Of Babies and Bathwater

An Extension of the Business & Society Research Forum on the Fortune Reputation Database

EUGENE SZWAJKOWSKI
University of Illinois, Chicago
RAYMOND E. FIGLEWICZ
University of Michigan, Dearborn

A research forum published in *Business & Society* in 1995 (Issue 2) analyzed whether *Fortune* magazine's annual Reputation Survey (FRS) is viable as a corporate social performance (CSP) research database. We examine plausible alternative interpretations for a number of assertions and conclusions by the forum authors, including the premise for Brown and Perry's proposed transformation: that the *Fortune* data are confounded by the presence of a financial "halo," which biases ratings of nonfinancial attributes. Finally, we examine the appropriate roles of the two primary corporate rating systems in the context of CSP.

In 1995, *Business & Society* (Issue 2) undertook a laudable project: a research forum of five articles (Baucus, 1995; Brown and Perry, 1995a, 1995b; Logsdon and Wartick, 1995; Sodeman, 1995) with an introduction by editor Donna Wood (1995). Each of the studies focuses on either the *Fortune* reputation database in general (see Fisher, 1996) or its "community and environmental responsibility" (CER) component measure in particular, or both, and on their appropriate roles in the analysis of corporate social performance (CSP). The major stimulus for the project was Brown and Perry's (1994) proposal for adjusting CER values by statistically removing the contamination of an alleged financial "halo." Our article is intended to be an extension of the forum's evaluations of both the original and transformed databases. We also assess the place of such databases in the context of the CSP literature.

Although we agree with many of the points made in the *B&S* forum, we also believe that there are some questionable assertions among the critiques. Put another way, in their zealous, and admirable, efforts to identify possible sources of Type I research error in *Fortune* magazine's annual Reputation Survey (FRS) studies (drawing conclusions from invalid data), the forum authors have created a vulnerability to Type II error (rejecting conclusions derived from valid data, or "throwing the baby out with the bathwater").

Because of these problems, we contend that the viability of the *Fortune* reputation survey (hereafter designated as FRS or "the survey") should still be considered an open question. By our reckoning, FRS is not "terminally ill" (Baucus, 1995: 227), even as a source of CSP data, at least not yet. Its limitations do not justify total abandonment, "provided that researchers know what they are using, how to use it, and how to carefully interpret their findings and link results back to theory" (Wood, 1995: 198).

The primary aims of this article, then, are to draw attention to relevant evidence that was undiscovered, misinterpreted, or undervalued by the forum authors, and to examine plausible alternative arguments to those put forth by the authors. First, though, we address the value of the research forum in general.

THE RESEARCH FORUM

The *B&S* forum makes a significant contribution to the literature in a number of ways. First, it brings together targeted analyses by some of the best-known scholars in social issues research. Such articles would ordinarily be scattered across several journals or not published at all because of their degree of specialization. Second, true to their authors' credentials, each article conveys valuable insights on proper investigative methods in the related areas of reputation and CSP. Finally, taken together, the forum provides an object lesson in the complexities and demands imposed by scientific rigor in the context of an important real-world topic. This is particularly valuable to less experienced scholars and, of course, more so to those contemplating corporate reputation or CSP as an area of research focus.

However, the same concentrated attention the forum gives to its authors' criticisms also allows us a better opportunity to appreciate whatever limitations those criticisms contain. Of course, that is a healthy outcome, too. It is why we peer-review and publish our work. Indeed, the forum's authors themselves critique each other's comments. This article

is offered in the same spirit, but it is not just a "me, too" effort. In the course of reading the *B&S* studies, we found, along with the cogent scholarship mentioned above, some serious questions of logic, methodological validity, and interpretation. We submit our evaluations of these issues in the following pages, beginning with the forum's premise.

PREMISE: THE FINANCIAL HALO

The impetus for Brown and Perry's (1994) article, and subsequently for the forum, is the contention that there is a "financial halo" in the *Fortune* data. That is, when the survey's respondents are ostensibly rating a firm's innovativeness or quality of management or CER, they are unduly influenced by the company's financial performance. Empirical support for this position is based on two replicated research findings. First, when factor analyzed, the survey's eight attributes consistently load on a single factor accounting for an overwhelmingly large (more than 80%) proportion of variance. Second, commonly used indicators of financial performance correlate significantly with the FRS reputation measures. We examine each of these findings in the context of the relevant literatures.

Single-Factor Structure

Fombrun and Shanley (1990) were the first to discover the single-factor phenomenon, which was later confirmed by Fryxell and Wang (1991, 1994) using data from different years. As evidence of a halo effect, a single-factor structure is, at first blush, quite convincing. Brown and Perry (1994: 1349) argue that a "halo is present whenever principal components analysis results in 'a common general factor showing high loadings on nearly all attributes which accounts for appreciable variance' " (Dillon, Mulani, and Frederick, 1984: 194). In most cases, we would concur, but not here.

Szwajkowski and Figlewicz (1995) contend that this one-factor finding is spurious because factor analysis is being inappropriately applied to the data. Specifically, it is being applied to the wrong level of the data. Although there are eight attributes and hundreds of companies in the sample, these data are not suited to examining rater bias or financial halo because they have already been aggregated, with only the central tendencies for each company/attribute remaining from the information contained in the original survey responses. Statistically, a factor analysis of means is not the same as, and cannot be assumed to reflect, a factor analysis of their underlying observations. To make such an assumption is to engage

in "ecological inference" (also called "the ecological fallacy" in the literature on research methods: see King, 1997; Robinson, 1950, 1951; Smith, 1981). Ecological inference is the phenomenon of drawing individual-level conclusions from aggregated data.

The problem of ecological inference is not pervasive in the organizational literature, where most analyses are conducted on unaggregated data (that is, the degrees of freedom are associated with the lowest level of analysis). However, a number of disciplines, such as education, epidemiology, geography, political science, and sociology, have confronted this issue for decades. Gary King (1997), a political scientist at Harvard, puts it thus:

> The ecological inference problem has been among the longest standing, most actively pursued, and consequential in quantitative social science. It was originally raised over 75 years ago as the first statistical problem in the nascent discipline of political science, and it has held back research agendas in most empirical subfields of the discipline. Almost all researchers who use aggregate data have encountered some form of the ecological inference problem. (p. 1)

Virtually every text on the subject explicitly states that factor analysis is distinctly a function of the correlation matrix. Therefore, the factor structure of means can accurately mirror the corresponding rating data only if the respective correlation matrices are equal. There is no statistical reason to expect this to be true. Put another way, if we can demonstrate that a single-factor structure of means can result from even unbiased or inversely correlated ("negative-halo") data, then the interpretation of a positive-correlation halo in the underlying ratings becomes just one possibility of many. Let us, then, construct a simulated, but extreme and artificial, sample of ratings (see Tables 1a and 1b).

