Small Group Research

http://sgr.sagepub.com

A Meta-Analytic Review of the Productivity and Satisfaction of Democratic and Autocratic Leadership

John Gastil

Small Group Research 1994; 25; 384
DOI: 10.1177/1046496494253003

There exists a wealth of research examining the effects of democratic and autocratic leadership on group productivity and member satisfaction; however, past reviews of this literature have not systematically integrated the results of available quantitative studies. This essay uses a meta-analysis to provide such an integration. Analysis reveals no correlation between democratic/autocratic leadership style and productivity, except when taking into consideration the influence of study setting and task complexity. Results also suggest that democratic leadership has a moderate positive correlation with member satisfaction, but this relationship may be moderated by task complexity. The conclusion discusses the limits of experimentally manipulating democratic leadership and the need for conceptual refinement.

A META-ANALYTIC REVIEW OF THE PRODUCTIVITY AND SATISFACTION OF DEMOCRATIC AND AUTOCRATIC LEADERSHIP

JOHN GASTIL
University of Wisconsin–Madison

The tumultuous political events of the past 3 years have raised hopes for the creation and revitalization of democratic institutions across the globe. In Czechoslovakia, Namibia, El Salvador, and elsewhere, windows of opportunity have opened wide. Unless all of these windows close, social scientists are presented with a remarkable opportunity to assist those persons and organizations who seek to study and create democratic groups and societies.

Just as political scientists and mass media scholars can make valuable contributions to the study of democracy, small group

AUTHOR'S NOTE: *I would like to acknowledge the assistance of James P. Dillard and an anonymous reviewer. The author can be reached at the University of Wisconsin–Madison, Department of Communication Arts, 821 University Avenue, Madison, WI 53706.*

SMALL GROUP RESEARCH, Vol. 25 No. 3, August 1994 384-410

researchers can aid those who seek to understand the democratic process. One body of theory and research that may prove fruitful is the literature on democratic and autocratic methods of leadership. Krips (1992) recently showed the relevance of this subject to current international political developments by explaining the limitations that autocratic leadership places on political and economic reform in Estonia.

The social and scholarly significance of democratic leadership is reflected in the fact that modern academic writings on the subject appeared in the United States as early as the 1920s and 1930s. At the time, theorists sought to discover a form of authority more conducive to the development and maintenance of a democratic society (Busch, 1934; Smith, 1926; Tead, 1935; Whitehead, 1936).

Undoubtedly, the most famous conceptual distinctions were made by Kurt Lewin and his associates, Ronald Lippitt and Ralph K. White (Lewin & Lippitt, 1938; Lewin, Lippitt, & White, 1939; White & Lippitt, 1960). Lewin and his colleagues conducted a pair of experiments that operationalized autocratic, democratic, and laissez-faire styles of leadership. Laissez-faire leadership resulted in unproductive groups and unsatisfied group members, so the contrast between democratic and autocratic leadership became the primary conceptual distinction.

Their simple model was a watershed in the empirical study of democratic leadership and led to a flood of studies and theoretical essays (for reviews, see Anderson, 1959; Bass, 1990; Gibb, 1969; Stogdill, 1974). Since 1939, dozens of studies have investigated the effects of democratic and autocratic leadership, and hundreds of books and articles have discussed these forms of leadership.[1] Even today, the body of research on democratic leadership continues to grow in the United States (Garland & Barry, 1990), Canada and Japan (Chelladurai, Imamura, Yamaguchi, Oinuma, & Miyauchi, 1988), India (Belwal, 1985; Srinivasan & Kamalanabhan, 1986), Portugal (Serpa, Pataco, & Santos, 1991), Poland (Pawlicka, 1987), Finland (Isohanni, 1990), France (Loiseau, 1988), and the former Soviet Union (Chernyshev & Suryaninova, 1990).

THE EFFECTS OF DEMOCRATIC LEADERSHIP

Most of the empirical research in this literature explores the effects of democratic leadership on task-oriented groups and their individual members. Productivity and satisfaction have received the most attention, and this emphasis is not surprising given the historical importance of these outcomes in management theory (e.g., George, 1972) and the larger society.

Unfortunately, the empirical literature on the productivity and satisfaction of democratic leadership has not been well summarized by those who have reviewed or cited it. Political theorists who discuss democratic leadership typically ignore existing research altogether (e.g., Barber, 1984; Nagel, 1987). Textbooks on small group behavior typically mention democratic and autocratic leadership, but their summaries are vague or misleading. Brilhart and Galanes (1989) suggest that democratic leadership is more effective, without giving any clear justification or qualification; others altogether ignore the possible effects of democratic leadership (e.g., Cragan & Wright, 1986; Husband, 1988; Lawler, 1984; Quinn, 1984). Some authors discuss effects but focus on the early experiment by Lewin and associates to the exclusion of all other studies (e.g., Beebe & Masterson, 1990; Brown, 1988; Forsyth, 1990; French, Kast, & Rosenzweig, 1985). Finally, there are those authors who discuss the effects of democratic leadership and possible moderator variables, but moderators are chosen without reference to the broad base of empirical research on democratic leadership (e.g., Barker, Wahlers, Watson, & Kibler, 1987; Jensen & Chilberg, 1991; Johnson & Johnson, 1991; Tubbs, 1988; Wilson & Hanna, 1986).

Even the most recent studies on autocratic and democratic leadership fail to build themselves on the foundation provided by previous research. One example of this pattern is the recent series of studies in sport psychology that followed the introduction of Chelladurai and Saleh's (1978, 1980) leadership scale (e.g., Chelladurai et al., 1988; Dwyer & Fischer, 1990; Garland & Barry, 1990; Schliesman, 1987; Serpa, Pataco, & Santos, 1991; Weiss & Friedrichs, 1986).

Finally, the most current and extensive review of the literature on democratic leadership has serious limitations. In his updated and expanded version of Stogdill's (1974) classic *Handbook of Leadership*, Bass (1990) argues that democratic leadership is, on balance, less productive but more satisfying than autocratic leadership. According to Bass, these relationships are moderated by variables such as the group's history and the maturity, ability, and expectations of the membership. Bass attempts to reconcile the findings of every study reviewed by positing the existence of an unwieldy number of moderators. He mixes traditional qualitative review methods with a primitive version of meta-analysis—a box-score tabulation of positive, zero, and negative correlations (see Hunter & Schmidt, 1990). Because this approach ignores effect sizes, sample sizes, moderators, and outliers, its results are extremely unreliable. Bass's review also suffers from the exclusion of relevant studies and the inclusion of unrelated studies (e.g., studies in which member participation, not democratic leadership, is the independent variable; see Bass, 1990, Table 21.3, p. 423). In sum, current writings on democratic leadership do not base themselves on a thorough and precise summary of the existing literature. This essay aims to provide such a review.

METHOD

Meta-analysis (Hunter & Schmidt, 1990; Rosenthal, 1984) is a useful method for reviewing quantitative data. It allows one to compare the findings of different studies, while taking into account relative sample sizes and other important contextual features (e.g., type of subjects, type of group task). Meta-analysis can specify the average effect sizes of a body of studies, but it can also identify complex moderator relationships that explain seemingly contradictory findings.

