ELSEVIER

An Empirical Comparison of Published Replication Research in Accounting, Economics, Finance, Management, and Marketing

Raymond Hubbard
DRAKE UNIVERSITY
Daniel E. Vetter
CENTRAL MICHIGAN UNIVERSITY

The results of a large-scale content analysis of 18 leading business journals covering the 22-year time period 1970 to 1991 show published replication and extension research is uncommon in the business disciplines. For example, such research typically constitutes less than 10% of published empirical work in the accounting, economics, and finance areas, and 5% or less in the management and marketing fields. Further, when such work is undertaken the results usually conflict with existing findings. This raises the prospect that empirical results in these areas may be of limited value for guiding the development of business theory and practice. Strategies for cultivating a replication research tradition to facilitate knowledge development in the business disciplines are suggested. J BUSN RES 1996. 35.153–164

Whereas uncorroborated empirical research outcomes must be considered tentative, their successful replication promotes confidence in the veracity of a discipline's cumulative knowledge base. Other things being equal, replication protects against the uncritical assimilation of specious empirical results into the literature. Replications with extensions serve to determine the scope and limits of empirical findings by seeing if they can be generalized to other populations, contexts, time periods, geographical areas, and so on. Indeed, the principle of replicability is widely acknowledged to be the touchstone of the scientific method (Kane, 1984), the hallmark of science (Blaug, 1992), and the most important criterion of genuine scientific knowledge (Rosenthal and Rosnow, 1984).

Two contemporaneous developments supporting the need for greater attention to the topic of replication in the business fields warrant emphasis. First are the recent appearances of a number of research articles on the subject. For example, contributions published in the accounting (Lindsay, 1994), economics (Dewald, Thursby, and Anderson, 1986; Feigenbaum and Levy, 1993; Levy and Feigenbaum, 1990; Mirowski and Sklivas, 1991; Poirier, 1988), finance (Kane, 1984), marketing (Hubbard and Armstrong, 1994; Hubbard, Brodie, and Armstrong, 1992; Raman, 1994; Zinkhan et al., 1990), and statistical (Ehrenberg and Bound, 1993; Lindsay and Ehrenberg, 1993) literatures discuss the important role of replications and extensions in the knowledge development process. Second are the changes in journal publication policies adopted in accounting (*Accounting Review*, 1990), consumer behavior (Monroe, 1992a, 1992b), economics (Ashenfelter, 1986), and marketing (Kinnear, 1992) to facilitate data-sharing and replication and extension research.

Given the concerns expressed previously by both researchers and journal editors about the issue of replication in the business disciplines, the objectives of the present article are timely. First, the importance of replication and extension research is outlined, especially as it pertains to safeguarding the empirical literature against Type I and other errors, and also as a means for gauging the generalizability of results. Second, several hypotheses are formulated regarding the conduct of replication research in the five business disciplines of accounting, economics, finance, management and organizational behavior (hereafter "management"), and marketing. More specifically, an examination is made of the quantity, timeliness, and outcomes of such published work for the period 1970 through 1991. Third, a description of the data collection procedures used in the study is provided. Fourth, the results of the hypothesis tests are discussed. Finally, steps deemed necessary to inculcate a replication tradition in the business fields are suggested.

Address correspondence to: Raymond Hubbard, College of Business and Public Administration, Drake University, Des Moines, IA 50311

Journal of Business Research 35, 153–164 (1996)

655 Avenue of the Americas, New York, NY 10010

ISSN 0148-2963/96/$15.00
SSDI 0148-2963(95)00084-6

Importance of Replication and Extension Research

Replication and extension research can play a major role in ensuring the integrity of a discipline's empirical results. This is accomplished in two major ways: first, by serving as a guard against Type I errors (erroneous rejections of the null hypothesis) and other questionable findings; second, and more importantly, by assessing the robustness and empirical generalizability of results, thus contributing to the growth of knowledge.

Protection against Type I and Other Errors

The need for replication as a protection against Type I and other erroneous or questionable results would seem to be especially pronounced in many business areas where the lack of powerful theories places an extraordinary burden on the ability of exploratory-empirical studies to sustain the disciplines. Typically, the researcher in these studies hopes to obtain statistically significant ($p < .05$) findings. Such outcomes are important because of the widespread perceived bias against the publication of null results. In response to this perception, some researchers may "strain for statistical significance" and thus inflate the occurrence of published Type I errors. The upshot is that Type I errors in the accounting (Burgstahler, 1987; Lindsay, 1994), economics (Feige, 1975), management (Mazen et al., 1987), and marketing (Hubbard and Armstrong, 1992) literatures are believed to be in excess of those prescribed by nominal alpha levels such as .05.

Similar concerns have been raised in a number of other disciplines. Greenwald (1975), in fact, maintains that Type I error proliferation in the social and behavioral sciences is so widespread as to call into question the scientific basis of much of the published literature. Data mining procedures for achieving statistically significant outcomes, combined with a reluctance among researchers to submit papers with "insignificant" ($p > .05$) results for publication, leads to what Rosenthal (1979) calls the "file drawer problem." Taken to the extreme, this problem states that journals are filled with the 5% of studies that are Type I errors, whereas the other 95% with null results languish in the file drawers of researchers. Yet as Walster and Cleary (1970) emphasize, only the publication of replications and failures to replicate will uncover Type I errors in the literature.

Some erroneous results, Type I or otherwise, eventually reach the textbooks. This is particularly unfortunate because once in the textbooks these errors take on the appearance of established truths. Some even become "classics," going for years without detection. Examples from the management area include the Hawthorne effect (Franke and Kaul, 1978) and Frederick W. Taylor's account of the loading of pig iron (Wrege and Perroni, 1974).

An example from the marketing literature underscores the ongoing potential damage to a discipline that can occur when uncorroborated journal results filter into textbooks. The example relates to Gorn's (1982) experiment claiming that product preferences can be classically conditioned through a single pairing with background music. Kellaris and Cox (1989) were concerned that Gorn's findings may have been due in part to demand artifacts, a concern heightened by the article's influential standing. They noted, for instance, that the study had been cited at least 34 times between 1982 and 1988. Perhaps more importantly, they said it was being presented in a number of consumer behavior textbooks, one of which used it as a basis for asserting that classical conditioning of product preferences is "well established and widely used." In failing to replicate Gorn's results in three well designed experiments, Kellaris and Cox concluded that single-exposure conditioning of product preferences is far from being "well established."

