

**Replications and Extensions in Marketing –
Rarely Published But Quite Contrary^{*}**

Raymond Hubbard
Drake University, Des Moines, IA 50311, USA

J. Scott Armstrong
The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract

Replication is rare in marketing. Of 1,120 papers sampled from three major marketing journals, none were replications. Only 1.8% of the papers were extensions, and they consumed 1.1% of the journal space. On average, these extensions appeared seven years after the original study. The publication rate for such works has been decreasing since the 1970s. Published extensions typically produced results that conflicted with the original studies; of the 20 extensions published, 12 conflicted with the earlier results, and only 3 provided full confirmation. Published replications do not attract as many citations after publication as do the original studies, even when the results fail to support the original studies.

"Replicability is almost universally accepted as the most important criterion of genuine scientific knowledge..."

Rosenthal and Rosnow, 1984, p. 9

"Replicability . . . is the Supreme Court of the scientific system."

Collins, 1985, p. 19

1. Introduction

As suggested by the above quotations, replication is held in high regard by some scientists. Other things being equal, the failure to obtain similar findings in a replication indicates the need for further work in the area. A successful replication, on the other hand, promotes confidence in the reliability of the results, and suggests the need to study whether the findings can be generalized to different populations, products, geographical areas, and so on. Replications

* The authors thank Richard Bagozzi, Gregory Boller, Stephen Brown, Gilbert Churchill, Andrew Ehrenberg, James Engel, Anthony Greenwald, Robert Mittelstaedt, Leonard Reid, Robert Rosenthal, Daniel Vetter, the editor, and three anonymous reviewers for comments on earlier versions of this paper. Daniel Vetter also assisted with the classification of the replications and extensions. Editorial assistance was provided by Jennifer Armstrong and Phan Lam. Any remaining errors are our responsibility.

with extensions serve this function of assessing whether outcomes can be generalized beyond the original context. They help to determine the scope and limits of the findings. Thus, replications and extensions play a valuable role in ensuring the integrity of a discipline's empirical results.

Physical scientists are said to be more concerned about issues of replicability than social scientists (Chase, 1970). The 1989 "discovery" of cold fusion (Close, 1991) provides an example of this concern. The benefits of replication research are further discussed in the next section.

2. Some benefits of replications and extensions

2.1 A guard against the perpetuation of erroneous and questionable results

The publication of replications and extensions helps to protect the literature from the uncritical acceptance and dissemination of erroneous and questionable results. It is unrealistic to expect the peer review system to shoulder this burden alone. The members of a discipline have a collective responsibility to ask whether a given result is plausible, reproducible and/or generalizable.

Replication helps guard the published literature from a proliferation of Type I errors; that is, erroneous rejections of the null hypothesis. Researchers from a variety of disciplines allege that journal policies are biased in favor of the publication of statistically significant ($p < 0.05$) results, thus contaminating their respective literatures with Type I errors whose incidence exceeds that prescribed by nominal alpha levels such as 0.05 (Feige, 1975; Greenwald, 1975; Hubbard and Armstrong, 1992). This is because researchers faced with insignificant results may be tempted to engage in data mining, or "hunting for p values" less than 0.05 (Salsburg, 1985, p. 220), and thus contribute toward Type I error proliferation.

Greenwald (1975) comments that inflated Type I error rates in the social and behavioral sciences raise grave questions about the scientific basis for much of the published literature. Similar conditions may apply to marketing (Hubbard and Armstrong, 1992). Walster and Cleary (1970, p. 17) submit that "only by allowing publication of replications and failures to replicate will Type I errors in the literature be uncovered."

The publication of errors, Type I or otherwise, and questionable results in scholarly journals damages a discipline's reputation. Some of these results find their way into the textbooks, thereby increasing the damage. For once in the textbooks, these results take on a life of their own, are presented as established fact, and are passed on unwittingly in the classroom each term. In the social sciences the results of some "classic" studies were accepted for years even though they were erroneous. Among the more notable examples are J.B. Watson's conditioning of Little Albert, Cyril Burt's "twins" research (Samelson, 1980), and the Hawthorne Effect (Franke and Kaul, 1978).

In the marketing literature, the subliminal advertising study conducted at a New Jersey movie theater in 1956 became a "classic" work that is presented routinely in consumer behavior textbooks. This study, reported in the popular press (see Wilkie, 1986), claimed to show that the

subliminal messages "Hungry? Eat Popcorn" and "Drink Coca-Cola," flashed repeatedly on the screen for 1/3,000 of a second, boosted the sales of popcorn and Coca-Cola by 58% and 18%, respectively. Notwithstanding such "impressive" results, subsequent replication efforts have found no evidence to support the influence of subliminal advertising and persuasion on buyer behavior (DeFleur and Petranoff, 1959; Moore, 1982). Fortunately, disconfirming results like these now appear in the textbooks also.

More recently in marketing, Gorn (1982) published the results of an experiment stating 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 partially due to demand artifacts, a concern heightened by the fact that the paper had become quite influential. Kellaris and Cox observed that the study was cited at least 34 times between 1982 and 1988 (Social Science Citations Index). Perhaps more importantly, they claimed, it was being presented in some 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 three well-designed experiments, Kellaris and Cox failed to replicate Gorn's results. They concluded that single exposure conditioning of product preferences is not "well established."