STUDY SELECTION

To locate the available studies on democratic leadership, several search strategies were adopted. Computer searches of *Psychlit*

(1974-1992) and the *On-Line Catalog of the Library of Congress* (1950-1992) were conducted. Titles and abstracts were searched for the keywords "democratic leader(ship)" or "autocratic leader(ship)." An attempt was also made to find studies that lacked these particular labels (see O'Keefe, 1991). Additional studies were located by combing the references in relevant reviews (Anderson, 1959; Bass, 1990; Eagly & Johnson, 1990; Gibb, 1969; Miller & Monge, 1986; Stogdill, 1974). As each article was located, its citations were searched for additional studies. This process identified 94 possible studies, spanning from 1939 to 1991.

Following O'Keefe (1991), exclusion criteria were specified. Although a complete list of all excluded studies is not provided, some examples are given below. In addition, study inclusion followed the guidelines of Hale and Dillard (1991): apparent methodological differences were not treated as reasons for exclusion; instead, they were considered potential grounds for partitioning heterogeneous groups of studies.

Five exclusion criteria were employed. First, unpublished materials (i.e., dissertations and master's theses) and studies the author was unable to translate (e.g., Vujtech, 1972) were excluded.[2] Second, studies were omitted if they were qualitative (e.g., Flanders, 1951; Narain, 1964) or lacked the numerical data necessary for extracting effect sizes. Third, studies were not included if they focused on different independent and/or dependent variables (e.g., I. Bhushan, 1971; R. Bhushan & Verma, 1972; Fowler & Rosenfeld, 1979; Srinivasan & Kamalanabhan, 1986; Yukl & Kanuk, 1979).

Fourth, studies were excluded if they failed to operationalize the independent variable (democratic vs. autocratic leadership style). This includes defining styles in terms of a single behavior or using an operationalization too distant from leadership style.[3] An example of the latter is Adams (1952), who assumed that bomber crews' attitudes toward authority directly reflected the leadership styles of their superior officers. Using this logic, Adams operationalized leadership style with measurements of crew member attitudes.

Finally, studies were excluded if they operationalized neither productivity nor satisfaction, per se. For example, low and inconsistent correlations between absenteeism and satisfaction (Hackett &

Guion, 1985) made it difficult to include studies that only operationalized satisfaction as absenteeism (e.g., Ley, 1966).

Applying these five criteria, 39 independent samples from 37 studies were selected for meta-analysis (see Appendix A for a complete list of studies). All of these studies reported correlations or transformable statistics quantifying the association between democratic/autocratic leadership style and group productivity and/or member satisfaction.

EFFECT SIZES

To compare the studies directly, it was necessary to use a common statistic for all effect sizes. The Pearson *r* was chosen, because most readers are already familiar with its interpretation (for a discussion, see Rosenthal, 1984). Using techniques discussed in Cohen (1988), Dillard (1991), Hunter and Schmidt (1990), and Rosenthal (1984), raw data, *t*s, *q*s, *p*s (transformed into *z*s), chi-squares, and—when necessary—*F*s were converted into *r*s for each study. The right-hand columns of Tables 1 and 2 show the statistics from each study that were used to compute *r*, and the third column from the left shows the uncorrected *r* values that were extracted. Positive *r* values indicate a positive relationship between the democratic leadership style and the dependent variable in question.

Reliabilities for the independent and dependent variables were also extracted from the studies. For the studies without reported reliabilities, an artifact distribution was created (Hunter & Schmidt, 1990). Studies were divided into two categories, those that used confederates (leaders given explicit instructions) and those that used untrained leaders whose styles "naturally" differed. The first group had three reported reliabilities, which ranged from .80 to .91, and the other studies in this category were assigned the average alpha, .84. The second category, not surprisingly, had lower reliabilities, which ranged from .62 to .77, and the studies in this category without reported reliabilities were assigned the .71 average.

For productivity, a meaningful artifact distribution was not possible, as only one study (Haythorn, Couch, Haefner, Langham,

text continues on page 395

TABLE 1: *Ns*, *rs*, Alphas, Moderators, and Other Information on Studies Used in Meta-Analysis of Productivity

Author(s) (Date)	N *(Groups)*	r	*Opr*	*Alpha* *IV*	*Alpha* *DV*	*Dich*	*Manip*	*Site*	*Size*	*Cmplx*	*Time*	*A/O*	*sex*	*Statistics Used*
Adams (1954)	25	.20	3	.75	1	—	2	1	11	2	4	4	2	*d* from *x*, *sd*
Argyle et al. (1958)	81	.24	3	.71	1	—	2	1	25	1	5	3	2	*X2*
Berkowitz (1953)	72	.02	1	.71	1	—	2	1	9	3	2	3	1	*r* averaged
Bevan et al. (1958)	8	–.23	2	.84	1	—	1	3	11	3	1	2	1	*q* from *t*
Calvin et al. (1957)	12	.00	1	.84	1	—	1	3	8	2	2	2	1	*r* from *x*
Fox (1957)	4*	.08	3	.84	1	—	1	3	9	3	3	2	2	*d* from *x,t,df*
Haythorn et al. (1956a)	16	.23	3	.62	.71	—	2	3	4	3	2	2	2	*t*
Haythorn et al. (1956b)	16	.15	3	.62	.71	—	2	3	4	3	2	2	2	*d* from *x,sd*
Hendrix et al. (1982)	106	.30	3	.71	1	—	2	1	8	2	5	4	1	*r* from *F*
Kipnis et al. (1981)	113	–.04	1	.84	1	—	1	3	6	1	1	2	1	*r*
Lyle (1961)	12	.00	2	.84	1	50/50	1	3	8	3	1	2	1	*X2* from raw
Mahoney (1967)	283	–.01	3	.71	1	—	2	1	x	9	6	3	9	*r*
McCurdy & Eber	12	–.27	3	.84	1	—	1	3	3	1	1	2	1	*d* from raw
(1953)		–.51	2											
		.00	1											
McCurdy & Lambert														
(1952)	66	–.01	1	.84	1	—	1	3	3	1	1	2	1	*d* from raw
Meade (1967)	4	–.14	3	.84	1	part	1	2	6	3	4	1	2	*X2*
		–.28	2			50/50								
		.00	1											
Meade (1985)	16	.10	3	.84	1	part	1	2	6	3	4	1	2	*X2*
		–.11	2			50/50								
		–.16	1											

Morse & Reimer (1956)	48	–.27	3	.84	1	—	1	2	50	2	6	3	0	*d* from raw
Mullen (1965)	2	.00	1	.71	1	—	2	1	188	1	6	3	1	*X2* from raw
Rosenbaum et al. (1971)	20	.20	1	.84	1	—	1	3	4	2	1	2	2	*r* from *F*
	20	.30	1											
	20	.10	2											
Shaw (1955)	48	–.33	3	.84	1	—	1	3	4	1	1	2	2	*r* from *F*
		–.24	2											
		–.42	1											
Torrance (1953)	33	–.11	2	.84	1	50/50	1	2	13	3	1	4	2	percentages
Wispe (1951)	8*	–.07	2	.91	1	—	1	2	20	3	5	2	9	*z* from *p*
Wu-Tien et al. (1978)	8*	.08	2	.77	1	—	2	1	30	2	5	1	1	*q* from *r*