Generalizing Results

The building block of science is empirical generalization, and replication and extension research is the key to generalization (Lindsay and Ehrenberg, 1993). Unreplicated research findings, including those with "high" levels of statistical significance, are necessarily speculative in nature (Hubbard and Armstrong, 1994). They remain "virtually meaningless and useless" in themselves (Lindsay and Ehrenberg, 1993, p. 219). A research orientation focusing instead on establishing the scope and limits, or generalizability, of findings would benefit the business disciplines because generalization is an integral part of knowledge discovery. Systematically conducted replications and extensions serve this purpose.

Unfortunately, critics argue, the majority of the business (and social science) literature consists of fragmented and isolated findings. These uncorroborated studies, even though routinely accompanied by statistically significant results (Hubbard and Armstrong, 1992), provide a weak foundation for the development of business theory and practice (Hubbard, 1994).

Notwithstanding the above, success stories demonstrating how replication and extension research can lead to empirically generalizable results do exist. Consider, for example, the field of buyer behavior. Ehrenberg's (1988) work outlines the progress made in modeling brand purchase incidences and repeat buying patterns using the negative binomial distribution up through the increasingly more general results on individual purchase frequencies and brand choice made possible by the comprehensive Dirichlet model.

The beauty of the Dirichlet model lies in its ability to parsimoniously account for several empirical patterns of buyer behavior including, for example, those described by the duplication of purchase law and the double jeopardy (DJ) phenomenon. Consider the latter. The DJ phenomenon, which states that brands with smaller market shares not only are bought by fewer people in a given time period (penetration level) but are also bought less often (average frequency of purchase), has been subjected to much replication and extension research. For example, Ehrenberg and Bound (1993) note that DJ holds for over 50 different products (both convenience and shopping goods, differentiated and undifferentiated items, tangible goods and services), as well as for different distribution channels, different countries, different time periods, and so on. They also note

J Busn Res 155
1996:35:153–164

a few exceptions where DJ does not apply or applies only partially. Hence they show both the (still increasing) scope of the DJ findings as well as some of its limits. The work reported in Ehrenberg and Bound (1993) provides perhaps the quintessential example of the value of systematic replication and extension research in producing robust and generalizable results.

Hypotheses

In this section a number of hypotheses concerning replication and extension research in the business disciplines are formulated. These hypotheses relate to the quantity, timeliness, and outcomes of such work.

Quantity of Published Replication and Extension Research

When considering the developmental status of the various business disciplines it is not clear, on a priori grounds, as to which might be expected to publish the most replications and extensions. On the one hand, it could be argued that disciplines that have developed reasonably strong theoretical bases (for example, economics and finance) will publish more replications as a result of the need to continually test and refine established models. Empirical testing of the Phillips curve, the demand for money (economics), and the capital asset pricing model (finance), are exemplary in this regard. On the other hand it is just as plausible to assume that the less theory-driven disciplines (for example, management and marketing) will, of necessity, engage in repeated empirical explorations precisely because of the relative absence of well-developed theoretical frameworks.

The preceding observations suggest that replication research will be viewed as being equally important by all. This leads to the first hypothesis.

H1: Other things being equal, there is no difference in the publication frequencies of replications and extensions in the accounting, economics, finance, management, and marketing disciplines.

The nature of the principal data types (primary or secondary) used by the five disciplines might also have a bearing on the quantity of published replication work. Hubbard and Armstrong (1994) speculated that because of the additional time and costs that could be involved when relying chiefly on primary data sources, those disciplines utilizing mostly secondary data might have a comparative advantage in conducting replication research. This leads to the next hypothesis.

H1a: Disciplines relying mostly on secondary data sources will exhibit a higher publication frequency of replications and extensions than those more dependent on primary data.

Although presently considered "immature" (Lindsay and Ehrenberg, 1993), the five business disciplines nevertheless have aspirations of attaining scientific status. Given this goal, it might

	Successful Replication/Extension Attempt? Yes	No
Short-Term	Promotes confidence in the stability of initial results. Further work necessary to assess generalizability. Examples: *Accounting*: Jiambalvo (1982) supports Maher, Ramanathan & Peterson (1979). *Finance*: Servaes (1991) corroborates Lang, Stultz & Walking (1989).	Promotes lack of confidence in the stability of initial results. Further work necessary to resolve conflicts. Examples: *Management*: Woodman and Wayne (1985) contradict Terpstra (1981). *Economics*: Fishe & Wohar (1990) contradict Mankiw, Miron & Weil (1987).
Long-Term	Provides ongoing, cumulative evidence for the existence of a phenomenon, together with its scope and limits. Examples: *Management*: Bird & Fisher (1986) support Kirchner & Dunnette (1954). *Marketing*: Ehrenberg & Bound (1993) demonstrate the empirical robustness of the Dirichlet Model of Buyer Behavior.	May be of questionable value when too much time has elapsed. Earlier replication generally more useful. Example: *Marketing*: Webster & von Pechmann (1970) fail to replicate Haire (1950).

Figure 1. The relative utilities of replications over time.

be expected that the publication incidence of replications in these areas would increase over time. Hence the following hypothesis is offered.

H1b: The publication frequency of replications and extensions in accounting, economics, finance, management, and marketing will be higher for the time period 1980–1991 than it was during 1970–1979.

Timeliness of Replication Research

Other things being equal, rapid attempts to replicate and extend the findings of original studies are preferable to those conducted at later dates. Figure 1 depicts the relative utilities associated with successful and unsuccessful replication attempts performed over the short and long hauls.

Short-term, successful replications and extensions generate confidence in the reliability of initial research outcomes. In the accounting literature, for example, Jiambalvo's (1982) work on measuring the accuracy and congruence in the performance evaluation criteria of CPA personnel supported the results of an earlier study in this area by Maher, Ramanathan, and Peterson (1979). Similarly, in an example from the finance literature, Servaes (1991) was able to corroborate Lang, Stulz, and Walking's (1989) findings on the relationship between takeover gains and the q ratios of targets and bidders.

Short-term, unsuccessful replications cast doubt on the findings of the original study (and possibly on those attempting to replicate it). Failed replications may signal the need for additional work in that area. For example, Terpstra (1981) posited the existence of an inverse relationship between degree of methodological rigor and reported success of organization development interventions. Woodman and Wayne's (1985) subsequent research indicated no support for such an assertion. Likewise, the results of Fishe and Wohar's (1990) economics

study contradicted the claim of Mankiw, Miron, and Weil (1987) that the creation of the Federal Reserve System in 1914 marked the beginning of a new stochastic structure of interest rates.