The values in the data columns of Panel A represent the hypothetical observations (survey responses) underlying three attribute ratings for three out of n companies in the distribution. Assume that the ratings for all of the firms in the database follow a similar configuration. Then, for each firm, the responses for Dimension A are perfectly but inversely correlated with B. Variable C is completely random with respect to the others, and therefore uncorrelated with them. However, both the means and the standard deviations of all three attributes in each company subsample are identical. Even though the means vary from firm to firm, their intracompany values are identical. Thus, the correlation matrix of the means (bottom, right of Panel B) has all positive values of 1.0. A factor analysis on this matrix, such as those performed by Fombrun and Shanley

Table 1a
Example of Ecological Inference as Ecological Fallacy:
Comparison of Sample Distributions

	Respondent Ratings								
	Attribute (Company 1)			Attribute (Company i)			Attribute (Company n)		
Respondent	A	B	C	A	B	C	A	B	C
1	2	10	10	1	9	9	0	8	8
2	2	10	4	1	9	3	0	8	2
3	2	10	10	1	9	9	0	8	8
4	2	10	2	1	9	1	0	8	0
5	3	9	4	2	8	3	1	7	2
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
32	9	3	7	8	2	6	7	1	5
33	10	2	9	9	1	8	8	0	7
34	10	2	3	9	1	2	8	0	1
35	10	2	8	9	1	7	8	0	6
36	10	2	3	9	1	2	8	0	1
Mean	6.00	6.00	6.00	5.00	5.00	5.00	4.00	4.00	4.00
SD	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58

Table 1b
Example of Ecological Inference as Ecological Fallacy: Zero-Order Correlations

	Respondent Ratings											
	Company 1^a			Company i^a			Company n^a			Rating Means[b]		
Attribute	A	B	C	A	B	C	A	B	C	A	B	C
A	1.00			1.00			1.00			1.00		
B	−1.00	1.00		−1.00	1.00		−1.00	1.00		1.00	1.00	
C	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00

a. The sample size for this matrix is the total number of respondents rating this company.
b. The sample size for this matrix is the total number of companies rated.

(1990) and Fryxell and Wang (1991, 1994), would result in a single factor, with all loadings positive. Factor analyzing the ratings-level matrix, however, produces a two-factor structure with A and B having opposite loadings on one factor, and C being a unique factor. Clearly, the factor structure of the means gives a distorted picture of the individual ratings.

Obviously, actual FRS attribute means do not correlate perfectly, but they are quite close to this extreme. For example, the zero-order correlations of the eight dimensions in the 1996 survey (Fisher, 1996) range from

0.66 to 0.95. Excluding the CER attribute from the matrix yields r values that all exceed 0.70, and 86% of them are above 0.80. These are extremely high correlations for such a large sample ($N = 417$ companies).

In general, the greater the within-attribute rating variance, given a fixed and closed-ended scale (0-10 in FRS), the more the means of those attributes will tend to regress toward the scale midpoint because a high or low mean would require greater agreement. If the mean values across the attribute set (within company) regress toward the midpoint, they would also tend to be similar to each other. Replicate this scenario over a substantial majority of the firms in the sample and the resulting mean data would be highly intercorrelated, with a probable single-factor structure, even if the raw data are essentially uncorrelated and free of any halo. Thus, the single-factor-with-all-positive-loadings phenomenon can derive from a correlation matrix of ratings that could have any configuration whatsoever, one of them—but only one—being a pervasive halo (financial or otherwise). The converse, however, is not true; namely, that all configurations of raw data (attribute categories, reliability levels, central tendencies, and distributions) will necessarily produce a single-factor structure of means, unless the means themselves are highly and positively correlated, as in FRS.

In addition, there is new empirical evidence that the single-factor findings do not reflect FRS data at the rater level. The aggregated data, overall and for each attribute, are now available on *Fortune*'s World Wide Web site ("America's Most Admired Corporations," 1997). For the first time, though, means also are given separately for the executive (presumably including directors) and analyst subsamples. We factor-analyzed these 16 variables (8 attributes × 2 rater types) and obtained an intriguing result. As in earlier studies, all eight attributes loaded positively and very heavily on a single factor, but each subset of eight loaded on a separate factor.

This is strong evidence that an ecological fallacy was present in the earlier factor structure research because there must be higher inter-attribute correlation within rater type than intra-attribute correlation across rater type to produce our findings. That conclusion is confirmed by paired (by rater type) intra-attribute t tests, in which we found that rating means on five of the eight attribute pairs differed significantly at $p < .01$ or better. On two of these (quality of products and services, innovativeness), the analysts tended to rate firms higher than executives did, whereas the opposite was true for the other three (value as long-term investment, community and environmental responsibility, use of corporate assets). Neither our factor analysis nor our t-test results would have occurred if

there were as much interrater agreement as would be necessary if the single-factor structure had been an accurate reflection of rater data.

Perhaps the most convincing evidence of statistical artifact is that an analysts-versus-executives factor analysis cannot even be conducted at the rater level. This is, of course, because each rater can be either an analyst or an executive but cannot be both. Thus, no observation or record in the database can have all 16 variables necessary to create the complete correlation matrix, and the aggregate-level factor analysis cannot reflect reality at the rater level.

Conclusion: We must reject the factor analysis arguments because of their dependence on ecological inference and because of the new evidence cited in the preceding paragraphs. The method as executed is statistically invalid, and the interpretation of the results is misleading because no single definitive conclusion can be drawn from those tests. Factor analysis is a viable method for detecting halo effects, but it must be conducted at the appropriate measurement level.

Intercorrelation With Financial Performance

The other justification for the halo premise is that the literature has frequently reported a high degree of correlation between financial performance indicators and social components of reputation. If nonfinancial attributes such as innovativeness and CER are strongly correlated with common measures of financial performance, does that signal the existence of a halo? The answer is "yes and no." There is certainly the suggestion of a halo, but as Baucus (1995) points out, citing Murphy, Jako, and Anhalt (1993), the real issue is whether it is a *true* halo or an *illusory* one. Brown and Perry (1995a) assume that it is wholly illusory, or error, despite the fact that they acknowledge "there may well be true association between the constructs" (p. 203).