NOTE: Rows with blank spaces have features identical to those immediately above them.
* = These studies used *S*s as the units of analysis, so reported *r*s should be smaller due to greater variance within cells.
Opr = Operationalization of productivity; 1 = quantity, 2 = quality, 3 = mixed quantity and quality.
Alpha = Coefficient alpha (reliability); IV = independent variable (leadership style), DV = dependent variable (productivity); italicized alphas are reported in studies (or derived from information in studies); alphas of 1 reflect the impossibility of conducting an artifact distribution; all other alphas were obtained through an artifact distribution, using information about *k* (e.g., number of judges or items) when reported in the original study.
Dich = Correction for dichotomization; — = no correction, 50/50 = correction for a fifty-fifty split, part = two or more *r*s were averaged, but only some of them needed to be corrected for dichotomization.
Manip = Manipulation of IV; 1 = instructions given to confederates, 2 = leaders given no instructions, then rated according to style.
Site = Site of the study; 1 = field study, 2 = field experiment, 3 = experiment in laboratory.
Size = Group size (i.e., number of *S*s in each group); *x* = hypothetical experiment (i.e., no actual group size).
Cmplx = Complexity of group task; 1 = simple (e.g., simple information processing, assembly line, memory exercise), 2 = moderate complexity, 3 = complex (e.g., idea generation, creativity requirement, unfamiliar, and complicated task), 9 = unknown.
Time = Duration of the group task studied; 1 = less than one hour, 2 = a period of hours (within one day), 3 = over a period of days, 4 = weeks, 5 = months, 6 = years.
A/O = Age and occupation; 1 = grade school children, 2 = college students, 3 = business employees, 4 = military personnel.
Sex = Sex of *S*s; 0 = all female, 1 = male and female, 2 = all male, 9 = not reported and difficult to infer.
Stats. Used = The statistics used to compute an effect size; raw = raw data in article, nsd = no significant differences reported.

TABLE 2: *N*s, *r*s, Alphas, Moderators, and Other Information on Studies Used in Meta-analysis of Satisfaction

Author(s) (Date)	N (Ss)	r	*Opr*	*Alpha* *IV*	*Alpha* *DV*	*Dich*	*Manip*	*Site*	*Size*	*Cmplx*	*Time*	*A/O*	*Sex*	*Statistics Used*
Berkowitz (1953)	648*	–.07	1	.71	.77	—	2	1	9	3	2	3	1	*r* averaged
Day & Hamblin (1964)	48	.11	2	.84	.81	—	1	3	4	2	1	2	0	*r* from *F*
Foa (1957)	217	.22	3	.71	.90	50/50	2	1	18	2	5	4	2	*X2* from raw
Fox (1957)	36	.34	9	.84	.53	—	1	3	9	3	3	2	2	*t*
		.26	1											
		.08	2											
		.62	3											
Gibb (1951)	200*	–.12	2	.71	.53	—	2	3	10	3	3	4	2	*rho*
Gibb (1951)	100*	–.23	3	.71	.53	—	2	3	10	3	3	2	2	*rho*
Haythorn et al. (1956a)	64*	.15	9	.62	.71	—	2	3	4	3	2	2	2	*d* from *t*
		.26	1											
		.50	2											
		–.34	3											
Haythorn et al. (1956b)	64*	.40	9	.62	.71	—	2	3	4	3	2	2	2	*d* from *x*, *sd*
		.13	1											
		.54	2											
		.48	3											
Hendrix et al. (1982)	848*	.37	9	.71	.81	—	2	1	8	2	5	4	1	*r* from *F*
Kushell & Newton (1986)	144	.31	9	.84	.81	—	1	3	36	2	1	2	1	*r* from *p*
		.31	2											
		.31	3											
Lyle (1961)	132*	.26	1	.84	.81	50/50	1	3	8	3	1	2	1	*X2* from raw
Meade (1967)	24	–.24	9	.84	.81	—	1	2	6	3	4	1	2	*X2*
		–.48	4											
		–.00	3											

Meade (1985)	96	–.06	9	.84	.53	—	1	2	6	3	4	1	2	*X2*
		–.06	3											
		–.06	4											
Morse & Reimer (1956)	208	.10	9	.84	*.82*	—	1	2	50	2	6	3	0	*d* from raw
	201	.09	1		*.73*									
	214	.10	3		*.91*									
Mullen (1965)	376	.30	1	.71	.97	50/50	2	1	188	1	6	3	1	*X2* from raw
Roberts et al. (1968)	115	.19	1	.71	.85	—	2	1	58	3	6	3	9	*d* from *x*, *sd*
Rosenbaum et al. (1971)	60	.20	2	.84	.81	—	1	3	4	2	1	2	2	*r* from *F*
Scontrino (1972)	60	.46	9	.84	*.91*	—	1	3	5	3	1	2	1	*d* from *x*, *sd*
		.41	2											
		.67	3											
		.29	4											
Shaw (1955)	192*	.43	1	.84	.53	—	1	3	4	1	1	2	2	*d* from *F*
Snadowsky (1974)	320*	.19	9	*.80*	.81	—	1	3	4	2	1	2	2	*d* from *F*
		.20	1											
		.18	3											
Stitt et al. (1983)	678*	.19	9	.84	.81	—	1	3	6	1	1	2	1	*z* from *p*
		.16	2											
		.22	3											
Sudolsky & Nathan (1971)	64	.29	9	.84	.53	—	1	x	x	x	x	2	1	*z* from *p*
		.29	1											
		.29	3											
Thelen & Withall (1949)	10	.25	1	*.81*	1	—	1	2	10	3	2	1	9	*X2* from raw
Vroom & Mann (1960)	240*	–.39	9	.71	.69	—	2	1	10	1	6	3	2	*r*
		–.37	1											
		–.41	3											
Vroom & Mann (1960)	1120*	.42	9	.71	.69	—	2	1	40	1	6	3	2	*r*
		.42	1											
		.41	3											

continued

TABLE 2: Continued

Author(s) (Date)	N (Ss)	r	Opr	Alpha IV	Alpha DV	Dich	Manip	Site	Size	Cmplx	Time	A/O	Sex	Statistics Used
White & Lippitt (1960)	20	.45	3	.84	.53	—	1	2	5	3	4	1	2	*X2* from raw
Wischmeier (1955)	40	–.15	3	.84	*.85*	—	1	3	5	3	2	2	9	*z* from *p*
		.03	2											
		–.27	3											
Wispe (1951)	160	–.13	1	*.91*	.90	—	1	2	20	3	5	2	9	*z* from *p*
Ziller (1957)	168*	.22	9	.84	.69	—	1	2	11	3	1	4	2	*z* from *p*
		.00	2											
		.41	3											

NOTE: Rows with blank spaces have features identical to those immediately above them.
* = These studies used groups as the units of analysis, so reported *r*s should be larger due to less variance within cells.
Opr = Type of satisfaction measured; 1 = global/morale, 2 = task/goal progress, 3 = leader(ship), 4 = membership in group, 9 = combination of types.
Alpha = Coefficient alpha (reliability); IV = independent variable (leadership style), DV = dependent variable (satisfaction); italicized alphas are reported in studies (or derived from information in studies); the alpha of 1 was for a unique measure of satisfaction that had no reported reliability and was not comparable to other studies' methods; all other alphas were obtained through an artifact distribution, using information about *k* (e.g., number of judges or items) when reported in the original study.
Dich = Correction for dichotomization; — = no correction, 50/50 = correction for a fifty-fifty split.
Manip = Manipulation of IV; 1 = instructions given to confederates, 2 = leaders given no instructions, then rated according to style.
Site = Site of the study; 1 = field study, 2 = field experiment, 3 = experiment in laboratory, x = hypothetical experiment.
Size = Group size (i.e., number of *S*s in each group); x = hypothetical experiment (i.e., no actual group size).
Cmplx = Complexity of group task; 1 = simple (e.g., simple information processing, assembly line, memory exercise), 2 = moderate complexity, 3 = complex (e.g., idea generation, creativity requirement, unfamiliar and complicated task), x = hypothetical experiment.
Time = Duration of the group task studied; 1 = less than one hour, 2 = a period of hours (within one day), 3 = over a period of days, 4 = weeks, 5 = months, 6 = years, x = hypothetical experiment.
A/O = Age and occupation; 1 = grade school children, 2 = college students, 3 = business employees, 4 = military personnel.
Sex = Sex of *S*s; 0 = all female, 1 = male and female, 2 = all male, 9 = not reported and difficult to infer.
Stats. Used = The statistics used to compute an effect size; raw = raw data in article, nsd = no significant differences reported.

