On the Use of the Coefficient of Variation as a Measure of Diversity

ARTHUR G. BEDEIAN
KEVIN W. MOSSHOLDER
Louisiana State University

*The coefficient of variation (*V*) is a statistical measure commonly used for comparing diversity in work groups. It has been employed by organizational researchers to index the relative internal variability of top-management teams, task groups, boards of directors, departments, and other social aggregates on numerous dimensions. Commenting on its widespread application, this article reviews cautions and pitfalls associated with its use for this purpose. Research implications associated with using* V *are also discussed.*

It is generally recognized that workforce diversity in the United States will increase well into the next century. Accordingly, organizational researchers have turned their attention to developing ways of identifying and accentuating the positive benefits while simultaneously minimizing the negative effects of demographic diversity on work group dynamics (e.g., Milliken & Martins, 1996; Pelled, 1996; Shaw & Barrett-Power, 1998; Tsui & Gutek, 1999). In doing so, researchers have conceptualized and measured diversity in various ways (Lawrence, 1997). The purpose of this article is to highlight cautions and pitfalls associated with the use of a popular statistical index, the coefficient of variation (V), for comparing diversity across work groups. We suggest that the selection of this index for measuring relative variation between groups appears to be based on common practice rather than sound methodological considerations. As we will show, failure to appreciate these cautions and pitfalls could have research implications.

Authors' Note: The helpful vetting of Paul D. Allison, Terry C. Blum, Glenn R. Carroll, and J. Richard Harrison, and especially the reproofing of David A. Harrison, on an earlier draft manuscript is gratefully acknowledged. We are particularly indebted to Dan B. Rinks, Michael Smithson, and Michael C. Sturman for invaluable technical assistance. Correspondence concerning this article should be addressed to the authors at Department of Management, Louisiana State University, Baton Rouge, LA 70803-6312, USA. Electronic mail may be sent to abede@lsu.edu.

Organizational Research Methods, Vol. 3 No. 3, July 2000 285-297

The Coefficient of Variation Defined

According to Williams and O'Reilly (1998), *V* is among the most commonly used statistics for assessing the consequences of group-based demographic differences in organizations. As an indicator of diversity, *V* has been used by organizational researchers to index and compare the internal variability of top-management teams, task groups, boards of directors, departments, and other social aggregates on numerous dimensions. As typically estimated, *V* for a sample is defined as a proportion, $V = s/\bar{x}$ (Equation 1), where s designates the sample standard deviation and $\bar{x}$ is the sample (arithmetic) mean such that $\bar{x} \neq 0$.[1] Thus, *V* indicates how large within-group differences among scores on a response variable tend to be in comparison to their average magnitude. In finite populations or samples, *V* has a lower bound of zero, indicating complete uniformity, and a natural upper bound defined by Equation 1, which as demonstrated by Martin and Gray (1971, Equation 4) simplifies to $\sqrt{n-1}$, where n is sample size, when all cases have zero values save one (i.e., a social aggregate in which only one member possesses the dimension under consideration). Because $s/\bar{x}$ may be a small quantity, some sources express *V* as a percentage by multiplying the right-hand side of Equation 1 by 100. This, however, may be inappropriate because, as a function of n, *V* can exceed unity (Martin & Gray, 1971).

In so much as the variance and standard deviation of a frequency distribution are dependent on the magnitude of their corresponding data, they are considered absolute measures of dispersion. In counterpoint, *V* was specifically invented to eliminate the influence of absolute size on variability (e.g., to compare variability in the body weight of males and females; Pearson, 1897). It does so by indexing the dispersion of a data set relative to its own mean. Furthermore, in that s and $\bar{x}$ are expressed in identical units, *V* is divorced from any scale or unit of measurement; that is, given that the numerator and denominator for *V* are dimensionally equivalent, their units of measurement cancel (Simpson, Roe, & Lewontin, 1960, p. 90). The fact that *V* has no units underscores that it is a relative measure and considered a pure number, independent from units of any measure.