To accept the idea that such a halo is entirely error, one must believe either that there is no material relationship between financial performance and social performance to the degree that the latter is measured by the FRS attributes or that the survey attributes do not accurately measure social performance and thus cannot reflect a true halo even if one exists. There is no conclusive direct evidence to support or refute either belief, but indirect and circumstantial clues weigh against both.

The concept of halo as error dates back to the work of Thorndike (1920), who described it as the influence that an overall impression exerts on a rater's evaluation of specific attributes. Asch's (1946) "warm-cold"

experiments showed how a single characteristic can create an impression so strong as to influence a rater's evaluation of other attributes. Given these scenarios, two elements must be present for an illusory financial halo to exist. First, most of the executives, outside directors, and analysts who are the FRS respondents must substantially believe that there is a strong positive correlation between financial performance and all of the nonfinancial attributes in the survey. Second, and more important, that belief must be *in error* in order for the halo to be illusory. There is ample behavioral evidence to suggest that a true link exists between social performance and financial performance (SP/FP). Firms such as Union Carbide (workplace safety at Bhopal, India), Manville (workplace and product safety regarding asbestos), Exxon (*Valdez* oil spill), Beech-Nut (product quality regarding adulteration of apple juice), and Archer-Daniels-Midland (price-fixing) all experienced a financial impact from negative social performance events and practices. Johnson & Johnson overcame the economic damage that resulted from the Tylenol tampering scare by its adept and thorough implementation of corrective measures.

One of the strongest arguments for undertaking socially responsive actions is "enlightened self-interest," that is, the likelihood of a material payoff for doing good and the threat of negative consequences from socially harmful behavior. Enlightened self-interest is a sham if a financial halo is illusory. Frooman's (1997) meta-analysis of 27 event studies, however, convincingly argues that irresponsible or illegal behavior has significant , and negative, financial (market-value) effects.

Numerous studies of the relationship between financial performance and the overall reputation ratings, including some commissioned by *Fortune* itself and those cited by Brown and Perry (1994: 1348), have consistently concluded that financial measures account for only about half of the variance in reputation ratings. According to Fisher (1996), "reputation entails much more than just minting money. As measured in the survey, half of it comes from intangibles like the way a company treats its employees, how much it spends on research and development, and the strength of its management team" (p. 91).

Reason dictates the need for caution in claiming that positive SP/FP relationships are true. Few would argue against the notion that better employees make better and more innovative products, or that better products increase profitability and market value. Yet, the idea that community involvement and environmental responsibility are financially beneficial to the firm remains controversial. Indeed, business spokespersons frequently argue that such investments hinder competitiveness. We believe it is only logical to conclude that *some* social behaviors positively

impact financial performance as a general rule, others relate negatively, and still others are unrelated.

Conclusion: We must reject the argument that intercorrelation with financial performance is conclusive proof of an *illusory* financial halo. There is no clear evidence that whatever halo might exist in the FRS data either is entirely illusory or emanates only from financial performance. A plausible alternative interpretation, one that is perhaps even more credible, is that both financial and social performance are driven by effective organization, which is in turn determined by management values and practices and how they fit with the demands of the market and other stakeholders. Data transformation is unnecessary if the halo is not error, and partialling out financial performance is inappropriate for even an illusory halo if that halo is not known to be solely financial.

THE FORTUNE SURVEY AS A CSP DATABASE

Much of the *B&S* forum centers on the halo issues discussed above, but many of its authors' criticisms of the FRS database are independent of the halo premise. These deal with such concerns as sampling artifacts, construct validity, and reliability. As with the premise, these criticisms bring to light important points, but we question a number of the subsequent interpretations. The following major criticisms advanced by the forum authors will be discussed in turn: (a) representativeness of the respondent sample, (b) representativeness of the company sample, and (c) lack of attribute clarity.

Representativeness of the Respondent Sample

Virtually throughout the *B&S* forum, one can find claims that the kinds of respondents *Fortune* chooses to survey are subject to biases that call into question the validity of their ratings. Wood (1995), for example, detects the likelihood of a predisposition toward self-interest:

> It seems to me that the nature of the input—reputational ratings by CEOs, other top executives, and financial analysts—is itself a fatal blow to using the measure in any objective sense. It seems the same as asking the foxes how well they take care of things down at the henhouse. (p. 198)

This is a generalization that has a ring of truth to it, but on further examination, there are plausible alternative arguments. Admittedly, the

innate human need for self-esteem suggests that raters would tend to overvalue the attributes of their own firm (and probably industry). All of the respondent types figure to be subject to such an industry bias, but only the chief executive officers and outside directors identify with a specific company. More important to the criticism, though, is the fact that self-esteem also would influence the same two groups to *underrate* their intraindustry competitors, who typically outnumber their own firm in the sample by a ratio of nine to one. Thus, there might be a counterbalancing of biases, but we cannot say at this point whether they cancel each other out or whether the net effect is undervaluation or overvaluation.

Perhaps the most salient issue here is that the absolute magnitude of any of the ratings, as far as we know, has never been an important question to researchers, managers, or any stakeholders. For example, in the 1995 survey, Rubbermaid received an aggregate reputation rating of 8.65, with attribute values ranging from 8.22 to 9.13. Given the 0-10 scale, does that mean that the company's reputation is within 13.5% of being perfect, or at least the best it can be? Of course not. What is meaningful to people, including researchers, can be inferred from the descriptive statistics that *Fortune* reports, and those treat *relative* rather than *absolute* standing, especially within industry. Indeed, for the cited year, *Fortune* devoted seven pages to intraindustry comparisons (Jacob, 1995). Our own primary research concern regarding self-interest bias is that it could vary in strength across industries, signaling the advisability of controls for industry type. However, such controls are already recognized as important to reputation research. For example, a reasonable approach would be to use industry-matched, paired samples of firms, such as in Szwajkowski and Figlewicz (1995).

A second type of respondent bias attributed to the FRS sample is unrepresentativeness, specifically, a perspective that takes into account only business (read "financial") concerns. The following passage is typical of this argument:

> *Fortune* uses a large sample of industry stock analysts, senior executives, and outside directors in its annual survey, so the survey results are representative of the views of the corporate business community only. The perceptions and judgments of consumers, employees, middle management, labor union leaders, suppliers, community leaders, environmental activists, government officials, the media, academics, and any of a host of other stakeholder groups are simply not represented in the *Fortune* survey. (Logsdon and Wartick, 1995: 223)

We question the claim that the work roles of these respondents focus solely on finance, even when one accepts the Friedmanesque concept that they are ultimately answerable only to shareholders (Lee and McKenzie, 1994; Schlossberger, 1994; Van Buren, 1995). The responsibility of each category of respondent—executive, director, analyst—is to recognize interests of every stakeholder group that may have material importance to the market value, and hence to the shareholders, of the companies they are rating.