The value of long-term, unsuccessful replications may be questionable. For example, in view of the enormous social, economic, and technological changes that transpired in the intervening 20 years, it is likely that few researchers were surprised by Webster and von Pechmann's (1970) failure to reproduce the findings of Mason Haire's (1950) classic "shopping list" study involving instant coffee. A failed or successful replication shortly after Haire's study was published would have been more informative.

Long-term, or ongoing, successful replication efforts obviously constitute the most desirable outcome. Empirical generalization is basic to knowledge development. Thus, Bird and Fisher (1986) found remarkably similar results to those reported by Kirchner and Dunnette (1954) some 30 years earlier concerning attitudes toward the employment of older workers. The ability of the Dirichlet model to subsume many empirical patterns of buyer behavior from 1959 to the present is a powerful example of the replicability of results over time (Ehrenberg and Bound, 1993).

With respect to the timeliness of replications across disciplines, the following null hypothesis is advanced:

H2: Other things being equal, there is no difference in the publication time lags between the original and replicated/extended studies in the accounting, economics, finance, management, and marketing disciplines.

Because failures to replicate call into question the original results, they may be considered more important than those providing confirmation. Consequently, their timeliness is of particular interest. Thus, the next hypothesis is suggested.

H2a: Replications and extensions whose results conflict with those of the original work will have a shorter publication time lag than those supporting it.

Owing to the possible time and cost economies associated with replications using secondary versus primary data, and those authors replicating their own work, the following two hypotheses are proposed.

H2b: Replications and extensions using secondary data will have shorter publication time lags than those using primary data.

H2c: Authors publishing replications and extensions of their own work will have shorter publication time lags than those carried out by independent researchers.

Outcomes of Replication and Extension Research

Basically, the results of a replication either support, partially support, or conflict with those of the original work. Because there is no compelling a priori argument to suggest that differences should exist in the frequency of these three outcomes across the five business disciplines encompassed in this study, the following hypothesis is postulated.

H3: Other things being equal, there is no difference in the proportion of replications and extensions that support, partially support, or conflict with their predecessors among the accounting, economics, finance, management, and marketing disciplines.

The type of data used, secondary or primary, may affect the outcomes of replication research. It is conceivable, for example, that studies using secondary data such as government statistics and CRSP tapes will produce a higher rate of supportive results given the relatively "objective" nature of these generally public sources of information. In contrast, studies utilizing primary data (as in survey and experimental research) are perhaps more vulnerable to criticisms that "subjective" factors involved in such things as questionnaire-wording or experimental demand artifacts may have operated to bias the outcomes. Unfortunately, it is difficult to predict whether these biases, if present, are more likely to result in studies confirming or disconfirming each other. In view of this indeterminacy, the following hypothesis is offered.

H3a: There is no difference in the proportion of replications and extensions that support, partially support, or conflict with the original studies by the type of data, secondary or primary, employed.

By helping to avoid any biases that may have been associated with the original work from carrying over to the replication, some researchers maintain that replications and extensions performed by independent (different) scholars generally are preferable to those where authors replicate their own work (Hubbard and Armstrong, 1994). In their discussion of the problem of "correlated replicators," Rosenthal and Rosnow (1984) question the value of 10 replications conducted by the same investigator versus 10 replications each performed by a different researcher.

Some journal editors agree (McCabe, 1984). Indeed, McCabe states that the editorial policy of the *Quarterly Journal of Business and Economics* is to reject the publication of replications and extensions conducted by the same authors. The likelihood is that replications by the same authors will usually confirm the results of their previous endeavors. Hence, the next hypothesis.

H3b: Authors replicating and extending their own work will show a higher incidence of supportive results than those conducted by independent investigators.

Method

Definitions

The definitions of replication and replication with extension are based on those used by other researchers interested in estimating their publication incidence (Brown and Coney, 1976; Reid et al., 1981). A replication is a duplication of a previously published empirical study whose purpose is to determine whether the findings of that study are repeatable. These "exact"

J Busn Res
1996:35:153–164

or "direct" replications attempt to reproduce, as faithfully as possible, the conditions of the original investigation by using the same variable definitions, settings, subjects, measurement instruments, sampling methods, analytical techniques, and so on. Repeating a study with another sample taken from the same population is an example of a replication.

A replication with extension is a duplication of a previously published empirical study whose purpose is to evaluate the generalizability of earlier research results. The intent is not to alter the conceptual relationships analyzed in the original study, but to test them differently by modifying certain aspects of the initial design. Examples would be making changes in either the manipulated or measured variables, but not both, investigating the influence of additional variables, repeating the study using different populations, contexts, geographical areas, time periods, or using any combination of the aforementioned changes.

The preceding paragraph provides illustrative rather than exhaustive examples of the kinds of modifications in research designs subsumed by the definitions. These definitions are also congruent with Brinberg and McGrath's (1985) ideas about the validity and robustness of empirical results.

Samples

The publication frequency of replications and extensions in the business literature was examined by content-analyzing a simple random sample of 25% of the issues from leading journals in the accounting, economics, finance, and management disciplines from 1970 through 1991. For marketing, a 50% simple random sample of issues was used. The original intention was to use a 50% random sample for all five disciplines. It was soon discovered, however, that this was overly ambitious. As it was, collection of the data for the present project took several years.

Leading journals were targeted in this study because they enjoy the widest readership, prestige, and influence on scholarly thought and practice. Replications and extensions published in the lower tiers of the journal hierarchy are unlikely to attract attention.

In all, some 472 issues drawn from 18 business journals were involved in the present study (see Appendix A for further sample information). These 472 issues yielded a total of 6,400 research papers (articles, notes, and commentaries), 4,270 (66.7%) of which were empirical. The current work focuses on the empirical studies only.

Both authors examined and classified independently each of the empirical papers. Although the use of more than two judges might be expected to improve the reliability of the estimates of published replication research in the business literature, the onerous nature of the task precluded this possibility. Following Reid et al. (1981), an article was not classified as a replication or extension unless it contained a specific citation of the original study. This did not mean that the replicating authors had to explicitly identify their work as being a replication or extension (this was the present authors' responsibility). When uncertainty about an article's status occurred, it was resolved by both of the present authors. Typically, such articles were included in the study. For example, a particularly liberal practice of counting as extensions those studies that modified both the manipulated and measured variables was followed. Thus, to the extent that a misclassification of papers may have occurred, the effect has been to overstate the incidence of published replications and extensions.