& Carter, 1956a, 1956b) reported reliabilities for this variable. As a result, a reliability of 1 was assigned to all the other studies' measures of productivity. In many cases, this maximal reliability might have been obtained had the authors calculated one. For instance, Kipnis, Schmidt, Price, and Stitt (1981) measured group productivity by counting the number of objects produced by each group, a task that should result in high interrater reliabilities. In other cases, reliability would have been much lower, as when Meade (1967, 1985) assessed productivity on the basis of teachers' subjective ratings of the quality of children's art work. One might have given measures such as these speculative reliabilities, but there exists no acceptable procedure for doing so. Assigning values of 1 was considered the lesser evil.

By contrast, it was possible to construct an artifact distribution for questionnaire measures of satisfaction. The eight reported reliabilities ranged from .71 to .91 and produced an average alpha of .81 and an average interitem correlation of .53 (see Rosenthal, 1984). These figures are similar to those obtained for similar measures of satisfaction (e.g., Smith, Kendall, & Hulin, 1969, esp. chapter 4). For each study without an alpha but with a reported number of questionnaire items, the average interitem correlation and Rosenthal's (1984) reliability table (Table 3.10, p. 57) were used to compute a reliability. Following Rosenthal, "effective" reliability was calculated as $n\bar{r}/[1 + (n-1)\bar{r}]$, where n is the number of items, and $\bar{r}$ is the average interitem correlation. Those studies that reported neither reliabilities nor the number of items in their questionnaires were assigned the average alpha of .81.

The reliabilities obtained through these artifact distributions are reported in the fifth and sixth columns of Tables 1 and 2. Reliabilities that were reported in the original studies are italicized in these tables.

MODERATOR VARIABLES

Previous research reviews have posited that several variables moderate the effects of leadership style on productivity and satisfaction. Unfortunately, many hypothesized moderators and their

effects were not reported in a sufficient number of studies, which made it impossible to analyze them. This included geographic origin, cultural background, personality, and leadership expectations. Other moderators were imperfectly reported but worth including in the analysis. Although data were rarely broken down according to sex, researchers typically reported whether groups were all male, all female, or mixed. Many other variables could be placed into simple coding categories, including the measure of productivity (quantity or quality), forms of satisfaction measured (global, task, leadership, and group satisfaction), study setting (field, field experiment, and laboratory experiment), and group task complexity (simple, moderate, complex). Not surprisingly, the ages and occupations of subjects were closely related, so they were combined into a single age/occupation variable (i.e., grade school students, college students, military personnel, business employees). In addition, the operationalizations of leadership style were dichotomized into experimental and quasi-experimental, the latter consisting of field studies analyzing differences between preexisting leaders. Finally, task duration, group size, and publication date were included as possible moderators.[4] The moderator values assigned to each of the studies are reported in the right-hand columns of Tables 1 and 2.

ANALYSIS PROCEDURE

The analysis of the studies began by focusing on the different measures of productivity and satisfaction. It was necessary to determine whether the unit of analysis was individual studies (and multiple studies within a single article, as in Gibb [1951] and Vroom & Mann [1960]) or studies broken down by their different operationalizations. For instance, if quantity and quality measures of productivity appeared to constitute two distinct dependent variables, it would be necessary to analyze them separately. Otherwise, those studies that measured both quantities and qualities would be assigned a single r, the average of the z scores for the quantity and quality correlations (Rosenthal, 1984). Thus Tables 1 and 2 report

both overall *r*s and, when available, *r*s for different measures of productivity and satisfaction.

Once this issue was resolved, the meta-analysis proceeded according to the method outlined by Hunter and Schmidt (1990) and employed by Segrin (1990). Correlations were corrected for dichotomization, weighted by sample size, then summarized with a mean effect size and 95% confidence interval. The homogeneity of the distribution was tested using a chi-square and a comparison of expected and observed variances. A significant chi-square suggested that there was substantial variation among the studies. Variation was also presumed to be present if the ratio of expected to observed variance was less than 75%, meaning that less than three quarters of the observed variance could be attributed to sampling error.

When such variation was present in a distribution, an attempt was made to partition studies according to one or more moderator variables. These subgroups of studies were tested for homogeneity, and if found heterogeneous, split once again. Partitioning ceased when a distribution was homogeneous, the number of studies became too small to divide, and/or it was impossible to increase the homogeneity of distributions through further division.

This method was modified to acknowledge the criticism presented by Hall and Rosenthal (1991): "A nonsignificant overall heterogeneity test does not mean that no significant contrast can be extracted from the effect sizes" (p. 440). To address this problem, homogeneity tests were used as guidelines, rather than strict rules. When (a) prior theory and research made partitioning seemingly homogeneous studies plausible, (b) there was no overall loss in homogeneity, and (c) the effect sizes of the subgroups differed, groups of studies were divided, and subgroup features reported.

Finally, the file drawer statistic (Rosenthal, 1984) was calculated and reported for all groups and subgroups of studies. Typically, this statistic shows how many studies of average sample size and an effect size of zero would be necessary to reduce a substantively significant positive or negative average effect size down to the point where it enters a range of practical insignificance, that is,

$-.10 < \bar{r} < .10$.[5] If the file drawer statistic is low, it means that the results of the meta-analysis could be contradicted by pulling a few studies with nonsignificant findings from academia's ubiquitous file drawers of unpublished studies.

In this essay, the file drawer equation is also modified to calculate the number of studies with significant effect sizes it would take to create an $r = \pm.20$, what Cohen (1988) would call a "small to medium" effect size. This is not so much a file drawer statistic as a drawing board statistic: It shows how many studies with significant results would have to be on the drawing board for a future meta-analysis to detect a moderate average effect size when a current analysis shows a near-zero $\bar{r}$.[6]

RESULTS

PRODUCTIVITY

Despite their differences, the various measures of productivity were highly correlated. Five studies reported correlations between leadership and measures of both quantity-based and quality-based productivity. Correlating these correlations resulted in an r of .43, $p > .05$, $k = 5$. In addition, mean correlations between leadership and productivity were calculated separately for the different measures of productivity, and the results showed similar means and variances. Consequently, it was decided to treat the different measures of group productivity as operationalizations of the same dependent variable.

The overall distribution of the 23 studies measuring productivity is reported in Table 3. (Correlations corrected for attenuation are reported in the right-hand column of Tables 3 and 4. Analyses of these effect sizes resulted in subgroups identical to those found for correlations corrected only for dichotomization.) The chi-square and ratio of expected to observed variance suggested a homogeneous distribution.