Sensitivity to Scale

Given that *V* is scale free, multiplying each input (x) from which *V* is computed by a constant (k) will leave its value unchanged (i.e., $x \rightarrow kx$, $k > 0$). This occurs because uniform increases (or decreases) in scale (e.g., changing measurement units from inches to centimeters) do not constitute a real change in a frequency distribution's shape or form. Correspondingly, as all inputs are increased by the same proportion, their relative differences remain the same and, thus, so does the value of *V* (Stine, 1989).

Sensitivity to Location

Whereas *V* is invariant under multiplication by a constant, it is not invariant to scale transformations that involve adding (or subtracting) a constant to each input (i.e., $x \rightarrow x + k$, $k \neq 0$). When a positive constant is added to the inputs from which *V* is computed, the mean of each input increases, yet the standard deviation of the inputs' underlying distribution remains the same. Consequently, although the degree of dispersion is unchanged, *V* decreases as the mean value of its inputs increases relative to their underlying distribution's standard deviation. For this reason, it is generally advised that *V* may be meaningfully calculated only for ratio-scale variables, such as annual income in dollars, having a true (i.e., naturally occurring) fixed zero-point (Allison, 1978). *V* may yield a spurious value when applied to variables measured on an interval scale, such as political attitudes or occupational status, whose means and standard deviations are established by arbitrary convention, or variables measured on either an ordinal scale, such as rankings, or a nominal scale, such as race or religion, for which ratios between numbers are meaningless (Agresti & Agresti, 1977; Allison, 1982).

The Coefficient of Variation: Problems in Application

As an index of diversity, *V* has been widely used in the fields of economics and sociology to compare social inequality across nations, cities, and other social aggregates. Inspired by Pfeffer's (1981, 1983) theoretical essays describing the likely effects of demographic diversity on work group dynamics, organizational researchers have since compared the consequences of different demographic distributions on a wide array of work-related outcomes (see, e.g., Carroll & Harrison, 1998). In addition to measuring income inequality, *V* has been used by organizational researchers to compare variation across work groups on such ratio-scale variables as employment levels (measured as number of employees), and demand stability and return on assets (both measured in dollars). Our review of the demographic diversity literature suggests, however, that although *V* has been used appropriately in such instances, both analytical and theoretical considerations have been seemingly overlooked in other applications. We address two analytical considerations: using *V* with data reflecting a less than ratio-level structure, and adjusting for differences in *n* when using *V* to compare diversity across different size groups. Commenting on the theoretical underpinnings most commonly used to explain demographic diversity effects, we then question whether *V* should be used to index the relative internal variability of work groups.

Level of Measurement

As measurement techniques are transmitted over time, details relating to their application are sometimes inadvertently omitted, new information is overlooked, and errors in reporting become solidified (Bedeian, Armenakis, & Randolph, 1988). The impact of these errors may remain undetected as they are disseminated throughout a discipline, even in the face of counterfactual information. The unqualified notion that *V* is an appropriate measure for interval-level data appears to be such a case.

Interval-level measurement allows for the analysis of differences between variables, but does not generally permit the study of their proportionate magnitudes (i.e., ratio comparisons).[2] Allison (1978), among others, has noted that this typically rules out the use of *V* with interval-level data, which by definition have arbitrary zero points and ranges. Confusion over this in the demographic diversity literature seems widespread. For example, one finds *V* described as "a measure of dispersion for interval-scaled data" (Wiersema & Bird, 1993, p. 1010). Interestingly, Allison (1978) is often cited to support the assertion that because *V* is scale invariant, it is preferred to a sample's standard deviation or variance for use with interval-level variables (e.g., Boeker, 1997; Jackson et al., 1991; Wiersema & Bantel, 1992). Although Allison

(1978) acknowledges that standard deviation and variance statistics are not scale invariant, he does not conclude that *V* is preferred for variables measured on an interval scale.