Cohen's (1960) kappa (k) was used as a measure of rater agreement because it possesses a number of desirable properties. Fleiss (1981) discusses several of these. For example, it incorporates a means of correcting for chance agreement among raters. It is also easy to interpret, for a value of $k = 1$ indicates complete agreement, a value of $k > 0$ indicates that observed agreement is greater than chance agreement, and a value of $k < 0$ indicates that observed agreement is less than chance agreement. As Fleiss (1981) also notes, values of k greater than 0.75 are generally interpreted as representing excellent agreement beyond chance, values between 0.40 and 0.75 suggest good agreement beyond chance, and values below 0.40 reflect poor agreement beyond chance.

Excellent levels of rater agreement were obtained for each of the accounting ($k = .76$, $z = 6.3$), economics ($k = .83$, $z = 10.4$), finance ($k = .87$, $z = 9.4$), management ($k = .85$, $z = 9.4$), and marketing ($k = .85$, $z = 6.4$) disciplines. All rater agreements were statistically significant at the $p < .001$ level. Furthermore, there was no statistically significant difference between the highest (0.87) and lowest (0.76) k values ($z = 1.6$).

Results

Quantity of Published Replication and Extension Research

There were no exact replications in the sample of research reports. This outcome is not as surprising as it may appear to be at first glance. It is virtually impossible to conduct an exact replication, if only because of the different time periods that must necessarily be involved in executing original and replicated studies. Lindsay and Ehrenberg (1993) and Rosenthal and Rosnow (1984) discuss this issue further. Some 266 articles, however, qualified as replications with extensions. This represents 6.2% of the published empirical papers (Table 1). Appendix B lists replication rates for each of the 18 journals involved in the study.

No differences were hypothesized to exist in the publication frequency of replications and extensions across the five business disciplines (H1). When considering the results for the accounting (8.6%), economics (8.4%), finance (9.7%), management (5.3%), and marketing (2.9%) disciplines, H1 was rejected ($\chi^2_{(4)} = 46.7$, $p < .001$). Utilizing the additivity of independent chi-squares, the unplanned contrast between marketing (2.9%) and the average proportion of extensions for the other four disciplines (7.4%) was statistically significant ($\chi^2_{(4)} = 30.1$, $p < .001$).

Had the contrast investigated previously been formulated on an a priori basis, rather than suggested by the data as in

Table 1. Replications and Extensions in Five Business Disciplines

Discipline	Number of Empirical Studies	Replications and Extensions	Percentage of Replications and Extensions	95% Confidence Interval
1970–1991				
Accounting	373	32	8.6	2.8
Economics	980	82	8.4	1.7
Finance	556	54	9.7	2.5
Management	1,222	65	5.3	1.3
Marketing	1,139	33	2.9	1.0
Total	4,270	266	6.2	0.7
1970–1979				
Accounting	115	8	7.0	4.7
Economics	402	41	10.2	3.0
Finance	203	19	9.4	4.0
Management	602	31	5.1	1.8
Marketing	516	17	3.3	1.5
Total	1,838	116	6.3	1.1
1980–1991				
Accounting	258	24	9.3	3.5
Economics	578	41	7.1	2.1
Finance	353	35	9.9	3.1
Management	620	34	5.5	1.8
Marketing	623	16	2.6	1.2
Total	2,432	150	6.2	1.0

this case, the use of one degree of freedom instead of the four used here would have been appropriate. But as the contrasts examined were post hoc in nature, Miller (1981) recommends the use of c-1 degrees of freedom (where c is the total number of groups in the study), hence the use of four degrees of freedom. This more stringent procedure reduces the likelihood of committing a Type I error.

The contrast between the proportion of extensions for the accounting, economics, finance, and management areas was also statistically significant ($\chi^2_{(4)} = 16.6$, $p < .005$). It was therefore necessary to further partition the chi-square for these four groups in order to discover the source(s) of this difference. The data from Table 1 suggest that the management area is the likely candidate. An unplanned contrast between management's (5.3%) proportion of replications versus the average proportion for accounting, economics, and finance (8.8%) is statistically significant ($\chi^2_{(4)} = 15.5$, $p < .005$), whereas that between the latter three disciplines is not ($\chi^2_{(4)} = 1.1$).

In summary, marketing publishes fewer extensions than the other four business fields of accounting, economics, finance, and management. In turn, management publishes a smaller proportion of extensions than occurs in accounting, economics, and finance.

H1a stated that those disciplines relying predominantly on secondary data sources might be expected to publish a greater frequency of replication research than those more dependent on primary data. In addressing this hypothesis, a calculation was made for all 18 journals of the percentage of empirical research papers that used secondary data (x) (Appendix B). These data were then regressed (OLS) against the percentage of empirical studies, by journal, that were extensions (y). The resultant equation, $y = 0.039 + 0.049x$, yielded a statistically significant coefficient for x ($t = 1.94$, $p < .05$, one-tail test). Thus H1a is supported. Note, however, that the goodness of fit ($r^2 = 0.19$) is unimpressive, thereby illustrating the distinction between statistical and substantive significance.

Because inattention to the issue of replication research could have deleterious consequences for knowledge development, H1b predicted a higher publication incidence of such works in the five business fields during 1980 to 1991 than during 1970 to 1979. Contrary to expectations, economics' share of replications in fact declined (10.2% to 7.1%), as did the share for marketing (3.3% to 2.6%). The other three disciplines—accounting (7.0% to 9.3%), finance (9.4% to 9.9%), and management (5.1% to 5.5%)—posted increases (Table 1). None of these shifts, however, was statistically significant at the .05 level. The analyses do not support H1b.

Timeliness of Replication Research

Analysis of variance was used to evaluate hypotheses H2, H2a, H2b, and H2c on the timeliness of replication research. The continuous dependent variable is the time lag between the publication dates of the original and replication studies, whereas discipline (the five business areas), replication outcome (sup-

Table 2. ANOVA on Publication Time Lags Between Original and Replication Studies

Source of Variation	Mean Publication Time Lag[a]	Sample Size	Eta[b]	*df*	*F*	Significance of *F*
Discipline			0.15	4	1.42	0.23
Accounting	3.6 (2.0)	32				
Economics	4.2 (3.6)	82				
Finance	3.9 (3.6)	54				
Management	4.6 (5.0)	65				
Marketing	5.6 (4.6)	33				
Outcome			0.06	2	0.49	0.61
Support	4.8 (5.4)	72				
Partial Support	4.3 (3.3)	73				
Conflict	4.1 (3.4)	121				
Data type			0.05	1	0.61	0.44
Primary	4.9 (4.8)	103				
Secondary	4.0 (3.3)	163				
Investigator			0.25	1	15.08	0.0001
Independent	4.6 (4.2)	215				
Nonindependent	3.0 (2.6)	51				
Explained				29	1.69	0.02
Residual				236		
R^2 = 0.17						

[a] Time lag in years. Values in parentheses are standard deviations.
[b] Calculated following Rosenthal and Rosnow (1984).

port, partial support, conflict), data type (primary, secondary), and status of the investigator (independent, nonindependent) are the independent variables.