Nevertheless, these statistics could have concealed a meaningful subgroup analysis, according to Hall and Rosenthal (1991). Thus

TABLE 3: Effect Sizes and Other Summary Data for Studies Measuring Productivity

Sets of Studies	k	$\bar{r}$	*95% Conf. Interval*	N	X2	*% Var. S. E.*	$\bar{r}_c$
Overall	23	–.02	.35 to –.31	1013	29	81	–.03
Setting of study							
Field study	7	.10	.35 to –.16	577	10	70	.11
Field experiment	5	–.19	–.04 to –.34	109	1	100	–.20
Lab experiment	11	–.04	.29 to –.37	327	9	100	–.04
Complexity of lab experiments							
Simple	4	–.10	.14 to –.34	239	4	100	–.11
Moderate complexity/ complex	7	.12	.43 to –.19	88	2	100	.16

NOTE: $p < .05$, k = number of studies, $\bar{r}$ = correlation corrected only for dichotomization, N = number of groups, % Var. S. E. = the percentage of variance accounted for by sampling error, $\bar{r}_c$ = correlation corrected for dichotomization and attenuation in independent and dependent variables.

the overall distribution was checked for moderator variables, a process that involved the examination of correlations, ANOVAs, and breakdowns of means. Given the limited theoretical clarity on the issue of moderators, all possible relationships among the effect sizes and the moderators were explored. The distribution of effect sizes could be broken down into more homogeneous piles by sorting according to two variables, the method of operationalizing leadership style and the setting of the study. Three analyses were conducted, sorting the distribution according to confederate versus naive leadership operationalization, study setting, and a combination thereof (akin to an interaction in ANOVA).

Comparison of these three analyses led to a decision to report only the second breakdown (see Table 3), which accounted for the most variance and produced the most distinguishable effect sizes. Field studies that examined the effects of preexisting leadership differences showed a small positive relationship between democratic leadership and productivity ($\bar{r}$ = .10), whereas field studies that involved quasi-experimental leadership manipulations resulted in a moderate negative average correlation ($\bar{r}$ = –.19). Lab experiments had an average correlation at the midpoint between the correlations of the other two settings ($\bar{r}$ = –.04).

As with the overall distribution, the subgroups appeared homogeneous—with the possible exception of the seven field studies. An attempt to partition these subgroups produced one possible breakdown, dividing the lab experiments into studies employing simple tasks and studies involving moderately or highly complex tasks (see Table 3). Simple lab tasks produced a small negative average correlation between democratic leadership and productivity ($\bar{r} = -.10$), and relatively complex lab tasks resulted in an equally small positive average correlation ($\bar{r} = .12$).

The file drawer and drawing board statistics for these distributions suggest that the findings are not entirely robust (see Rosenthal, 1984). For the overall distribution of 23 studies, it would take 18 additional studies with $N = 44$ and $\bar{r} = .20$ to change $\bar{r}$ from $-.02$ to $-.10$. However, less than 10 additional studies would significantly change the $\bar{r}$ for any of the subgroups of studies.

SATISFACTION

As Smith et al. (1969) have observed, measures of satisfaction are often highly correlated. This may have been the case with the set of studies examined herein, but the nature of the data made it difficult to obtain such correlations. Correlations within studies that measured more than one of the four forms of satisfaction were compared. None of these correlations involved more than eight studies, but there was one nearly significant correlation, between global satisfaction and satisfaction with leadership, $r = .61$, $p = .055$, $k = 8$. A variety of statistics was employed to explore ways in which the operationalizations might have related to effect sizes, but no clear relationships were apparent. Because partitioning the studies according to satisfaction operationalizations complicated analysis and revealed no clear distinctions between the studies, it was decided to average different measures of satisfaction within studies.

The first attempt at meta-analysis of this overall set of studies identified the existence of an outlier. Many psychometricians and meta-analysts (e.g., Hunter & Schmidt, 1990, p. 207) argue that almost all data sets contain extreme values that should be viewed as outliers and removed to ensure a meaningful analysis. Following

this principle, one of Vroom and Mann's (1960) studies was dropped due to its distance from all other studies ($r = -.39$).[7]

The overall pool of studies showed an average effect size of $\bar{r} = .23$ and a heterogeneous distribution (see Table 4). One possible explanation for the heterogeneity was explored initially. Thirteen studies used the group as the unit of analysis, which resulted in modest *N*s (these studies are shown with asterisks in Table 2). Because all studies needed to be based on the same unit of analysis, those studies using the group as the unit were included by multiplying their *N* by the number of people per group. This resulted in four relatively large weightings, which did not appear to skew the distribution of effect sizes.

As with productivity, moderator variables were examined with a battery of tests. Only two variables appeared vaguely related to effect sizes—the complexity and duration of group tasks. The overall distribution was broken down in three different ways: Studies were divided according to complexity, duration, and the interaction of complexity and duration.

The breakdown that most effectively reduced heterogeneity and differentiated effect sizes split studies according to their complexity. As shown in Table 4, studies with simple and moderately complex tasks resulted in a moderate correlation between democratic leadership and member satisfaction ($\bar{r} = .32$). Studies with highly complex group tasks averaged a near-zero correlation ($\bar{r} = .03$).

As with productivity, the overall average effect size was somewhat robust. It would take 36 studies with $N = 221$ and $\bar{r} = 0$ to reduce the overall $\bar{r}$ from .23 to .10. It would take 24 studies with $N = 383$ and $r = 0$ to reduce the $\bar{r}$ for simple/moderately complex tasks from .32 to .10, and 11 studies with $N = 121$ and $\bar{r} = .20$ would raise the high task complexity $\bar{r}$ from .03 to .10.

DISCUSSION

The results of this analysis support three empirical generalizations and point to possible subjects for future research. First, neither democratic nor autocratic leadership is, on average, more produc-

TABLE 4: Effect Sizes and Other Summary Data for Studies Measuring Satisfaction

Sets of Studies	k	$\bar{r}$	*95% Conf. Interval*	N	X2	*% Var S. E.*	$\bar{r}_c$
Overall	28	.23	.61 to –.15	6212	258*	11	.30
Complexity of task							
Simple/moderate complexity	11	.32	.53 to .10	4211	61*	18	.42
Complex	16	.03	.42 to –.35	1937	75*	22	.03

NOTE: * $p < .05$, k = number of studies, $\bar{r}$ = correlation corrected only for dichotomization, N = number of groups, % Var. S. E. = the percentage of variance accounted for by sampling error, $\bar{r}_c$ = correlation corrected for dichotomization and attenuation in independent and dependent variables. Recall that one study, using a hypothetical situation, could not be coded in terms of complexity.

tive; however, democratic leadership is most productive when it is a naturally occurring phenomenon, as opposed to a researcher's manipulation during a field or laboratory experiment. Moreover, experimentally imposed democratic leadership may be relatively effective (compared to autocratic leadership) when laboratory groups are given moderately or highly complex tasks. Second, the link between democratic leadership and satisfaction is neither large ($\bar{r} = .23$) nor uniform. The variance in r across studies is substantial, and it is difficult to organize the reported effect sizes into a coherent pattern. Third, the analyses of productivity and satisfaction do not support the common view that a large number of variables unambiguously and powerfully moderate the effects of democratic and autocratic leadership (e.g., Bass, 1990).

If this meta-analysis prompts future research, such investigation should consider the influence of task complexity and recognize the importance of study design and setting. Two decades ago, Bormann (1970) argued that small group research ignored the limitations of experimental manipulations at its peril. The artificiality of interventions, whether in the laboratory or the field, can distort findings, and it appears that this has been the case with studies of the productivity of democratic leadership. The dampening effect of experimental manipulation on the productivity of democratic leadership should not be surprising, because fully democratic leadership is implemented democratically rather than by an experimenter's random assignment to leadership conditions. Therefore,

future studies should look at preexisting differences in leadership style, and democratic leadership should be chosen by the membership, rather than imposed on the group.