In organizational research, common perceptual and attitudinal variables such as intended turnover or job involvement are generally measured at best on an interval- rather than ratio-scale. Using *V* to describe work group variability with regard to interval-scale data relating to such response variables is a case in point. For example, Harrison, Price, and Bell (1998) presented results of an investigation into the impact of attitudinal diversity on group social integration. They hypothesized that time would enhance the effects of attitudinal diversity (i.e., variability in a work group's overall satisfaction, organizational commitment, satisfaction with supervision, and work satisfaction) on group cohesiveness and used *V* to index group variability on all four attitudinal variables. Elsewhere, in a study of the relationship between values and organization structure, Hinings, Thibault, Slack, and Kikulis (1996) computed *V* scores to assess within-group unanimity across 35 amateur sports organizations for a set of seven organizational values (e.g., commitment to professional involvement, commitment to government involvement).

Although the application of *V* to variables measured on an interval-level for the purpose of comparing relative variation between groups is generally inappropriate, Allison (1978) does show that even in the absence of identical means, such comparisons can be valid under certain circumstances. These circumstances arise in comparisons of inequality for two distributions that are each measured from the same arbitrary zero point and on a common interval scale. Moreover, it must be assumed that there is a non-negative ratio scale underlying the interval scale being observed. Given that specific algebraic inequalities are satisfied, a test for differences in relative variation across distributions can be conducted (Allison, 1978; Equation 14). As an example, Allison (1978) cited the case of social power (i.e., occupational status) measured on an interval scale. He suggested that whereas zero power or powerlessness is a meaningful concept, negative power or occupational status is not. Accordingly, he concluded that it is reasonable to assume that a non-negative ratio scale underlies the measurement of social power when operationalized as occupational status. Using data on racial inequalities, Allison (1978) showed that occupational status was more unequally distributed among a sample of Black females than among a sample of White females. Once again, this assumes that a non-negative ratio scale underlies occupational status, an assumption with which some may not agree.

Although the preceding discussion has focused on instances where interval-level data were involved, examples may also be cited wherein nominal-level and dichotomous measures have been used to calculate and compare *V*s for such variables as educational background (e.g., Haleblian & Finkelstein, 1993) and heterogeneity of top-management team tenure (e.g., Wiersema & Bantel, 1992). Whereas diversity measures for nominal data are available, it is generally recommended that they not be considered equivalent to measures applied to ratio-level data. Not being scale invariant, their values are quite sensitive to the number and construction of their underlying categories and are, thus, more similar to measures of dispersion than diversity (Allison, 1981).

Comparisons of Variability Across Groups of Differing Size

Depending on one's theoretical perspective, a second potentially problematic consideration involves comparing *V*s across groups of differing size. Because the maximum value of *V* is a function of *n*, some have suggested normalizing or adjusting such comparisons for differences in sample or group sizes (Martin & Gray, 1971). Given that the upper limit of *V* is $\sqrt{n-1}$, a work group in which $N = 5$ may thus receive a lower variation value than a larger group in which $N = 10$ simply due to its size and, consequently, because the upper limit of *V* is 2.00 in the first group but 3.00 in the second. At the same time, as Smithson (1982) notes, when using *V* for comparative purposes, the point is not to simply set an upper limit, but to ensure that the same variation value is obtained when comparing groups whose sole difference is sample size. This prompted Smithson (1982) to create V'', defined as follows:

$$\left(\frac{\left(\sum_{i=1}^{N} X_i^2 - N\overline{X}^2\right)}{\left(\sum_{i=1}^{N} Q_i^2 - N\overline{X}^2\right)}\right)^{1/2}, \tag{2}$$

where Q_i equals the *i*th data value in a set of researcher specified *N* values that produce the maximum variation in a response variable. V'' ranges from zero to unity and, thus, may be interpreted as a proportion of its maximum possible value. Moreover, because V'' is normalized with respect to *n*, it yields the same value when comparing groups that differ only in sample sizes.

Accessing the literature indexed by UnCover and in the Web of Science using "coefficient of variation" as a key term, as well as cross referencing bibliographies and literature reviews (e.g., Tsui & Gutek, 1999; Williams & O'Reilly, 1998), we were able to identify 36 studies published between 1984 and 1999 on work group diversity. *V* was used as a measure of diversity in 29 of these studies. The studies may be characterized as focusing on top-management teams, work groups/units, and project/product teams. Table 1 summarizes the attributes included in these studies, as well as the specific effects analyzed. Tenure (organization and team) was the most often analyzed attribute, with age being the second most frequently studied. Other attributes investigated included cultural values, functional background, attitudinal diversity, and education level. Mean group sizes and standard deviations varied across all studies for which this information was provided. None of the studies corrected for *n* in computing *V*s for groups of different sizes.