There is no question that nonfinancial issues can have a marked impact on market value. The findings of more than a decade of event studies on product recalls, executive succession, lawsuits, union representation, corporate name changes, and the like, bear this out. Furthermore, the examples of Union Carbide, Exxon, Manville, Archer-Daniels-Midland, and Johnson & Johnson, cited earlier, are also relevant here. Both the reputations and finances of these firms were significantly affected by primarily nonfinancial events or conditions.

Virtually every business school curriculum has a core capstone course, usually called Strategic Management, which requires its students to draw on their full range of functional training and consider the interests of all stakeholder sectors. The clear consensus is that organizations can succeed only with such a broad managerial perspective. It follows that what is true for companies must also be true for executives, directors, and analysts.

Conclusion: Respondent biases might indeed exist, but they should not be assumed, nor assumed away, without direct evidence. The evidence that has been compiled to date is not pertinent to rater bias (factor structure studies) or could occur without rater bias (financial performance intercorrelation). Furthermore, we have presented counterarguments to the speculation that financial performance bias is a natural expectation of the raters' functional positions. Failure to recognize nonfinancial stakeholder influences would make executives, directors, or analysts poor stewards of their fiduciary responsibilities.

Representativeness of the Company Sample

"It is a grave error to assume that *Fortune* ever intended to advance the science of management or the field of business and society" (Sodeman, 1995: 217). Although the word "grave" in this quote might be a bit extreme, Sodeman's point is well taken. One should not use FRS for reputation or CSP investigation without a thorough understanding of the

specific limitations it has relative to the research questions at hand. Indeed, the factor structure studies analyzed above suffer from just such a problem.

At the same time, this advice is applicable to all empirical examinations that employ secondary data sources. We know that all methodologies are flawed. Archival inquiries have the advantage of minimizing reactivity (influence of researcher biases on data collection), but they are subject to the lack of control over the availability and configuration of pertinent information. Clearly, CSP scholars would design the survey differently if given the opportunity. Unfortunately, this is not an option. If we refused to work with or to accept as useful anything but perfect research designs, there would be no literature at all.

Because it is an imperfect world, investigators must assess the degree to which credible information that addresses their research questions can be found in a database and then proceed from that point. To their credit, the forum authors take on this task to the benefit of all. We now assume the role of evaluating their efforts.

Baucus (1995: 228-9), in particular, questions the appropriateness of using FRS, based primarily on inconsistencies in the company sample over time. The author notes marked growth in both the number of companies evaluated (now more than double the original sample) and the number of respondents (now nearly doubled). At the same time, the response rate dropped somewhat from roughly half to 43.6% in 1988, after which this statistic was no longer reported. Baucus also points to attrition effects over time from mergers, bankruptcies, and the modification of *Fortune*'s criteria for configuring industries, as well as possible changes in the reported values of the respondents. All of these she cites as reasons for abandoning the FRS rankings, which we must let "die with dignity" (Baucus, 1995: 230).

All of Baucus's observations are accurate, but we question the severity of the conclusion. In essence, she would have us apply the control requirements of true experimental design (Campbell and Stanley, 1963) to archival analysis. Indeed, there is considerable correspondence between the flaws she cites and those in Campbell and Stanley's list of artifacts (p. 5) affecting internal validity of experiments (for example, history, maturation, instrumentation, and mortality). As an archival researcher herself (see Baucus and Near, 1991), Baucus has had to accept the fact that these effects cannot be totally controlled outside the laboratory, especially with real (and large) organizations as the subjects, as they are in FRS.

Furthermore, most of her criticisms focus on factors that are likely to be internal validity concerns only in longitudinal studies. In a cross-sectional analysis, they would limit only generalizability to other years.

Yet, regardless of the application, these constraints, we argue, are not fatal to the use of FRS. In fact, many of the same characteristics are present, even inescapable, in other widely used databases: for example, COMPU-STAT (financial statement data), and CRSP (share price data). They are shortcomings resulting from trade-offs (control versus realism) that are accepted as unavoidable in macrolevel research. It is certainly worthwhile to remind ourselves periodically of these limitations, but we should not be paralyzed by them. Instead, we give them their due, recognizing them as the potential sources of error that they are and conceptually evaluating the likelihood that they have affected our results.

Some partial solutions are available for a variety of these problems. For example, one can use multiple research designs or statistical methods that have complementary strengths and weaknesses; this allows for the detection of convergent and discriminant validity. One can control for artifact effects, such as by using event history analysis to minimize the consequences of data attrition. Finally, one can employ listwise deletion, analyzing only those observations with complete data. Inevitably, though, any corrective measures undertaken will carry with them their own baggage of shortcomings (such as loss of information), which also must be examined.

Conclusion: There is no doubt that FRS data must be used with the utmost of caution because virtually every aspect of them is volatile from year to year. This, however, is not cause for abandoning the database entirely; all large-sample studies of real-world organizations face most of the same problems. The solution is to incorporate controls where possible, evaluate the extent to which each shortcoming might systematically bias findings, and otherwise acknowledge all such potential sources of error, putting forth one's best judgment as to their impact.

Lack of Attribute Clarity

There is no attempt in FRS to define the meanings of the eight component attributes. As Sodeman (1995) puts it in the following quote (about the term "environment" in the CER attribute),

> We cannot assume that the editors of *Fortune*, the respondents of this survey, and other researchers who use this data set agree on what this term represents. (p. 218)

Actually, it would be rather safe to say that there would be a great deal of *dis*agreement among these groups about the meanings of each of the

FRS attributes, including those that are directly financial. Again, the crucial methodological concern is how serious an obstacle this creates for sound empirical research. Sodeman implicitly acknowledges the concern and responds as follows:

> Letting the subject express his or her own concept of what a criterion means is an acceptable research technique, but the researcher must also attempt to gain some understanding of what the subject believes those meanings are. (Sodeman, 1995: 217)

We would argue that "some understanding" about meanings does not require, as Sodeman claims, "a clear, unambiguous, and robust definition of the criterion" (p. 217). There is ample precedent for such a conclusion in the literature on the market value of the firm (shareholder valuation). Today, virtually every study that evaluates a company's total worth or value (sometimes labeled as overall effectiveness) employs as a primary measure some variant of the stock market return (price appreciation plus dividends). It is often described as the best indicator, and sometimes as the only true indicator, of effectiveness (see Szwajkowski and Figlewicz, 1995; Wokutch and McKinney, 1991). In essence, the argument is that a firm's value is whatever the market is willing to pay, or perhaps more accurately, is committed to pay.