Ceteris paribus, H2 stated that the time lag between the publication of the original and replication studies would not be different across the five business fields. This hypothesis was supported by the data. Table 2 shows that there was no statistically significant main effect of discipline type on publication time lags (F = 1.42, p = .23, eta = 0.15). The average time lag in years for each of the disciplines is as follows: accounting (3.6), economics (4.2), finance (3.9), management (4.6), and marketing (5.6). The average publication time lag for all five disciplines was 4.3 years.

Because they question the credibility of the original findings, and consequently may be viewed as being more important, it was hypothesized that replications with conflicting outcomes would have shorter publication time lags than those confirming earlier results (H2a). This prediction was borne out. Replications whose outcomes conflicted with previous research had an average time lag of 4.1 years, whereas those that partially (4.3 years) or fully (4.8 years) supported their predecessors had longer lags. There was, however, no statistically significant main effect of replication outcome on publication time lag (F = 0.49, p = .61, eta = 0.06). H2a is therefore rejected (Table 2).

It was also anticipated that owing to potential time and cost economies, replications and extensions using secondary data sources would exhibit a shorter publication time lag than those using primary data (H2b). This was the case, but as Table 2 reveals, there was no statistically significant main effect for replications using secondary data (4.0 years) versus those using primary data (4.9 years) on publication time lags (F = 0.61, p = .44, eta = 0.05). H2b is therefore rejected.

Researchers replicating or extending their own previously published work should enjoy comparative time and cost economies over those performed by independent (different) investigators (H2c). Table 2 shows a statistically significant main effect of investigator status on the time lag between the publication of original and replication studies (F = 15.08, $p < .0001$, eta = 0.25). For independent replicators this time lag was 4.6 years; for those replicating their own published work it was 3.0 years. H2c receives support.

Outcomes of Replication and Extension Research

Hypotheses H3, H3a, and H3b, concerning replication outcomes, were tested via multiple logistic regression. In its most general form, the regression relationship stated that replication outcome is a function of three variables, namely, the discipline involved, the type of data used, and the status of the investigator. Based on the conclusions reached by the replicating author(s), the polytomous dependent variable had three possible outcomes: those that supported, partially supported, or conflicted with the results of the original study. The "support" response acted as the reference category. The variable "discipline"

consisted of the five subject areas, with marketing serving as the reference group. Primary (reference) or secondary data were the two options for the variable "data type," whereas independent and nonindependent (reference) replicators defined the "investigator status" variable. The SAS (1989) PROBIT procedure applied to the 266 observations on replication outcomes yielded the maximum likelihood parameter estimates and associated information presented in Table 3.

H3 stated the null hypothesis of no statistically significant differences in the proportion of replications and extensions supporting, partially supporting, and conflicting with the results of their predecessors among the five areas. H3 is rejected. The influence of the "discipline" variable on outcomes is statistically significant ($\chi^2_{(4)} = 10.87$, $p < .05$). Further examination shows that the management field contributes to this difference ($\chi^2_{(1)} = 7.02$, $p < .01$). The negative sign on the management parameter estimate (−1.09) indicates that replications in this field are less likely to conflict with the initial studies, but rather are more likely to support or partially support them.

Other things being equal, extensions whose results conflict with earlier studies raise questions about the robustness of published results. And as Table 4 shows, there are many conflicting outcomes across the disciplines, with only management bucking the trend. Overall, the average publication frequency of conflicting results is 45.5% (a figure rising to 55.7% when management is omitted from the calculations). When coupled with those studies that offer only partial confirmation of earlier findings (27.4%), it can be seen that an average of about 27% of replications support previous work (only 20% with management excluded).

H3a predicted that different data types, primary versus secondary, would have no influence on replication outcomes. This hypothesis was upheld. Table 3 indicates that whereas use of secondary data is more likely to produce conflicting results, the coefficient (0.21) is not statistically significant.

H3b stated that authors replicating their own work would be more likely to obtain supportive results than those performed by independent (different) researchers. The statistically significant coefficient (1.77, $\chi^2_{(1)} = 28.5$, $p < .0001$) for investigator status in Table 3 affirms H3b. Of the 215 extensions conducted by independent researchers, 54% (116) conflicted with, 27.4% (59) partially supported, and 18.6% (40) supported the results of the original studies. In stark contrast, of the 51 cases where authors extended their own previous work, only 9.8% (5) conflicted with earlier results. And even in these few instances, the conflicts were usually anticipated. About 28% (14) of these nonindependent replications provided partial support for earlier work, whereas 63% (32) were in full accord.

Table 3. Logistic Regression Analysis of Replication Outcomes

Variable	ML Parameter Estimate	Standard Error	χ^2[a]	*df*	Odds Ratio
Intercept 1 (conflict)	−1.61	0.044	13.74[b]	1	
Discipline			10.87[c]	4	
Accounting	0.07	0.53	0.02	1	1.07
Economics	0.21	0.51	0.16	1	1.23
Finance	−0.05	0.53	0.01	1	0.95
Management	−1.09	0.41	7.02[d]	1	0.34
Marketing[e]	–	–	–	–	
Data type					
Primary[e]	–	–	–		
Secondary	0.21	0.41	0.26	1	1.23
Investigator					
Independent	1.77	0.33	28.51[b]	1	5.87
Nonindependent[e]	–	–	–		
Intercept 2 (partial support)	1.47	0.16		1	
$n = 266$					
Log - likelihood = −249.05					

[a] The χ^2 tests for individual parameter estimates are Wald tests.
[b] $p < .0001$.
[c] $p < .05$.
[d] $p < .01$.
[e] Indicates reference category.

Developing a Replication Tradition in the Business Disciplines

The lack of a replication tradition frustrates knowledge development by casting doubt on the validity, reliability, and generalizability of business results. Regrettably, the findings presented in this article suggest that knowledge development will continue to be impeded in the business disciplines. For example, replication research typically constitutes less than 10% of published empirical work in the accounting, economics, and finance areas, and 5% or less in the management and marketing fields. In terms of the amount of journal research-space, in pages,

Table 4. Outcomes of Replications and Extensions: 1970–1991

Discipline	Support[a]	Partial Support[a]	Conflict[a]	Total
Accounting	6 (18.8)	10 (31.3)	16 (50.0)	32
Economics	15 (18.3)	17 (20.7)	50 (61.0)	82
Finance	12 (22.2)	11 (20.4)	31 (57.4)	54
Management	32 (49.2)	24 (36.9)	9 (13.8)	65
Marketing	7 (21.2)	11 (33.3)	15 (45.5)	33
Total	72 (27.1)	73 (27.4)	121 (45.5)	266

[a] Values in parentheses are percentages.