This problem with democratic leadership research derives in part from a lack of conceptual precision. In Lewin et al.'s (1939) original study of democratic leadership (where the concept originated), they studied adult leadership of boys' groups and presumed that these groups were incapable of choosing their own methods of leadership. For this and other reasons, reviewers have long lamented the inadequacy of existing conceptualizations of democratic and autocratic leadership (e.g., Gibb, 1969; M. Lewin, 1987).

To address this issue, Gastil (in press) recently provided a definition of democratic leadership. Taking a functional approach, Gastil argues that democratic group leadership amounts to giving group members responsibility, improving the general abilities and leadership skills of other group members, and assisting the group in its decision-making process. Taking a functional approach, the definition suggests that many or all group members can perform important leadership functions. If used in research, this definition could bridge the chasm between democratic theory and empirical research on democratic leadership.

In conclusion, it is appropriate to return to this essay's opening theme, the pressing need for social scientific insight into the problems faced by democratic activists across the globe. From this perspective, the most noteworthy findings in this analysis are the apparent effectiveness of emergent democratic leadership and the limitations of externally imposed democratic methods. As nations and organizations around the world "experiment" with democratic decision making, they might remember this important lesson.

APPENDIX A
Studies Used in the Meta-Analysis

Adams, S. (1954). Social climate and productivity in small military groups. *American Sociological Review, 19*, 421-425.

Argyle, M., Gardner, G., & Cioffi, F. (1958). Supervisory methods related to productivity, absenteeism, and labor turnover. *Human Relations, 11*, 23-40.

Berkowitz, L. (1953). Sharing leadership in small, decision-making groups. *Journal of Abnormal and Social Psychology, 48*, 231-238.

Bevan, W., Albert, R. S., Loiseaux, P. R., Mayfield, P. N., & Wright, G. (1958). Jury behavior as a function of the prestige of the foreman and the nature of his leadership. *Journal of Public Law, 7*, 418-449.

Calvin, A. D., Hoffmann, F. K., & Edgar, E. L. (1957). The effect of intelligence and social atmosphere on group problem solving behavior. *Journal of Social Psychology, 45*, 61-74.

Day, R. C., & Hamblin, R. L. (1964). Some effects of close and punitive styles of supervision. *American Journal of Sociology, 69*, 499-510.

Foa, U. G. (1957). Relation of workers [sic] expectation to satisfaction with supervisor. *Personnel Psychology, 10*, 161-168.

Fox, W. M. (1957). Group reaction to two types of conference leadership. *Human Relations, 10*, 279-288.

Gibb, C. A. (1951). An experimental approach to the study of leadership. *Occupational Psychology, 25*, 233-248.

Haythorn, W., Couch, A., Haefner, D., Langham, P., & Carter, L. F. (1956a). The behavior of authoritarian and equalitarian personalities in groups. *Human Relations, 9*, 57-73.

Haythorn, W., Couch, A., Haefner, D., Langham, P., & Carter, L. F. (1956b). The effects of varying combinations of authoritarian and equalitarian leaders and followers. *Journal of Abnormal Social Psychology, 53*, 210-219.

Hendrix, W. H., & McNichols, C. W. (1982). Organizational effectiveness as a function of managerial style, situational environment, and effectiveness criterion. *Journal of Experimental Education, 52*, 145-151.

Kipnis, D., Schmidt, S., Price, K., & Stitt, C. (1981). Why do I like thee: Is it your performance or my orders? *Journal of Applied Psychology, 3*, 324-328.

Kushell, E., & Newton, R. (1986). Gender, leadership style, and subordinate satisfaction: An experiment. *Sex Roles, 14*, 203-208.

Lyle, J. (1961). Communication, group atmosphere, productivity, and morale in small task groups. *Human Relations, 14*, 369-379.

Mahoney, T. A. (1967). Managerial perceptions of organizational effectiveness. *Management Science, 14*, B76-B90.

McCurdy, H. G., & Eber, H. W. (1953). Democratic versus authoritarian: A further investigation of group problem-solving. *Journal of Personality, 22*, 258-269.

McCurdy, H. G., & Lambert, W. E. (1952). The efficiency of small human groups in the solution of problems requiring genuine co-operation. *Journal of Personality, 20*, 478-494.

Meade, R. D. (1967). An experimental study of leadership in India. *Journal of Social Psychology, 72*, 35-43.

Meade, R. D. (1985). Experimental studies of authoritarian and democratic leadership in four cultures: American, Indian, Chinese and Chinese-American. *High School Journal, 68*, 293-295.

Morse, N. C., & Reimer, E. (1956). The experimental change of a major organizational variable. *Journal of Abnormal and Social Psychology, 52*, 120-129.

Mullen, J. H. (1965). Differential leadership modes and productivity in a large organization. *Academy of Management Journal, 8*, 107-126.

Roberts, K., Miles, R., & Blankenship, L. V. (1968). Organizational leadership satisfaction and productivity: A comparative analysis. *Academy of Management Journal, 11*, 401-414.

Rosenbaum, L. L., & Rosenbaum, W. L. (1971). Morale and productivity consequences of group leadership style, stress, and type of task. *Journal of Applied Psychology, 55*, 343-348.

Scontrino, M. P. (1972). The effects of fulfilling and violating group members' expectations about leadership style. *Organizational Behavior and Human Performance, 8*, 118-138.

Shaw, M. E. (1955). A comparison of two types of leadership in various communication nets. *Journal of Abnormal and Social Psychology, 50*, 127-134.

Snadowsky, A. (1974). Member satisfaction in stable communication networks. *Sociometry, 37*, 38-53.

Stitt, C., Schmidt, S., & Price, K. (1983). Sex of leader, leader behavior, and subordinate satisfaction. *Sex Roles, 9*, 31-42.

Sudolsky, M., & Nathan, R. (1971). A replication in questionnaire form of an experiment by Lippitt, Lewin, and White concerning conditions of leadership and social climates in groups. *Cornell Journal of Social Relations, 4*, 188-196.

Thelen, H. A., & Withall, J. (1949). Three frames of reference: The description of climate. *Human Relations, 2*, 159-176.

Torrance, E. P. (1953). Methods of conducting critiques of group problem-solving performance. *Journal of Applied Psychology, 37*, 394-398.

Vroom, V. H., & Mann, F. C. (1960). Leader authoritarianism and employee attitudes. *Personnel Psychology, 13*, 125-140.

White, R., & Lippitt, R. (1960). *Autocracy and democracy*. New York: Harper & Brothers.

Wischmeier, R. R. (1955). Group-centered and leader-centered leadership: An experimental study. *Speech Monographs, 22*, 43-48.

Wispe, L. G. (1951). Evaluating section teaching methods in the introductory course. *Journal of Education Research, 45*, 161-186.

Wu-Tien, W., & Hsiu-Jung, C. (1978). Teacher leadership behavior as related to students' expectation, achievement, and adjustment. *Bulletin of Educational Psychology, 11*, 87-104.

Ziller, R. C. (1957). Four techniques of group decision making under uncertainty. *Journal of Applied Psychology, 41*, 384-388.