Market returns operate in a manner that strongly parallels that of reputation in the *Fortune* survey. First, shareholders are effectively rating a firm's value through their investment choices. Yet, they cannot ask for, or receive, an official list of criteria on which to base such choices. Consequently, it is reasonable to assume, and perhaps unreasonable to deny, that individual shareholders buy and sell stocks at specific market prices for a wide variety of reasons. That makes market value just as perceptual as reputation as an operationalization of organizational effectiveness.

A collection of trained professionals, such as the FRS respondent sample, is likely to exhibit a higher level of agreement on criterion meaning than is an unorganized combination of individual and institutional investors or a disparate grouping of single-issue social performance advocates. The FRS raters are essentially key informants, acting as surrogates for the stakeholders, and, given the description, should be as well qualified as anyone to do so, as Brown and Perry (1994) acknowledge. They represent the market for reputation, much as shareholders represent the market for company valuation, and as acquiring and target firms represent the market for corporate control. In each of these markets, the

players all have their own assessments of what constitutes the firm's worth. The reputation market as defined by FRS differs from the reputation market in the real world in that it is more, not less, specific regarding attributes. That is, *Fortune* assumes that overall reputation is the unweighted arithmetic sum of the eight attributes. If this is true, then the combined score should correlate well with shareholder valuation over time. Szwajkowski and Figlewicz (1995) found strong evidence of just that kind of correspondence, both longitudinally and cross-sectionally.

Viewing CSP as an inherently perceptual assessment by external stakeholders is subject to criticism derived from the "perception versus behavior" controversy that is specifically addressed in the *B&S* forum. We will consider the issues relevant to that debate later in this article. For now, let us accept the perceptual definition, as a number of CSP scholars already do, as demonstrated by the following quote:

> CSP is a perception about the behavior of the corporation under consideration by those external observers able to influence outcomes associated with or perceptions about the firm, or influence the environment in which the firm operates. (Waddock and Mahon, 1991: 233)

In this definition, Waddock and Mahon suggest that CSP is less about what a firm does in the social arena (its behavior) than it is about the reputation it gains regarding what it does. It is in that sense that we tend to use the terms reputation and CSP interchangeably. Sodeman implies that an understanding of the criteria for attributes such as those in FRS must be developed a priori, as is done, for example, by Kinder, Lydenberg, and Domini & Co., Inc. in their SOCRATES social investment database.[1] That is not necessarily true. An alternative that is no less valid would be to determine the ways that social behaviors shape respondent-defined attribute ratings through empirical analysis. For example, if plant closings and layoffs significantly influence a firm's reputation for responsibility to its community, then the occurrence of such events, and perhaps their magnitudes, should be related to the perceptual rating. This is not a new idea; indeed, substituting "market value" for the dependent variable (reputation) would delineate a typical research question for a classic event study of shareholder valuation.

One scientific disadvantage to specifying social performance criteria a priori is, of course, that it probably will create excessively narrow descriptions of the attributes in question. Another weakness is that it frames the research in terms of the investigator's value system, which may be quite different from that of the respondent, society in general, or the most

relevant sets of stakeholders. As a result, we can compare companies on explicit dimensions but cannot assess the degree to which those dimensions really matter in terms of the firms' survival or success.

Conclusion: We must agree with Brown and Perry (1995a) that "the subjectivity of reputational indexes [and specifically FRS] is their strength in many ways" (p. 200), but we also concur with Sodeman (1995) that "this subjectivity is one of the great weaknesses of the [FRS] database" (p. 217). Such a position is not self-contradictory; it reflects the fact that every aspect of every major empirical methodology has substantial advantages and disadvantages. Social performance studies, as the forum authors point out, use one or more of the three ways to measure CSP: perceptual evaluations derived from rater-defined criteria, as in FRS; perceptual evaluations derived from researcher-defined criteria, as in SOCRATES; and observations of behavior relative to researcher-defined criteria, as in ad hoc investigations (Baucus and Near, 1991; Hill, Kelley, and Agle, 1990; Simpson, 1986). Fortunately, the strengths and weaknesses of each of these complement those of the other two. Thus, a multimethod approach combining these can go a long way toward overcoming the criticisms put forth in the forum.

THE BROWN AND PERRY DATABASE

The foregoing analyses might lead readers to believe that, in our minds at least, it is the Brown and Perry (1994, 1995a) database, rather than FRS, that is "terminally ill," in Baucus's words. We prefer to characterize their effort, even if perfectly executed, as premature. If, as we contend, there is insufficient evidence that any illusory halo, much less a financial one, afflicts the FRS data, then the rationale for halo removal does not yet exist. It should be clear, though, that we cannot and do not assert that such a halo has been proven *not to exist*.

If more credible evidence of an illusory financial halo eventually surfaces, then both the viability of the Brown and Perry (1995a) transformation as a research tool and the forum critiques of it deserve more serious scrutiny. Although we have argued above that such a condition does not yet exist, we undertake an evaluation of this aspect of the forum for the sake of comprehensiveness.

Of the forum authors, only Sodeman (1995) and Baucus (1995) take on the challenge of evaluating Brown and Perry's reasoning and methodology. Logsdon and Wartick (1995) limit the scope of their essays to the proper applications for the database and their attendant limitations. We

must also acknowledge that, as all good researchers do, Brown and Perry themselves (1995b) do their best to answer the questions raised in the Sodeman and Baucus articles. In this section, we give our viewpoint on these issues.

We noted earlier that Baucus (1995) cited volatility in the FRS data over time as invalidating their use in longitudinal studies. She also expresses similar concerns regarding the Brown and Perry residuals:

> The technique of partialling out variance . . . assumes that correlations between rating dimensions are inflated in a stable and consistent manner over time, and we have no evidence to support that assumption. (p. 231)

In reply, Brown and Perry (1995b) correctly point out that their method removes the financial performance effects for each year separately, so that this assumption is not inherent in the process. It is true, however, that each year's data encompass some degree of difference in the number of companies included, the configuration of the "halo" being removed, and so on. All of this compounds the volatility that Baucus cited (and that we examined above) regarding the characteristics associated with the under-lying FRS data. Our earlier conclusions about FRS volatility and longitu-dinal analysis are applicable in this context as well.