J Busn Res
1996:35:153–164

devoted to replication research, the figures are even lower: accounting (4.5%), economics (2.9%), finance (4.4%), management (3.6%), and marketing (1.3%).

Hubbard and Armstrong (1994) discuss several reasons for the paucity of replication and extension research in the literature. These include: (1) misinterpreting statistical significance levels (especially $p < .05$) as a measure of the replicability of any given finding, (2) assuming that the statistical power of replication studies is low, (3) that information needed to conduct a replication with extension is difficult to obtain, (4) that replication research is published outside the major journals (the authors plan to research this issue in the near future), (5) that replications with extensions are considered to be of little importance, (6) that original works are not worthy of replication, and (7) that because of an editorial bias against them, conducting replications with extensions is not career enhancing.

The last point above deserves amplification. It is possible that the amount of replication activity occurring in the business fields is underreported by its publication frequency. If scholars perceive an editorial bias against replication research, such work is unlikely to be written up and submitted for publication. And survey evidence suggests the existence of just such an editorial and reviewer bias (Kerr et al., 1977; Neuliep and Crandall, 1990, 1993).

The usual argument is that limited journal space should be reserved largely for original research. This is how it should be, for distinction is conferred upon originality. It has been shown, however, that the publication of original works in leading business journals is scarcely threatened by the amount of space appropriated by replications and extensions.

It is difficult to determine, in an operational sense, what the optimum proportion of replications and extensions should be in the business literature. Raman's (1994) recently proposed model, despite its apparent mathematical formality, fails to provide an answer to this question. Ultimately, the reader is counseled to use "judgment" in deciding on the optimum number of replications. Neuliep and Crandall (1993) suggest that about 15% of journal space should be reserved for replication research. Clearly, the amount of journal space allocated to such work in the business disciplines falls well below this figure.

Critics from the various business disciplines maintain that current levels of published replication research are insufficient (Dewald et al., 1986; Hubbard and Armstrong, 1994; Kane, 1984; Lindsay and Ehrenberg, 1993). To alleviate this problem, it is recommended that consideration be given to creating a separate section in business journals for replication and extension research. This would signal unequivocally to authors that the business professions are committed to the need for these works.

A "Replications and Extensions" section, which could possibly have its own editor, would not necessarily be featured in each issue of a journal. The papers would be shorter than regular articles, especially those confirming original results. They could be further shortened by using a smaller print for certain portions of the paper (for example, the methods section dealing with samples, instruments, etc.) as is the practice with a number of American Psychological Association journals. Occasionally, the "Replications and Extensions" section might be organized around a particular research theme.

Another benefit from the publication of a replications section would be the generation of sufficient studies to mitigate some of the construct validity and sample-size problems associated with meta-analyses. Replication and meta-analytic research are complementary, rather than competing, scientific practices (Allen and Preiss, 1993). Indeed, a meta-analysis may not be replicable (Bullock and Svyantek, 1985).

The fostering of a replication tradition in the business fields would be expedited by involving graduate students in the process. At present, the value and purpose of replication receives little attention in graduate education. This could be changed. For example, the economist Mayer (1980) proposed that some foundation could sponsor a program allowing graduate students to replicate approximately 10% of the empirical literature published in the previous year. Mayer noted that a policy of this nature would not only catch many mistakes, which Dewald et al. (1986) contend are commonplace in economics, but also would remind authors to be more vigilant in the conduct of their research. Dewald et al. state that since 1983, graduate students in advanced econometrics courses at the Ohio State University are required to replicate and extend a published empirical study. The same is true of econometerics courses at the University of Missouri (Feigenbaum and Levy, 1993). Similar policies of encouraging graduate students to replicate and extend earlier marketing research are supported at the University of Auckland, New Zealand (Hubbard, Brodie, and Armstrong, 1992). Finally, Reid et al. (1981) proposed that more Master's theses and Ph.D. dissertations be awarded for replication research in advertising. The preceding initiatives deserve support. Furthermore, endorsement is needed of the observation that replication research is not mundane; done well, this kind of research can be intellectually challenging and exciting (Lindsay and Ehrenberg, 1993).

Conclusions

Empirical generalization is central to knowledge development and depends heavily on replication and extension research. There is no shortage of empirical research in the business areas. Based on a sample for the period 1980 to 1991, the proportion of empirical work is as follows: accounting (70.7%), economics (57.3%), finance (65.1%), management (91.7%), and marketing (79.5%). Unfortunately, very few of these studies are replications with extensions. Hence, knowledge of business phenomena is restricted.

Those emphasizing the importance of replication research should not be accused of inhibiting creative and original studies. On the contrary, the intent is to stimulate interest in determining whether these original results are robust and capable of being generalized to other contexts. At present, many empirical findings in the business literature are isolated and fragile, as

they have been largely immune from examinations designed to assess their reproducibility and generalizability. This legacy impedes the scientific development of the disciplines and reduces their ability to assist in making important policy decisions. By encouraging systematic replications and extensions, the researcher is allowed the opportunity to synthesize or consolidate what is known and not known about relationships of interest, and thus add to knowledge. As Lindsay and Ehrenberg (1993) point out, finding generalizable results depends on actively looking for them.

Use of the pejorative expression "mere replication" is both unfortunate and unwarranted. It betrays a lack of understanding of the value of replication research. Until the practice of conducting replications and extensions becomes routine, and is afforded the respect commensurate with its role in knowledge development, the integrity of findings in the business literature is potentially compromised.

The authors thank Stuart Allen, Scott Armstrong, Rod Brodie, Andrew Ehrenberg, Bob Leone, Murray Lindsay, Eldon Little, Massoud Saghafi, and two anonymous reviewers for their constructive comments.