To appreciate the impact of sample size on *V* across groups (e.g., work teams, where *n* usually varies), consider the relative diversity of four groups with identically distributed scores (on a 7-item ratio measure with anchors that range from 0 to 5) but differing sample sizes. The sample size, scores, mean, and standard deviation for each group are given in Table 2 and are as follows: Group A ($\bar{x} = 28.0$, $s = 7.00$, $n = 3$); Group B ($\bar{x} = 28.0$, $s = 6.26$, $n = 6$); Group C ($\bar{x} = 28.0$, $s = 5.99$, $n = 12$); Group D ($\bar{x} = 28.0$, $s = 5.84$, $n = 24$). Despite having the same central tendency and score distributions, their varying sample sizes result in different standard deviations and, thus, varying *V*s. As also shown in Table 2, calculating *V*s for the four groups yields the values .25, .22, .21, and

Table 1
Attributes and Outcomes Analyzed in 29 Diversity Studies Using the Coefficient of Variation

Study	*Sample*[a]	*Attributes Studied*	*Outcomes Analyzed*
1. Ancona & Caldwell (1992)	45 product teams (≅10; 6.2)	team tenure	team performance
2. Bantel & Jackson (1989)	199 TMT (6.3;1.64)	age, team tenure	technical and administrative innovation
3. Boeker (1997)	67 TMT (NR; NR)	team tenure	strategic change
4. Elron (1997)	22 TMT (6-7;1.86)	cultural values	team and organizational performance
5. Glick, Miller, & Huber (1993)	79 TMT (4.4; NR)	age, company, and team tenure, functional background	organizational performance
6. Greening & Johnson (1996)	96 TMT (5.6; 2.52)	age, team tenure	crisis events
7. Haleblian & Finkelstein (1993)	47 TMT (3.39; 1.46)	CEO dominance	organizational performance
8. Harrison, Price, & Bell (1998)	39 work groups (11.26; 4.52)	attitudinal diversity	group cohesiveness
9. Haveman (1995)	5,833 managers (NR; NR)	tenure	managerial turnover
10. Hinings, Thibault, Slack, & Kikulis (1996)	35 NCO (NR; NR)	organizational values	organizational structure
11. Jackson et al. (1991)	93 TMT (7.44; 3.6)	age, team tenure, education	turnover, promotions
12. Keck (1997)	104 TMT (5.27; NR)	team tenure	team structure across time
13. Keck & Tushman (1993)	104 TMT (5.3; NR)	team tenure	team context and structure
14. Keller (1994)	98 project teams (NR; NR)	team tenure	project performance
15. Knight et al. (1999)	76 TMT (8.0; NR)	age, team tenure, education	strategic consensus

16. Kosnik (1990)	53 TMT (NR; NR)	team tenure	greenmail decisions
17. Michel & Hambrick (1992)	134 TMT (6.18; 2.68)	team tenure	TMT characteristics
18. Murray (1989)	84 TMT (NR; NR)	age, company, and team tenure	organizational performance
19. O'Reilly, Caldwell, & Barnett (1989)	20 work groups (4.11; .85)	age, team tenure	turnover, social integration
20. O'Reilly, Snyder, & Boothe (1993)	24 TMT (6.25; 2.24)	age, tenure	organizational change
21. Pelled, Eisenhardt, & Xin (1999)	45 TMT (≡10; 3.2)	age, team tenure	conflict
22. Simons (1995)	57 TMT (6.06; 2.80)	education, team tenure	organizational performance
23. Simons, Pelled, & Smith (1999)	57 TMT (6.06; 2.80)	company tenure, education	organizational performance
24. Smith et al. (1994)	53 TMT (5.2; 3.66)	company and industry experience	return on investment, sales growth
25. Wagner, Pfeffer, & O'Reilly (1984)	31 TMT (NR; NR)	date-of-entry, team tenure	turnover
26. Westphal & Zajac (1995)	413 TMT (NR; NR)	age	demographic similarity
27. Wiersema & Bantel (1992)	87 TMT (4.3; 0.95)	age, company, and team tenure	strategic change
28. Wiersema & Bird (1993)	40 TMT (5.6; NR)	age, company, and team tenure	turnover
29. Zajac & Westphal (1996)	413 TMT (13.77; 3.36)	age	demographic similarity