NOTES

1. The literature on democratic leadership has also produced an extended family of conceptual cousins, including Theory x versus Theory y management (McGregor, 1960), exploitative authoritative versus participative management (Likert, 1961, 1967), autocratic versus group decision making (Vroom & Yetton, 1973), conventional versus servant-leadership (Greenleaf, 1977), constituted versus nonconstituted leadership (Tucker, 1981),

transactional versus transformative leadership (Burns, 1978), power-over versus responsive leadership (Starhawk, 1986), and authority-based versus superleadership (Manz & Sims, 1989).

2. I would like to acknowledge the assistance of Arabel Huang, a colleague at the University of Wisconsin–Madison, who graciously translated essential details of Wu-Tien and Hsiu-Jung (1978).

3. The aforementioned studies in the psychology of sport are excluded due to the nature of their independent variables and reported data. These studies split the democratic/autocratic contrast into two separate behavioral dimensions. The individual effects of these scales could be combined, but the studies do not report these results (e.g., Chelladurai et al., 1988). Instead, the scales are typically entered into a regression equation with other independent variables (e.g., Dwyer & Fischer, 1990; Garland and Barry, 1990; Weiss & Friedrichs, 1986).

4. Publication date was used to code the date the study was conducted, unless more precise information was available. In the case of White and Lippitt (1960), for example, it was known that the original study was conducted in 1938-1939, so the 1938 date was used.

5. If a "small" effect size of .10 (Cohen, 1988) seems a low threshold for substantive significance, see Rosenthal's (1984) discussion of the potential impact of small effect sizes. As Rosenthal argues, if the dependent variable were simply success versus failure at a group task, an effect size of .10 is analogous to a 10% increase in a group's success rate.

The equation I use for the file drawer statistic is taken from Hunter and Schmidt (1990, p. 513): $x = k\,[(\bar{r}k/rt) - 1]$, where x is the file drawer statistic, k is the number of studies in the meta-analysis, rk is the average effect size of these studies, and rt is the threshold of the range of practical insignificance. For instance, if $\bar{r}k = .30$, the threshold would be $rt = .05$—the closest number in the range from .05 to –.05.

6. The drawing board statistic uses the same principle but a modified equation: $x = [k(\bar{r}k - rt)] / (rt - rs)$, where x is the drawing board statistic, rs is a significant effect size (for this analysis, rs is set at .20) and rt is the threshold of substantive significance, as opposed to insignificance. This is simply a transformation of the more intuitively appealing equation, $rt = [x(rs) + k(\bar{r}k)] / (x + k)$. For both the file drawer and drawing board statistics, x is always rounded up to the next whole number.

7. In retrospect, it is not surprising that this study is an outlier, because the researchers acknowledged that leaders previously identified as "authoritarian" frequently behaved in an egalitarian manner (Vroom & Mann, 1960, p. 132).

REFERENCES

Adams, S. (1952). *The American Psychologist*, *17*, 398.

Anderson, R. C. (1959). Learning in discussions: A resume of the authoritarian-democratic studies. *Harvard Educational Review*, *29*, 201-215.

Barber, B. (1984). *Strong democracy*. Berkeley, CA: University of California Press.

Barker, L. L., Wahlers, K. J., Watson, K. W., & Kibler, R. J. (1987). *Groups in process*. Englewood Cliffs, NJ: Prentice-Hall.

Bass, B. (1990). *Bass & Stogdill's handbook of leadership* (3rd ed.). New York: Free Press.

Beebe, S. A., & Masterson, J. T. (1990). *Communicating in small groups* (3rd ed.). Glenview, IL: Scott, Foresman.

Belwal, U. (1985). Study of the school climate in relation to administrative style. *Asian Journal of Psychology and Education, 15*(3), 13-16.

Bhushan, I. (1971). A study of leadership preference in relation to authoritarianism and intolerance of ambiguity. *Journal of the Indian Academy of Applied Psychology, 8*, 34-38.

Bhushan, R., & Verma, N. K. (1972). A study of social determinants of leadership preference. *Journal of the Indian Academy of Applied Psychology, 9*, 81-41.

Bormann, E. G. (1970). The paradox and promise of small group research. *Speech Monographs, 37*, 211-216.

Brilhart, J. K., & Galanes, G. J. (1989). *Effective group discussion*. Dubuque, IA: William C. Brown.

Brown, R. (1988). *Group processes*. New York: Basil Blackwell.

Burns, J. M. (1978). *Leadership*. New York: Harper & Row.

Busch, H. M. (1934). *Leadership in group work*. New York: Association.

Chelladurai, P., Imamura, H., Yamaguchi, Y., Oinuma, Y., & Miyauchi, T. (1988). Sport leadership in a cross-national setting: The case of Japanese and Canadian athletes. *Journal of Sport and Exercise Psychology, 10*, 374-389.

Chelladurai, P., & Saleh, S. D. (1978). Preferred leadership in sports. *Canadian Journal of Applied Sports Sciences, 3*, 85-92.

Chelladurai, P., & Saleh, S. D. (1980). Dimensions of leader behavior in sports: Development of a leadership scale. *Journal of Sport Psychology, 2*, 34-45.

Chernyshev, A. S., & Suryaninova, T. I. (1990). Genesis of the group subject of activity. *Soviet Journal of Psychology, 11*(2), 3-14.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cragan, J. F., & Wright, D. W. (1986). *Communication in small group discussions*. New York: West.

Dillard, J. P. (1991). *Retrieving effect sizes from ANOVA results*. Unpublished manuscript.

Dwyer, J. J., & Fischer, D. G. (1990). Wrestlers' perceptions of coaches' leadership as predictors of satisfaction with leadership. *Perceptual and Motor Skills, 71*, 511-517.

Eagly, A. H., & Johnson, B. T. (1990). Gender and leadership style: A meta-analysis. *Psychological Bulletin, 108*, 233-256.

Flanders, N. A. (1951). Personal-social anxiety as a factor in experimental learning situations. *Journal of Educational Research, 45*, 100-110.

Forsyth, D. R. (1990). *Group dynamics* (2nd ed.). Pacific Grove, CA: Brooks/Cole.

Fowler, G. D., & Rosenfeld, L. B. (1979). Sex differences and democratic leadership behavior. *Southern Speech Communication Journal, 45*, 69-78.

French, W. L., Kast, F. E., & Rosenzweig, J. E. (1985). *Understanding human behavior in organizations*. New York: Harper & Row.

Garland, D. J., & Barry, J. R. (1990). Personality and leader behavior in collegiate football: A multidimensional approach to performance. *Journal of Research in Personality, 24*, 355-370.

Gastil, J. (in press). A definition and illustration of democratic leadership. *Human Relations*.

George, C. S. (1972). *The history of management thought*. Englewood Cliffs, NJ: Prentice-Hall.

Gibb, C. A. (1951). An experimental approach to the study of leadership. *Occupational Psychology, 25*, 233-248.

Gibb, C. A. (1969). Leadership. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (pp. 205-282). Reading, MA: Addison-Wesley.

Greenleaf, R. K. (1977). *Servant leadership*. New York: Paulist.

Hackett, R. D., & Guion, R. M. (1985). A reevaluation of the absenteeism-job satisfaction relationship. *Organizational Behavior and Human Decision-Making Processes, 35*, 340-381.

Hale, J. L., & Dillard, J. P. (1991). The uses of meta-analysis: Making knowledge claims and setting research agendas. *Communication Monographs, 58*, 463-471.

Hall, J. A., & Rosenthal, R. (1991). Testing for moderator variables in meta-analysis: Issues and methods. *Communication Monographs, 58*, 437-448.