Appendix A. Further information on Journal Issues

The *ACCOUNTING* sample included 22 issues each of the *Accounting Review* (*AR*) and the *Journal of Accounting Research* (*JAR*), and 13 issues of the *Journal of Accounting and Economics* (*JAE*). The *ECONOMICS* sample comprised 22 issues each of the *American Economic Review* (*AER*), *Quarterly Journal of Economics* (*QJE*), and the *Review of Economics and Statistics* (*RES*), together with 33 issues of the *Journal of Political Economy* (*JPE*). The *FINANCE* sample consisted of 22 issues each of the *Journal of Finance* (*JF*), *Journal of Financial and Quantitative Analysis* (*JFQA*), and the *Journal of Money, Credit, and Banking* (*JMCB*), whereas for the *Journal of Financial Economics* (*JFE*) it was 18 issues. Twenty-two issues each of the *Academy of Management Journal* (*AMJ*) and the *Administrative Science Quarterly* (*ASQ*), in combination with 30 issues of the *Journal of Applied Psychology* (*JAP*) and 33 issues of *Organizational Behavior and Human Decision Processes* (*OBHDP*), formerly *Organizational Behavior and Human Performance* (*OBHP*), made up *MANAGEMENT'S* contribution. Finally, the *MARKETING* sample was composed of 44 issues each of the *Journal of Marketing* (*JM*) and the *Journal of Marketing Research* (*JMR*), and 37 issues of the *Journal of Consumer Research* (*JCR*).

A listing of these issues was supplied to the editor, and will also be made available to any interested reader. It can be seen that there is some variation in the number of issues selected for certain journals. This is due to factors such as (1) not all journals are published on a quarterly basis (e.g., *JPE* is published six times a year, whereas *JAP* fluctuated between four or six issues per year); (2) taking a 50% random sample from *JAR* because only two issues are published each year; and (3) the original publication dates of some journals, such as *JAE* (1979), *JCR* (1974), and *JFE* (1974), were not early enough to span the full time period, 1970 to 1991, used in this study.

Appendix B. Replications and Extensions in Five Business Disciplines by Journal: 1970–1991

Journal	Number of Empirical Studies	Replications and Extensions	Percentage	Proportion of Empirical Studies Using Secondary Data
Accounting				
Accounting Review	116	9	7.8	0.63
Journal of Accounting and Economics	44	3	6.8	0.99
Journal of Accounting Research	213	20	9.4	0.58
Economics				
American Economic Review	238	43	18.1	0.89
Journal of Political Economy	190	17	8.9	0.97
Quarterly Journal of Economics	93	7	7.5	0.92
Review of Economics and Statistics	459	15	3.3	0.96
Finance				
Journal of Finance	229	38	16.6	0.98
Journal of Financial Economics	69	5	7.2	0.96
Journal of Financial and Quantitative Analysis	119	3	2.5	0.98
Journal of Money, Credit, and Banking	139	8	5.8	0.99
Management				
Academy of Management Journal	278	12	4.3	0.13
Administrative Science Quarterly	109	7	6.4	0.37
Journal of Applied Psychology	564	38	6.7	0.03
Organizational Behavior and Human Decision Processes	271	8	3.0	0.01

Appendix B. Continued

Journal	Number of Empirical Studies	Replications and Extensions	Percentage	Proportion of Empirical Studies Using Secondary Data
Marketing				
Journal of Marketing	270	11	4.1	0.24
Journal of Marketing Research	555	12	2.2	0.21
Journal of Consumer Research	314	10	3.2	0.15

References

Allen, Mike, and Preiss, Raymond, Replication and Meta-Analysis: A Necessary Connection. *J. Soc. Behav. Personality* 8 (1993): 9–20.

Ashenfelter, Orley, Editorial Statement. *Am. Econ. Rev.* 76 (September 1986): v.

Bird, Charles P., and Fisher, Terri D., 30 Years Later: Attitudes Toward the Employment of Older Workers. *J. Appl. Psychol.* 71 (August 1986): 515–517.

Blaug, Mark, *The Methodology of Economics*: Or How Economists Explain, Cambridge University Press, New York. 1992.

Brinberg, David, and McGrath, Joseph E., *Validity and the Research Process*, Sage, Newbury Park, CA. 1985.

Brown, Stephen W., and Coney, Kenneth A., Building a Replication Tradition in Marketing, in *Marketing 1776–1976 and Beyond*, Kenneth L. Bernhardt, ed., American Marketing Association, Chicago, IL. 1976, pp. 622–625.

Bullock, R. J., and Svyantek, Daniel J., Analyzing Meta-Analysis: Potential Problems, an Unsuccessful Replication, and Evaluation Criteria, *J. Appl. Psychol.* 70 (February 1985): 108–115.

Burgstahler, David, Inference from Empirical Research. *Account. Rev.* 62 (January 1987): 203–214.

Cohen, Jacob, A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Measurement* 20 (Spring 1960): 37–46.

Dewald, William G., Thursby, Jerry G., and Anderson, Richard G., Replication in Empirical Economics: *The Journal of Money, Credit, and Banking* Project. *Am. Econ. Rev.* 76 (September 1986): 587–603.

Ehrenberg, Andrew S. C., *Repeat Buying: Facts, Theory, and Applications*, new edition. Charles Griffin, London. 1988.

Ehrenberg, Andrew S. C., and Bound, John A., Predictability and Prediction (with comments). *J. R. Statist. Soc. A.* 156 (1993): 167–206.

Feige, Edgar L., The Consequences of Journal Editorial Policies and a Suggestion for Revision. *J. Polit. Econ.* 83 (December 1975): 1,291–1,295.

Feigenbaum, Susan, and Levy, David M., The Market for (Ir)reproducible Econometrics (with comments). *Soc. Epistemology* 7 (1993): 215–292.

Fishe, Raymond P. H., and Wohar, Mark. The Adjustment of Expectations to a Change in Regime: Comment. *Am. Econ. Rev.* 80 (September 1990): 968–976.

Fleiss, Joseph L., *Statistical Methods for Rates and Proportions*. Wiley, New York. 1981.

Franke, Richard H., and Kaul, James D., The Hawthorne Experiments: First Statistical Interpretations. *Am. Sociol. Rev.* 43 (October 1978): 623–643.

Gorn, Gerald J., The Effects of Music in Advertising on Choice Behavior: A Classical Conditioning Approach. *J. Marketing* 46 (Winter 1982): 94–101.

Greenwald, Anthony G., Consequences of Prejudice Against the Null Hypothesis. *Psychol. Bull.* 82 (January 1975): 1–20.

Guttman, Louis, The Illogic of Statistical Inference for Cumulative Science. *Appl. Stoch. Models Data Anal.* 1 (1985): 3–9.

Haire, Mason, Projective Techniques in Marketing Research. *J. Marketing* 14 (April 1950): 649–656.

Hubbard, Raymond, The Dangers of Generalizing from Published Marketing Studies. *J. Market Res. Soc.* 36 (July 1994): 257–260.