Note. TMT = top-management teams; NR = not reported; NCO = national sports organizations.
a. Sample mean and standard deviation are reported in parentheses.

Table 2
The Impact of Sample Size on *V* Across Groups

Group	n	*Scores*	M	SD	V	V(n)	V″
A	3	21, 28, 35	28.0	7.00	.25	.18	.44
B	6	21, 21 28, 28 35, 35	28.0	6.26	.22	.10	.44
C	12	21, 21, 21, 21 28, 28, 28, 28 35, 35, 35, 35	28.0	5.99	.21	.06	.44
D	24	21, 21, 21, 21 21, 21, 21, 21 28, 28, 28, 28 28, 28, 28, 28 35, 35, 35, 35 35, 35, 35, 35	28.0	5.84	.20	.04	.44

.20, respectively. For comparison purposes, these values were divided by $\sqrt{n-1}$ to account for their conventional maxima (Martin & Gray, 1971), yielding normalized *V*s or *V*(*n*) values of .18, .09, .06, and .04, respectively. This contrasts with sample-size corrected *V*s (i.e., *V*″s) for all four groups that equal .44. Thus, when used for comparative purposes, *V* (whether normed or unnormed) may misrepresent the actual uniformity among groups whose sole difference is sample size.

With groups in some studies ranging widely in size and in other studies group size not reported, it is possible that the reported *V*s offer a different picture of the diversity present within the groups than would *V*″. As the preceding illustration demonstrates, as sample sizes across groups increase, corrections for *n* exert less influence. The grand mean of the sample sizes reported in Table 1 is 6.69, a small enough *n* such that *V*″ would differ from *V*. Of course, the exact conditions under which using *V*″ instead of *V* would impact study results needs further investigation, preferably in controlled simulations. Our position is simply that, because the group size adjustment is available, its computation would be recommended until such time that simulation results suggest otherwise.

Should *V* Be Used to Assess Demographic Diversity?

Growing out of the organizational demography research tradition, diversity studies have tended to focus on generic demographic variables, such as age, race, gender, tenure, and education, that traditionally have been theorized to correlate with behavior. Because individuals can differ along myriad dimensions, other variables (e.g., income, marital status, exempt/nonexempt, religion) have also been considered by researchers studying organizational diversity.[3] The theoretical underpinnings most commonly used to explain diversity effects, social integration and cognitive diversity (see Carroll & Harrison, 1998), are rooted in the idea that demographic variability systematically influences members' social or task interactions so as to affect work group functioning. In calculating empirical associations between demographic variability and such outcomes, organizational researchers have sought to determine which differences make a difference in group functioning.

The exact details of how these differences affect particular work group outcomes are seldom articulated (Lawrence, 1997). Moreover, as observed by Carroll and Harrison (1998, p. 637), many of the explanations that have been advanced are seemingly based on "questionable and unstudied assumptions." For instance, when organizational researchers employ V to compare diversity in age, tenure, or education across work groups, they are implicitly treating age/tenure/education at opposite ends of a frequency distribution as different in order or magnitude, essentially implying that diversity in these attributes has a negatively accelerating effect (D. A. Harrison, personal correspondence, January 22, 1999).