Haythorn, W., Couch, A., Haefner, D., Langham, P., & Carter, L. F. (1956a). The behavior of authoritarian and equalitarian personalities in groups. *Human Relations, 9*, 57-73.

Haythorn, W., Couch, A., Haefner, D., Langham, P., & Carter, L. F. (1956b). The effects of varying combinations of authoritarian and equalitarian leaders and followers. *Journal of Abnormal Social Psychology, 53*, 210-219.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.

Husband, R. L. (1988). Leading in organizational groups. In R. S. Cathcart & L. A. Samovar (Eds.), *Small group communication* (5th ed., pp. 493-507). Dubuque, IA: William C. Brown.

Isohanni, M. (1990). The organization of the ward treating psychotic patients. *Nordisk Psykiatrisk Tidsskrift, 44*, 261-267.

Jensen, A. D., & Chilberg, J. C. (1991). *Small group communication*. Belmont, CA: Wadsworth.

Johnson, D. W., & Johnson, F. P. (1991). *Joining together*. Englewood Cliffs, NJ: Prentice-Hall.

Kipnis, D., Schmidt, S., Price, K., & Stitt, C. (1981). Why do I like thee: Is it your performance or my orders? *Journal of Applied Psychology, 3*, 324-328.

Krips, H. (1992). Leadership and social competence in the declining years of communism. *Small Group Research, 23*, 130-145.

Lawler, E. E. (1984). Leadership in participative organizations. In J. G. Hunt, D. M. Hosking, C. A. Schriesheim, & R. Stewart (Eds.), *Leaders and managers* (pp. 316-332). New York: Pergamon.

Lewin, K., & Lippitt, R. (1938). An experimental approach to the study of autocracy and democracy: A preliminary note. *Sociometry, 1*, 292-300.

Lewin, K., Lippitt, R., & White, R. K. (1939). Patterns of aggressive behavior in experimentally created 'social climates.' *Journal of Social Psychology, 10*, 271-279.

Lewin, M. (1987). Kurt Lewin and the invisible bird on the flagpole: A reply to Graebner. *Journal of Social Issues, 43*, 123-139.

Ley, R. (1966). Labor turnover as a function of worker differences, work environment, and authoritarianism of foremen. *Journal of Applied Psychology, 50*, 497-500.

Likert, R. L. (1961). *New patterns of management*. New York: McGraw-Hill.

Likert, R. L. (1967). *The human organization*. New York: McGraw-Hill.

Loiseau, L. (1988). Styles de leadership et realite interiere: Analyse binaire classique des variations intrapsychiques [Leadership styles and internal reality: Classical binary analysis of intrapsychic variations]. *Revue de Psychologie Appliquee, 38*, 59-74.

Manz, C. C., & Sims, H. P. (1989). *Superleadership*. New York: Prentice-Hall.

McGregor, D. (1960). *The human side of enterprise*. New York: McGraw-Hill.

Meade, R. D. (1967). An experimental study of leadership in India. *Journal of Social Psychology, 72*, 35-43.

Meade, R. D. (1985). Experimental studies of authoritarian and democratic leadership in four cultures: American, Indian, Chinese and Chinese-American. *High School Journal, 68*, 293-295.

Miller, K. I., & Monge, P. R. (1986). Participation, satisfaction, and productivity: A meta-analytic review. *Academy of Management Journal, 29*, 727-753.

Nagel, J. H. (1987). *Participation*. Englewood Cliffs, NJ: Prentice-Hall.

Narain, I. (1964). Democratic decentralization and rural leadership in India: The Rajasthan experiment. *Asian Survey, 4*, 1013-1022.

O'Keefe, D. J. (1991). Extracting dependable generalizations from the persuasion effects literature: Some issues in meta-analytic reviews. *Communication Monographs, 58*, 472-481.

Pawlicka, I. (1987). Ocean funkcjonowania kierownikow przez podwladnych i produktywnosc grup zadaniowych w organizacji o strukturze dwuszczeblowej [Subordinates' evaluation of manager functioning and productivity of task groups in two-step structure organizations]. *Studia Psychologiczne, 25*, 41-51.

Quinn, R. E. (1984). Applying the competing values approach to leadership: Toward an integrative framework. In J. G. Hunt, D. M. Hosking, C. A. Schriesheim, & R. Stewart (Eds.), *Leaders and managers* (pp. 10-27). New York: Pergamon.

Rosenthal, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.

Schliesman, E. S. (1987). Relationship between the congruence of preferred and actual leader behavior and subordinate satisfaction with leadership. *Journal of Sport Behavior, 10*, 157-166.

Segrin, C. (1990). A meta-analytic review of social skill deficits in depression. *Communication Monographs, 57*, 292-308.

Serpa, S., Pataco, V., & Santos, F. (1991). Leadership patterns in handball international competition. *International Journal of Sport Psychology, 22*, 78-89.

Smith, P. C., Kendall, L. M., & Hulin, C. L. (1969). *The measurement of satisfaction in work and retirement*. Chicago, IL: Rand McNally.

Smith, T. V. (1926). *The democratic way of life*. Chicago: University of Chicago Press.

Srinivasan, P. T., & Kamalanabhan, T. J. (1986). Relationship of leadership styles to job involvement: An empirical study. *Journal of the Indian Academy of Applied Psychology, 12*, 53-58.

Starhawk. (1986). *Truth or dare*. New York: Harper & Row.

Stogdill, R. H. (1974). *Handbook of leadership*. New York: Free Press.

Sudolsky, M., & Nathan, R. (1971). A replication in questionnaire form of an experiment by Lippitt, Lewin, and White concerning conditions of leadership and social climates in groups. *Cornell Journal of Social Relations, 4*, 188-196.

Tead, O. (1935). *The art of leadership*. New York: McGraw-Hill.

Tubbs, S. L. (1988). *A systems approach to small group interaction* (3rd ed.). New York: Random House.

Tucker, R. (1981). *Politics as leadership*. Columbia: University of Missouri Press.

Vroom, V. H., & Mann, F. C. (1960). Leader authoritarianism and employee attitudes. *Personnel Psychology, 13*, 125-140.

Vroom, V. H., & Yetton, P. W. (1973). *Leadership and decision making*. Pittsburgh, PA: University of Pittsburgh Press.

Vujtech, J. (1972). Problem of style of management of a work group. *Synteza*, *5*, 58-66.

Weiss, M. R., & Friedrichs, W. D. (1986). The influence of leader behaviors, coach attributes, and institutional variables on performance and satisfaction of collegiate basketball teams. *Journal of Sport Psychology*, *8*, 332-346.

White, R. K., & Lippitt, R. (1960). *Autocracy and democracy*. New York: Harper & Bros.

Whitehead, T. N. (1936). *Leadership in a free society*. Cambridge, MA: Harvard University Press.

Wilson, G. L., & Hanna, M. S. (1986). *Groups in context*. New York: Random House.

Wu-Tien, W., & Hsiu-Jung, C. (1978). Teacher leadership behavior as related to students' expectation, achievement, and adjustment. *Bulletin of Educational Psychology*, *11*, 87-104.

Yukl, G. A., & Kanuk, L. (1979). Leadership behavior and effectiveness of beauty salon managers. *Personnel Psychology*, *32*, 663-675.

John Gastil is a doctoral candidate in communication arts at the University of Wisconsin–Madison. He received his B.A. in political science from Swarthmore College in 1989 and his M.A. in communication arts from the University of Wisconsin–Madison in 1991. He recently authored Democracy in Small Groups *(New Society Publishers, 1993).*