Hubbard, Raymond, and Armstrong, J. Scott, Are Null Results Becoming an Endangered Species in Marketing? *Marketing Letters* 3 (April 1992): 127–136.

Hubbard, Raymond, and Armstrong, J. Scott, Replications and Extensions in Marketing: Rarely Published but Quite Contrary. *Int. J. Res. Marketing* 11 (June 1994): 223–248.

Hubbard, Raymond, Brodie, Roderick J., and Armstrong, J. Scott, Knowledge Development in Marketing: The Role of Replication Research. *N. Zeal. J. Business* 14 (1992): 1–12.

Jiambalvo, James, Measures of Accuracy and Congruence in the Performance Evaluation of CPA Personnel: Replication and Extensions. *J. Account. Res.* 20 (Spring 1982): 152–161.

Kane, Edward J., Why Journal Editors Should Encourage the Replication of Applied Econometric Research. *Q. J. Bus. Econ.* 23 (Winter 1984): 3–8.

Kellaris, James J., and Cox, Anthony D., The Effect of Background Music in Advertising: A Reassessment. *J. Consumer Res.* 16 (June 1989): 113–118.

Kerr, Steven, Tolliver, James, and Petree, Doretta, Manuscript Characteristics Which Influence Acceptance for Management and Social Science Journals. *Acad. Management J.* 20 (March 1977): 131–141.

Kinnear, Thomas C., From the Editor. *J. Marketing* 56 (January 1992): 1–3.

Kirchener, Wayne K., and Dunnette, Marvin D., Attitudes Toward Older Workers. *Personnel Psychol.* 7 (Summer 1954): 257–265.

Lang, Larry H. P., Stultz, Rene M., and Walkling, Ralph A., Managerial Performance, Tobin's *q* and the Gains from Successful Tender Offers. *J. Financial Econ.* 24 (September 1989): 137–154.

Levy, David M., and Feigenbaum, Susan, Testing the Replication Hypothesis When the Data Set Is Subject to Gross Error. *Econ. Letters* 34 (1990): 49–53.

Lindsay, R. Murray, Publication System Biases Associated with the Statistical Testing Paradigm. *Contemp. Account. Res.* (in press).

Lindsay, R. Murray, and Ehrenberg, Andrew S.C., The Design of Replicated Studies. *Am. Statist.* 47 (August 1993): 217–228.

Maher, Michael W., Ramanathan, Kavasseri V., and Peterson, Richard B., Preference Congruence, Information Accuracy, and Employee Performance: A Field Study. *J. Account. Res.* 17 (Autumn 1979): 476–503.

Mankiw, N. Gregory, Miron, Jeffrey A., and Weil, David N., The Adjustment of Expectations to a Change in Regime: A Study of the Founding of the Federal Reserve. *Am. Econ. Rev.* 77 (June 1987): 358–374.

Mayer, Thomas, Economics as a Hard Science: Realistic Goal or Wishful Thinking? *Econ. Inq.* 18 (April 1980): 165–178.

Mazen, Abdelmagid M., Graf, Lee A., Kellog, Calvin E., and Hemmasi, Masoud, Statistical Power in Contemporary Management Research. *Acad. Management J.* 30 (June 1987): 369–380.

McCabe, George, Editor's Note. *Q. J. Bus. Econ.* 23 (Winter 1984): 79.

Miller, Rupert G., *Simultaneous Statistical Inference*. Springer-Verlag, New York. 1981.

Mirowski, Philip, and Sklivas, Steven, Why Econometricians Don't Replicate (Although They Do Reproduce). *Rev. Polit. Econ.* 3 (1991): 146–163.

Monroe, Kent B., Editorial. On Replications in Consumer Research: Part I. *J. Consumer Res.* 19 (June 1992a), 2 pages.

Monroe, Kent B., Editorial. On Replications in Consumer Research: Part II. *J. Consumer Res.* 19 (September 1992b), 2 pages.

Nelder, J. A., Statistics, Science and Technology (with comments). *J. R. Statist. Soc. A.* 149 (1986): 109–121.

Neuliep, James W., and Crandall, Rick, Editorial Bias Against Replication Research. *J. Soc. Behav. Personality* 5 (1990): 85–90.

Neuliep, James W., and Crandall, Rick, Reviewer Bias Against Replication Research. *J. Soc. Behav. Personality* 8 (1993): 22–29.

Poirier, Dale J., Causal Relationships and Replicability. *J. Econometrics* 39 (September/October 1988): 213–234.

Raman, Kalyan, Inductive Inference and Replications: A Bayesian Perspective. *J. Consumer Res.* 20 (March 1994): 633–643.

Reid, Leonard N., Soley, Lawrence C., and Wimmer, Roger D., Replication in Advertising Research: 1977, 1978, 1979. *J. Advertising* 10 (1981): 3–13.

Rosenthal, Robert, The "File Drawer Problem" and Tolerance for Null Results. *Psychol. Bull.* 86 (May 1979): 638–641.

Rosenthal, Robert, and Rosnow, Robert L., *Essentials of Behavioral Research: Methods and Data Analysis*, McGraw-Hill, New York. 1984.

SAS Institute, *SAS/IML User's Guide: Version 6 Edition*. SAS Institute, Carey, NC. 1989.

Servaes, Henri, Tobin's Q and the Gains from Takeovers. *J. Finance* 46 (March 1991): 409–419.

Terpstra, David E., Relationship Between Methodological Rigor and Reported Outcomes in Organizational Development Evaluation Research. *J. Appl. Psychol.* 66 (October 1981): 541–543.

Walster, G. William, and Cleary, T. Anne, A Proposal for a New Editorial Policy in the Social Sciences. *Am. Statist.* 24 (April 1970): 16–19.

Webster, Jr., Frederick E., and von Pechmann, Frederick, A Replication of the "Shopping List" Study. *J. Marketing* 34 (April 1970): 61–63.

Woodman, Richard W., and Wayne, Sandy J., An Investigation of Positive-Findings Bias in Evaluation of Organization Development Interventions. *Acad. Management J.* 28 (December 1985): 889–913.

Wrege, Charles D., and Perroni, Amedeo G., Taylor's Pig-Tale: A Historical Analysis of Frederick W. Taylor's Pig-Iron Experiments. *Acad. Management J.* 17 (March 1974): 6–27.

Zinkhan, George M., Jones, Marilyn Y., Gardial, Sarah, and Cox, Keith K., Methods of Knowledge Development in Marketing and Macromarketing. *J. Macromarketing* 10 (Fall 1990): 3–17.