and Research*, 22, 376–398.
- O'Reilly, C. A., III, Chatman, J., & 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. (1993a). Comparing correlations based on individual level and aggregated data. *Journal of Applied Psychology*, 78, 569–582.
- Ostroff, C. (1993b). 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., & Rothausen, T. J. (1997). The moderating effect of tenure in person–environment fit: A field study in educational organizations. *Journal of Occupational and Organizational Psychology*, 70, 173–188.
- Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin*, 114, 552–566.
- Roberts, K. H., Hulin, C. L., & Rousseau, D. M. (1978). *Developing an interdisciplinary science of organizations*. San Francisco: Jossey-Bass.
- Rousseau, D. M. (1985). Issues of level in organizational research: Multi-level and cross-level perspectives. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (pp. 1–37). Greenwich, CT: JAI Press.
- 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 (Eds.), *Teams: Their training and performance* (pp. 3–29). Norwood, NJ: Ablex.
- Schmitt, N. (1982). The use of analysis of covariance structures to assess beta and gamma change. *Multivariate Behavioral Research*, 17, 343–358.
- Schmitt, N., Pulakos, E. D., & Lieblein, A. (1984). Comparison of three techniques to assess group-level beta and gamma change. *Applied Psychological Measurement*, 8, 249–260.
- Sundstrom, E., De Meuse, K. P., & Futrell, D. (1990). Work teams: Applications and effectiveness. *American Psychologist*, 45, 120–133.
- 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.
- Yammarino, F. J., & Markham, S. E. (1992). On the application of within and between analysis: Are absence and affect really group-based phenomena? *Journal of Applied Psychology*, 77, 168–176.
- Zohar, D. (1980). Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology*, 65, 96–102.

Received December 2, 1996

Revision received October 30, 1997

Accepted October 31, 1997 ■