Sodeman (1995) also finds the transformed database to be unsuitable for time-series research, although primarily because "the full 10-year span of data is provided for only 75 companies" (p. 220). Although this represents a considerable loss of sample size, we do not assume automat-ically that the reason for the loss is related to all likely research questions and hypotheses, the operative validity criterion. Otherwise, it is not clear to us that a sample size of 75 firms is inadequate per se. Nor do we assume, as apparently Sodeman does, that the entire decade of data must be included in any longitudinal study. Much of the sample variation owes to the fact that FRS rated a much smaller group of companies in its first 2 years than it has in any subsequent year.

Although such criticisms can be defended or accommodated rather easily, both Sodeman (1995) and Baucus (1995) cite a shortcoming that figures to be crucial, if not fatal, to the internal validity of the halo-removal procedure. Of the two articles, Baucus discusses the issue in more detail, as in this passage:

> The halo-removal technique, as implemented, combines service and manu-facturing, high- and low-technology firms, and other types of firms in one regression equation for each year. . . . The authors need to partial out the

> effects of [industry variation], rather than simply averaging across all firms
> in a given year and assuming little variation between firms on these factors.

In an earlier section of this article, we alluded to the importance of industry differences; that lesson is applicable in this context as well and cannot be overemphasized. Going one step further, one might ask whether financial performance should be removed not just at the industry-specific level but also at the firm-specific level. We think not, but primarily on methodological, not conceptual, grounds. In the raw survey, FRS respondents evaluate all of the firms listed in the industry. Perhaps only the investment analysts would have comparable information regarding all of those companies, and even then they would be unlikely to be influenced by firm-centered financial performance. Whatever financial halo might exist, we expect that it would be focused on the industry level.

Of the forum authors, only Baucus addresses another concern we believe to be critical, and then only briefly. This is the actual empirical specification of the financial halo, best appreciated in Brown and Perry's (1994) original study. Their halo index consists of two variables averaged over a 3-year period (return on assets and sales growth) and three single-year measures (relative market value/book value, sales in logarithmic form, and the debt/equity ratio). The process that ultimately led to these five dimensions appears to be mostly a matter of brute empiricism, after first surveying the literature to find the financial measures most commonly used as independent variables. The authors then reduced the number of variables, apparently on the basis of two criteria, multicollinearity and statistical significance relative to the FRS attributes. It is not clear how they decided which of the collinear variables would be retained and which would be eliminated from the model.

It would be useful to try to gauge the appropriateness of this procedure; that is, how well it approximates whatever financial halo influences the attribute ratings. One way to do this would be to factor-analyze the transformed measures, under the assumption that if the halo is removed, the single-factor structure should disappear as well. Unfortunately, those data are not available.

Another method might be to examine the effect that halo removal has on the relationships among the FRS attributes. This technique is not as sophisticated, but the information already exists in Table 4 of Brown and Perry's (1994: 1356) primary study. In it, we can compare the 28 attribute intercorrelations prior to transformation (ranging from 0.68 to 0.95, 23 of them greater than 0.75) with those after transformation (0.47 to 0.91, with

15 greater than 0.75). If the influence of financial performance indeed had been removed, we would have expected a greater reduction in these correlations. The best explanation for this phenomenon is that the same response aggregation that masks variation in the factor analysis (as discussed in the halo premise section above) also serves to blunt efforts to remove intercorrelation.

Finally, Baucus (1995) suggests "that it would be simpler and easier to include performance measures as control variables in any analysis" (p. 232) as an alternative to Brown and Perry's (1994) halo-removal technique, but only assuming that the researcher has established that a halo related to those performance measures is present. Brown and Perry (1995b) reply that using control variables in this way is "econometrically equivalent" (p. 239) to their halo-removal process. We agree with their assessment *if* the source of the halo is equivalent for each. Unfortunately, we have already seen that there are serious doubts about halo specification in the residual data because of its lack of industry controls. If that obstacle were overcome, then applying the transformed measures would be superior to using control variables from an efficiency perspective because collecting additional predictor data for large samples is always a substantial undertaking.

Conclusion: We agree with Baucus (1995) in her call for a moratorium on the use of halo indexes, although we would limit the application of such a call solely to the index at issue here. Besides the crucial controversy over the existence of a halo, there is sufficient reason to suspect that the specific technique used by Brown and Perry (1994, 1995a) might not be very effective at extracting financial performance from FRS attributes. A different configuration of model variables might work better, but we would doubt that, given the huge obstacle created by the ecological inference artifact.

RESEARCH IN CSP AND THE ROLE OF FRS

At this point, we address the question that is always relevant in research: "So what?" That is, even if all that we say in these pages is absolutely true, what is its practical importance, and what should we do about it? In other words, what is the importance of this essay, and the *B&S* forum, in the context of the reputation and CSP literature? Our answer is that there certainly is much here that can and should be applied to future research on corporate social performance, including the clarification of our understanding of CSP itself. This is our final topic.

In our earlier discussion of the forum premise, we acknowledged the singular contribution of Baucus (1995) regarding the issue of true versus illusory halos. We believe that the point deserves considerable expansion because it is key to a conceptual understanding of CSP, financial performance, reputation, organizational effectiveness, and other related topics. Our contention is that there is a halo encompassing the full range of meaningful characteristics of a firm, but it is neither just financial nor just social. It might best be described as simply "excellence" (as in the Peters and Waterman, 1982, best-seller). Let us then posit a conceptual basis for the excellence halo phenomenon.

Ever since the publication of Fombrun and Shanley's (1990) article, we have been puzzled at the willingness of CSP scholars to assume that any significant direct relationship between social and financial performance measures must be erroneous. Don't we all advocate, in our classes and in our normative pronouncements, that organizations follow socially responsible courses of action? It seems inconceivable that anyone would do so if they believed that "taking the high road" is always hazardous to a company's economic health. If you feel otherwise, you must follow a line of reasoning in which at least some compatibility between social and financial performance is the general rule. If so, then it is only common sense that a competent management should pursue both.