To illustrate this point, consider the use of V to index age diversity across two work groups. Suppose two groups are of the same size and their age scores have identical standard deviations, 4.0. In the first group, the mean age is 21 years; however, in the second, the mean age equals 55. Applying the formula for V, one finds that despite identical standard deviations (within-group dispersion), the resulting values for V will be different, .190 and .073, respectively, indicating that age diversity in the younger group is more than 2.60 times that in the older group. This means that to have the same impact in diversity terms, as gauged by V, there must be greater age differences (i.e., greater dispersion) among a group of older workers as compared to a group of younger workers. Mathematically, the reason for this result is obvious. Whereas the groups' standard deviations (which measure each group's internal variability) are identical, their age scores (which in no way measure internal variability) cluster around different means.

There may be cases in which there are theoretical reasons for assuming that age diversity has a negatively accelerating effect. In such cases, because there is supporting theory, scale invariant indices such as V may be useful. In the absence of theory explaining why "units of age" for younger workers have a different meaning than for older workers, however, using V is inappropriate. Of course, although we selected age to demonstrate conceptual problems with using V, the same caution applies to other demographic variables (e.g., tenure, years of education, months of experience in an industry or with a company).

As noted, the theoretical framework most commonly invoked to explain diversity effects, social integration and cognitive diversity, are based on the notion that demographic variability systematically influences members' social or task interactions so as to influence work-related outcomes. Within these theoretical perspectives, demographic distinctions among individuals reflect differences without implying order or magnitude. As a consequence, the empirical relations underlying such distinctions cannot be mapped onto groups to form a meaningful numerical structure (Stine, 1989). Thus, despite differences, groups cannot be arrayed in terms of the dimensions on which diversity is calculated, because the dimensions themselves cannot be arranged in a defined order that connotes *more*, *higher*, or *greater*. This is easily understood in connection with variables like gender, which may be coded as either male or female, with neither designation reflecting an inherent ordering nor inequality. It is not as easily seen with variables like age or tenure, which are measured in units (e.g., calendar years) for which group differences (e.g., mean group age) are readily calculable. What must be remembered is that unless there is theory that specifies how age maps onto a numerical structure representing order or inequality among groups (e.g., more is better), then the units used to measure age are just as arbitrary as numerical designations for male or female.

Conclusion

Using V with data reflecting a less than ratio-level structure is generally inappropriate except under certain conditions. Moreover, given that V is a function of sample size, it may be useful to adjust for differences in n when using V to compare diversity across different size groups. Furthermore, those organizational researchers who, in the absence of theory, have used V as a relative diversity index seem to have unwittingly assumed that individual attributes are arranged in some defined order that is meaningful across groups. In doing so, it appears that they have implicitly treated attributes like age or tenure as having a negatively accelerating effect on work group outcomes. We reiterate the importance of remembering what V is and being sure that it is employed within a sound theoretical framework. For as Simpson et al. (1960, p. 90) have noted, although "V is in every case a good measure of relative dispersion, relative dispersion is not always a good measure of variability."

Whereas the present discussion focuses on the coefficient of variation, Tsui and Gutek's (1999) summary of the relevant literature reveals that other diversity measures have been employed by organizational researchers. Second in popularity to V, Taagepera and Ray's (1977) generalized index of concentration (C_n) has also been employed as a measure of relative dispersion. C_n, however, does not adjust for differences in sample or group sizes and implies some defined order of individual attributes, as do V and V''. There are, of course, numerous other measures of diversity. Drawing on different mathematical models, Coulter (1989) presents 31 intragroup inequality measures, 6 intergroup indexes, and 3 inequity indexes that have been used under various circumstances to gauge diversity. In doing so, he critiques their operational similarities and differences and clarifies their theoretical underpinnings. All 40 measures Coulter (1989) reviews have both desirable and undesirable properties depending on the theoretical context. Once again, the key point to be deduced is the importance of matching theoretical concept and statistical measure.

Unless there is a theoretical rationale suggesting that certain advantages accrue to individuals occupying specific points along the continua of typically assessed diversity variables (i.e., age, tenure), basic dispersion indices might well be considered. Thus, using an attribute's standard deviation as a diversity index would be an option, as would be using an attribute's average deviation index (average of absolute value of deviations from the mean) when extreme deviations were judged to unduly influence its standard deviation. Of course, this suggestion applies only to attributes that are assessed using at least an interval-level scale. In the end, the choice of a measure should rest on acknowledged analytical considerations. For, as we have shown, the interpretation of results stemming from the selection of one measure over another may be influenced as much by the measure chosen as by true or real differences in diversity.

Notes

1. Although Equation 1 defines V for a sample distribution such that $\bar{x} \neq 0$, V is most meaningful for variables from population distributions that take only positive values. Accordingly, in the

discussion that follows, zero is excluded to eliminate social aggregates in which no member possesses the characteristic under consideration.

2. The following discussion is based on the representational theory of measurement in which a relation between measurement scales and statistical procedures is implied. For a consideration of this and alternative theories of measurement, see Michell (1986).

3. For a multivariate extension of a diversity index for use with nominal variables, see Lieberson (1969). For a discussion of ordinal dispersion statistics, see Blair and Lacy (2000).

References

Agresti, A., & Agresti, B. (1977). Statistical analysis of qualitative variation. *Sociological Methodology, 10*, 204-237.

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

Allison, P. D. (1981). Inequality measures for nominal data. *American Sociological Review, 46*, 371-373.

Allison, P. D. (1982). Regression analysis of inequality measures applied to interval scales. *Quality and Quantity, 16*, 251-260.

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

Bantel, K. A., & Jackson, S. E. (1989). Top management and innovations in banking: Does the composition of the top team make a difference? *Strategic Management Journal, 10*, 107-124.

Bedeian, A. G., Armenakis, A. A., & Randolph, W. A. (1988). The significance of congruence coefficient: A comment and a statistical test. *Journal of Management, 14*, 559-566.

Blair, J., & Lacy, M. G. (2000). Statistics of ordinal variation. *Sociological Methods & Research, 28*, 251-280.

Boeker, W. (1997). Strategic change: The influence of managerial characteristics and organizational growth. *Academy of Management Journal, 40*, 152-170.

Carroll, G. R., & Harrison, J. R. (1998). Organizational demography and culture: Insights from a formal model and simulation. *Administrative Science Quarterly, 43*, 637-667.

Coulter, P. B. (1989). *Measuring inequality: A methodological handbook.* Boulder, CO: Westview.

Elron, E. (1997). Top management teams within multinational corporations: Effects of cultural heterogeneity. *Leadership Quarterly, 8*, 393-412.

Glick, W. H., Miller, C. C., & Huber, G. P. (1993). The impact of upper-echelon diversity on organizational performance. In G. P. Huber & W. H. Glick (Eds.), *Organizational change and redesign* (pp. 176-214). New York: Oxford University Press

Greening, D. W., & Johnson, R. A. (1996). Do managers and strategies matter? A study in crisis. *Journal of Management Studies, 33*, 25-51.

Haleblian, J., & Finkelstein, S. (1993). Top management team size, CEO dominance, and firm performance: The moderating roles of environmental turbulence and discretion. *Academy of Management Journal, 36*, 844-863.

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, 42*, 96-107.

Haveman, H. A. (1995). The demographic metabolism of organizations: Industry dynamics, turnover, and tenure distributions. *Administrative Science Quarterly, 40*, 586-618.

Hinings, C. R., Thibault, L., Slack, T., & Kikulis, L. M. (1996). Values and organizational structure. *Human Relations, 49*, 885-916.

Jackson, S. E., Brett, J. F., Sessa, V. I. , Cooper, D. M. , Julin, J. A., & Peyronnin, K. (1991). Some differences make a difference: Individual dissimilarity and group heterogeneity as correlates of recruitment, promotions, and turnover. *Journal of Applied Psychology, 76*, 675-689.

Keck, S. L. (1997). Top management team structure: Differential effects by environmental context. *Organization Science, 8*, 143-156.

Keck, S. L., & Tushman, M. L. (1993). Environmental and organizational context and executive team structure. *Academy of Management Journal, 36*, 1314-1344.

Keller, R. T. (1994). Technology-information processing fit and the performance of R&D project groups: A test of contingency. *Academy of Management Journal, 37*, 167-179.

Knight, D., Pearce, C. L., Smith, K. G., Olian, J. D., Sims, H. P., Smith, K. A., & Flood, P. (1999). Top management team diversity, group process, and strategic consensus. *Strategic Management Journal, 20*, 445-465.