Keep in mind that this is not the same as saying that socially beneficial corporate acts are not justified unless there is an economic payoff, or that social performance cannot be truly moral if it pays for itself or even makes a profit. Such thinking is indicative of what Wicks (1996; based on Freeman, 1994, and Sen, 1987) calls the separation thesis: "the view within which business and ethics are seen as conceptually distinct and separate realms with their own concepts, language, and logic" (Wicks, 1996: 113). We, instead, maintain that both ethics and economics ultimately are regulated by society (through stakeholders), using a common general mechanism, equally applicable to both. Excellence is finding the appropriate balance of financial and social performance and, wherever possible, choosing paths that serve both shareholders and stakeholders well.

Within this perspective, there is ample room for multiple and varied causation of behaviors. In the 1980s, when doing business in the Union of South Africa was controversial because of that country's apartheid policy, there was no uniform reaction by the firms that already had established operations there. Many such companies left because of real or threatened pressure from institutional investors (economic), dissatisfaction with the progress of their own efforts to create social change from

within (ethical), discomfort in the face of adverse publicity (discretion-ary), or some combination of these. Other firms stayed to benefit from low manufacturing costs or access to new markets (economic), to fulfill the obligations of their government contracts (legal), to provide better em-ployment opportunities for the South African workers than they would otherwise have (ethical), or some combination. Finally, some chose to leave/stay as a result of considering a mixture of the above points, and probably more, as trade-off issues.

Through various stakeholders, society judges these actions and their attributed motives, within the contexts of those stakeholders' value sys-tems. Because there is so much variance across those systems, the net effect on the firm and its reputation is nearly always mixed to some extent. It is only at this high level of abstraction that perceptual databases, such as FRS and SOCRATES, operate; it is only at the same level of abstraction that the interpretations drawn from them will have the ring of truth. Therefore, when researchers conclude that social performance is posi-tively, or negatively, related to financial performance, it should be with the realization that such a simple conclusion can only be a general one, masking a host of complex, underlying relationships.

For example, the public, especially the American public, wants a pristine environment, but not at any price. Twenty years ago, it would have been unthinkable that smoking would be prohibited anywhere in a public building. Yet, although such bans are commonplace today, a total ban on tobacco products remains an extremely thorny prospect, and the U.S. government still subsidizes the growing of tobacco. Academia has a role here, too. Indeed, since the first appearance of the *Fortune* survey, re-searchers have been using it in an attempt to understand the SP/FP link in all its variations. That brings us full circle to our current concern because it is the ultimate goal of investigations such as the forum, and this article, to enhance understanding by advancing the effectiveness of the methods we use.

The key to any successful application of FRS is cogently summarized in Wood (1995): "As always, it is up to us, the scholarly community, to generate valid, reliable measures of our constructs; to use those measures for what they are, no more and no less; and to keep in mind that our databases and empirical research exist so that we can build and confirm better theory" (p. 198).

With respect to FRS, this primarily amounts to avoiding the temptation of operationalizing all of corporate social performance as any single

stakeholder criterion. That is, CSP is not completely defined by individual stakeholder interests, such as environment or employee relations or customer satisfaction or community involvement. If our data are limited to such narrow perspectives, then our conclusions must be correspondingly narrow. If we intend to draw empirical conclusions about CSP in its multifaceted glory, then we must use a multifaceted operationalization. The multiple attributes of the FRS structure offer opportunities to do both single-stakeholder and comprehensive, though not exhaustive, CSP investigations.

The latter focus, however, presents a dilemma for researchers because on one hand, results of such studies would reflect the reality of multiple value systems. On the other hand, as Brown and Perry (1995b) note, investigators "yearn for valid, well-behaved data—data that are observable, relevant, free of measurement error and bias, normally distributed, and that meet the assumptions necessary to perform statistical tests. Unfortunately, data do not come made to order" (p. 238). This is certainly the case with FRS.

We agree with Brown and Perry that the validity of databases such as FRS should not be too readily dismissed, especially when the effects of their flaws are not systematic or can be overcome substantially by well-designed adjustment techniques. We also credit the same authors for their attempt to provide just such a technique for social performance scholars. Their intentions are meritorious. Unfortunately, they based their remedy on a diagnosis (financial halo) of questionable accuracy, and their method overlooked some important concerns (for example, control for industry differences). Thus, their specific prescription might have undesirable side effects. Indeed, from our perspective, we are unconvinced regarding the existence of a malady requiring treatment.

We share the concern of many of the forum authors about *potential* sources of invalidity in the *Fortune* data. Based on our investigation, however, it is questionable to conclude that invalidity is actually manifest. Indeed there are plausible arguments that suggest the opposite: that a positive relationship between financial performance and social performance is more likely to be true than illusory. We recommend that researchers should continue testing FRS for validity, especially if it can be done at the level of respondent ratings, because if validity can be shown, this database holds substantial promise as a powerful instrument in reputation, and CSP, research. By all means, throw out the bathwater, but only the bathwater.

NOTE

1. Kinder, Lydenberg, Domini & Co., Inc. (KLD) compiles a database of corporate ratings similar to FRS, except that only the latter includes some financial performance attributes. SOCRATES is the commercial version and formal name of the database and is marketed to investment advisors. It includes not only the ratings for each firm evaluated but also a detailed account of the rationales behind them. Researchers typically analyze the purely statistical version of the database, which does not have a separate name. This source usually is called the KLD database; it is a matrix of the same attribute variables and cases (companies) as are contained in SOCRATES.

The primary attributes rated in SOCRATES are community involvement, employee relations, product quality and safety, environmental responsibility, and diversity. These categories compare to the FRS subset of community and environmental responsibility; ability to attract, develop, and retain talented people; quality of products or services; and innovativeness. (For a detailed discussion of the KLD database underlying SOCRATES, including criticisms, see Wood and Jones, 1995, and Waddock and Graves, 1997).

REFERENCES

"America's Most Admired Corporations." 1997. [Online]. Available: http://www.fortune.com/. (Accessed January 3, 1997)

Asch, Solomon E. 1946. "Forming Impressions of Personality." *Journal of Abnormal and Social Psychology* 41:258-90.

Baucus, Melissa S. 1995. "Halo-Adjusted Residuals—Prolonging the Life of a Terminally Ill Measure of Corporate Social Performance." *Business & Society* 34:227-35.

Baucus, Melissa S., and Janet P. Near. 1991. "Can Illegal Corporate Behavior Be Predicted? An Event History Analysis." *Academy of Management Journal* 34:9-36.

Brown, Brad, and Susan Perry. 1994. "Removing the Financial Performance Halo From *Fortune*'s 'Most Admired' Companies." *Academy of Management Journal* 37:1347-59.

———. 1995a. "Halo-Removed Residuals of *Fortune*'s 'Responsibility to the Community and Environment'—A Decade of Data." *Business & Society* 34:199-215.

———. 1995b. "Some Additional Thoughts on Halo-Removed *Fortune* Residuals." *Business & Society* 34:236-40.

Campbell, Donald T., and Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin.

Dillon, William R., Narendra Mulani, and Donald G. Frederick. 1984. "Removing Perceptual Distortions in Product Space Analysis." *Journal of Marketing Research* 21:184-93.

Fisher, Anne B. 1996. "Corporate Reputations: Comebacks and Comeuppances." *Fortune* 133 (March 6): 90 ff.

Fombrun, Charles, and Mark Shanley. 1990. "What's in a Name? Reputation Building and Corporate Strategy." *Academy of Management Journal* 33:233-58.

Freeman, R. Edward. 1994. "The Politics of Stakeholder Theory: Some Future Directions." *Business Ethics Quarterly* 4:409-22.

Frooman, Jeff. 1997. "Socially Irresponsible and Illegal Behavior and Shareholder Wealth." *Business & Society* 36:221-49.

Fryxell, Gerald E., and Jia Wang. 1991. "What's in a Name Revisited: Evaluating the *Fortune* Corporate Reputation Data." Paper presented at the annual meeting of the Academy of Management, Miami Beach, August.

———. 1994. "The *Fortune* Corporate Reputation Index: Reputation for What?" *Journal of Management* 20 (Spring): 1-14.

Hill, Charles W. L., Patricia C. Kelley, and Bradley R. Agle. 1990. "An Empirical Examination of the Determinants of OSHA Violations." Paper presented at the annual meeting of the International Association for Business and Society, San Diego, March.

Jacob, Rahul. 1995. "Corporate Reputations." *Fortune* 131 (March 6): 54 ff.

Kinder, Peter D., Steven D. Lydenberg, and Amy L. Domini. 1992. *The Social Investment Almanac.* New York: Holt.

King, Gary. 1997. "Reprints and Reprints." [Online]. Available: http://gking.harvard.edu/preprints.shtml. (Accessed October 7, 1997)

Lee, Dwight R., and Richard B. McKenzie. 1994. "Corporate Failure as a Means to Corporate Responsibility." *Journal of Business Ethics* 13:969-78.

Logsdon, Jeanne M., and Steven L. Wartick. 1995. "Theoretically Based Applications and Implications for Using the Brown and Perry Database." *Business & Society* 34:222-6.

Murphy, Kevin R., Robert A. Jako, and Rebecca L. Anhalt. 1993. "Nature and Consequences of Halo Error: A Critical Analysis." *Journal of Applied Psychology* 78:218-25.

Peters, Thomas J., and Robert H. Waterman Jr. 1982. *In Search of Excellence: Lessons From America's Best-Run Companies.* New York: Harper & Row.

Robinson, William S. 1950. "Ecological Correlation and the Behavior of Individuals." *American Sociological Review* 15:351-7.

———. 1951. "The Logical Structure of Analytic Induction." *American Sociological Review* 16:812-8.

Schlossberger, Eugene. 1994. "A New Model of Business: Dual-Investor Theory." *Business Ethics Quarterly* 4:459-74.

Sen, Amartya Kumar. 1987. *On Ethics and Economics.* Oxford, UK: Basil Blackwell.

Simpson, Sally S. 1986. "The Decomposition of Antitrust: Testing a Multi-Level, Longitudinal Model of Profit-Squeeze." *American Sociological Review* 51:859-75.

Smith, H. W. 1981. *Strategies of Social Research.* 2d ed. Englewood Cliffs, NJ: Prentice Hall.

Sodeman, William A. 1995. "Advantages and Disadvantages of Using the Brown and Perry Database." *Business & Society* 34:216-21.

Szwajkowski, Eugene, and Raymond E. Figlewicz. 1995. "Reputation, Shareholder Valuation, and Performance: An Empirical Examination of Corporate Effectiveness." Paper presented at the annual meeting of the Academy of Management, Vancouver, BC., August.

Thorndike, Edward L. 1920. "A Constant Error in Psychological Ratings." *Journal of Applied Psychology* 4:25-9.

Van Buren, Harry J. III. 1995. "Business Ethics for the New Millennium." *Business and Society Review* 93 (Spring): 51-55.

Waddock, Sandra A., and Samuel B. Graves. 1997. "Quality of Management and Quality of Stakeholder Relations: Are They Synonymous?" *Business & Society* 36:250-79.

Waddock, Sandra A., and John F. Mahon. 1991. "Corporate Social Performance Revisited: Dimensions of Efficacy, Effectiveness, and Efficiency." Pp. 231-62 in James E. Post (ed.), *Research in Corporate Social Performance and Policy,* vol. 12. Greenwich, CT: JAI.

Wicks, Andrew C. 1996. "Overcoming the Separation Thesis: The Need for a Reconsideration of Business and Society Research." *Business & Society* 35:89-118.

Wokutch, Richard E., and Elizabeth W. McKinney. 1991. "Behavioral and Perceptual Measures of Corporate Social Performance." Pp. 309-30 in James E. Post (ed.), *Research in Corporate Social Performance and Policy*, vol. 12. Greenwich, CT: JAI.

Wood, Donna J. 1995. "The *Fortune* Database as a CSP Measure." *Business & Society* 34:197-98.

Wood, Donna J., and Raymond E. Jones. 1995. "Stakeholder Mismatching: A Theoretical Problem in Empirical Research on Corporate Social Performance." *The International Journal of Organizational Analysis* 3 (3): 229-67.

Eugene Szwajkowski received his Ph.D. from the University of Illinois at Urbana–Champaign and is a management consultant and adjunct professor of management at the University of Illinois at Chicago. He also has been a visiting professor at Washington University in St. Louis, Notre Dame, and the University of Illinois at Urbana–Champaign. His research interests are in organizational misconduct, business ethics, and the social issues area of strategic management, especially the relationship between corporate social and financial performance.

Raymond E. Figlewicz earned a DBA at Southern Illinois University. He is an assistant professor of accounting at the University of Michigan at Dearborn and has also taught at Kent University and Loyola University of Chicago. His current research interests include examining critical social, financial, and market factors associated with significant firm events; analyzing cash flows; and investigating multiple facets of the oil and gas industry.