Kosnik, R. D. (1990). Effects of board demography and directors' incentives on corporate greenmail decisions. *Academy of Management Journal, 33*, 129-150

Lawrence, B. S. (1997). The black box of organizational demography. *Organization Science, 8*,1-22.

Lieberson, S. (1969). Measuring population diversity. *American Sociological Review, 34*, 850-862.

Martin, J. D., & Gray, L. N. (1971). Measurement of relative variation: Sociological examples. *American Sociological Review, 36*, 496-502.

Michel, J. G., & Hambrick, D. C. (1992). Diversification posture and top management team characteristics. *Academy of Management Journal, 35*, 9-37.

Michell, J. (1986). Measurement scales and statistics: A clash of paradigms. *Psychological Bulletin, 100*, 398-407.

Milliken, F. J., & Martins, L. L. (1996). Searching for common threads: Understanding the multiple effects of diversity in organizational groups. *Academy of Management Review, 21*, 402-433.

Murray, A. I. (1989, Summer). Top management group heterogeneity and firm performance. *Strategic Management Journal, 10*, 125-141,

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

O'Reilly, C. A., III, Snyder, R. C., & Boothe, J. N. (1993). Effects of executive team demography on organizational change. In G. P. Huber & W. H. Glick (Eds.), *Organizational change and redesign* (pp. 147-175). New York: Oxford University Press.

Pearson, K. (1897). On the scientific measure of variability. *Natural Science, 11*, 115-118.

Pelled, L. H. (1996). Demographic diversity, conflict, and work group outcomes: An intervening process theory. *Organization Science, 7*, 615-630.

Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly, 44*, 1-28.

Pfeffer, J. (1981). Some consequences of organizational demography: Potential impacts of an aging work force on formal organizations. In S. B. Kiesler, J. N. Morgan, & V. K. Oppenheimer (Eds.), *Aging: Social change* (pp. 291-329). New York: Academic Press.

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

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

Simons, T. (1995). Top management team consensus, heterogeneity, and debate as contingent predictors of company performance: The complementarity of group structure and process. *Academy of Management Best Proceedings*, 62-66.

Simons, T., Pelled, L. H., & Smith, K. A. (1999). Making use of difference: Diversity, debate, and decision comprehensiveness in top management teams. *Academy of Management Journal, 42*, 662-673.

Simpson, G., Roe, A., & Lewontin, R. C. (1960). *Quantitative zoology* (Rev. ed.). New York: Harcourt Brace.

Smith, K. G., Smith, K. A., Olian, J. D., Sims, H. P., Jr., O'Brannon, D. P., & Scully, J. A. (1994). Top management team demography and process: The role of social integration and communication. *Administrative Science Quarterly, 39*, 412-438.

Smithson, M. (1982). On relative dispersion: A new solution for some old problems. *Quality and Quantity, 16*, 261-271.

Stine, W. W. (1989). Meaningful inference: The role of measurement in statistics. *Psychological Bulletin, 105*, 147-155.

Taagepera, R., & Ray, J. L. (1977). A generalized index of concentration. *Sociological Methods & Research, 5*, 367-384.

Tsui, A. S., & Gutek, B. A. (1999). *Demographic differences in organizations*. Lanham, MA: Lexington.

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

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

Wiersema, M. F., & Bantel, K. A. (1992). Top management team demography and corporate strategic change. *Academy of Management Journal, 35*, 91-121.

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

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

Zajac, E. J., & Westphal, J. D. (1996). Who shall succeed? How CEO/board preferences and power affect the choice of new CEOs. *Academy of Management Journal, 39*, 64-90.

Arthur G. Bedeian is a Boyd Professor in the E. J. Ourso College of Business Administration at Louisiana State University and A&M College. His academic interests include research methodology and workplace behavior.

Kevin W. Mossholder is H. Norman Saurage/Community Coffee Chair in the E. J. Ourso College of Business Administration at Louisiana State University and A&M College. His research interests include multilevel issues and helping behaviors in the workplace.