2.2 Assessment of the generalizability of marketing results

Leone and Schultz (1980) note that replication is the key to generalization. They further remark that while much empirical research takes place in marketing, little is generalizable, and thus we have little knowledge of marketing phenomena. This theme was echoed in a special session at the Marketing Science Conference held in London in 1992. Knowledge generation in the discipline would be aided by replications and extensions of previous studies. Unfortunately, as Jacoby (1978) observes, most of the consumer behavior and marketing literature is composed of unreplicated "single-shot" studies that should not be taken too seriously. Uncorroborated research results, including those with "high" levels of statistical significance, are necessarily tentative or speculative. Kollat et al. (1972), for example, comment that they would not be surprised if 90% of the findings in the consumer behavior literature prove to be wrong. We agree with Ehrenberg (1990) that researchers should focus on whether their results are generalizable across different data sets.

The merits of uncorroborated results should be examined via replications and extensions. As an example of a study whose results have been shown not to generalize thus far, consider Robertson and Bellenger's (1978) work regarding contributions to charitable organizations and mail survey responses. Here the weight of the evidence argues against their initial finding that the "incentive" of promised contributions to charities significantly increases mail survey responses over those of a no-incentive control group. Replications with extensions provided no support for Robertson and Bellenger's original results (Furse and Stewart, 1982; Hubbard and Little, 1988).

Conversely, replications concerning the Double Jeopardy phenomenon (which states that in terms of market share; a small brand typically has far fewer buyers than a larger brand and that its customers tend to buy it less often) have shown it to exhibit remarkable empirical regularity (Ehrenberg et al., 1990). Similarly, Robertson (1971) presents evidence from 21 independent

studies spanning a wide range of product categories and populations from which generalizable characteristics of consumer innovators emerge. Innovators, for example, typically have higher education and income levels, are much more likely to be opinion leaders, and have a more venturesome personality, than non-innovators. A final example of how cumulative empirical research can lead to generalizable results is provided by studies examining the capabilities of spatial-interaction, or gravity, models to predict consumer shopping patterns and the market shares of shopping outlets. The success of repeated empirical tests of these models make them " . . . a powerful research tool and decision aid for sales forecasting and strategy formulation" (Ghosh and McLafferty, 1987, p. 93). Spatial-interaction models now play an important role in the planning of retail facilities in the United Kingdom.

Replication and extension research helps to determine which results are valid, which are reliable, and which do or do not generalize (Campbell and Jackson, 1979). Our study therefore examines the publication of replications in marketing and consumer behavior from 1974 through 1989. First, we present evidence on the publication frequency of replications and extensions in marketing and consumer behavior, including whether this frequency has changed in recent years. Second, we compare the results of original and replicated marketing studies to see if they agree. This has a bearing on the need for replication. Third, we discuss some reasons for the scarcity of replications and extensions in marketing and consumer research. Finally, we speculate on ways to encourage their publication.

3. Frequency of replications and extensions

As a benchmark, we reviewed the literature on the publication of replications in the social sciences. Sterling (1959) did not discover a single replication in his inspection of 362 published psychology articles. Bozarth and Roberts' (1972) analysis of 1,046 psychology papers revealed that fewer than one percent of them were replications. In marketing, Brown and Coney (1976) examined 649 reports appearing in the 1971-1975 issues of the *Journal of Marketing* and *Journal of Marketing Research*, and the 1974-1975 issues of the *Journal of Consumer Research*. They concluded that only two percent of the articles were replications or extensions. Reid et al. (1981) found that of 501 advertising articles published in major marketing and communications journals and proceedings during the years 1977 through 1979, six percent qualified as replications or extensions.

We examined replication frequency in the marketing literature. To ensure comparability with earlier studies, the definitions of "replication" and "replication with extension" are based on those used by Reid et al. (1981, p. 7), who, in turn, had used a somewhat revised version of Brown and Coney's (1976, p. 622) approach. We defined a replication as a duplication of a previously published empirical study that is concerned with assessing whether similar findings can be obtained upon repeating the study. This definition covers what are variously referred to as "exact," "straight," or "direct" replications. Such works duplicate as closely as possible the research design used in the original study by employing the same variable definitions, settings, measurement instruments, analytical techniques, and so on. An example would be repeating the study with another sample drawn from the same population.

A *replication with extension* is a duplication of a previously published empirical research project that serves to investigate the generalizability of earlier research findings. The extension does not alter the conceptual relationships involved in the original study, but instead tests them by making changes in some aspects of the initial design. Examples would be modifying either the manipulated (Kellaris and Cox, 1989) or measured (Peat et al., 1975) variables, but not both, analyzing the impact of an additional variable or variables (Furse and Stewart, 1982), and repeating the study with a sample drawn from different populations whether these be in terms of gender (Zinkhan and Shermohamad, 1986), socioeconomic background (Bearden et al., 1979), special versus general consumer groups (Furse and Stewart, 1982), different geographical settings within the US (Shuptrine and Samuelson, 1976), and between the US and other countries (Clarke and Soutar, 1982). Also included in this definition are studies that wish to consider explicitly whether the passage of time has had an impact on earlier results (Nagashima, 1977), as well as those that have gathered data with different methods (Cunningham and Green, 1974), and those that examined consumer behavior in varied situations (Foxall, 1975).¹ These definitions are consistent with those of Mittelstaedt and Zorn (1984), and with Brinberg and McGrath's (1985) ideas about validity and the robustness of results.

While we regard it as important, no attempt was made to determine whether an author had replicated his or her own work by means of cross validation, performing a second experiment, etc. within the context of the *original* article. This is congruent with the previous assessments of replications. Because we are concerned with the willingness of journals to publish replications as separate or "independent" contributions, we did include in our counts instances where authors published replications of their own work as separate articles. This procedure affected three of the articles listed in our appendix (Blasko and Patti, 1984; Nagashima, 1977; and Solomon et al., 1976).

Thus, the major focus of our paper concerns replications and extensions carried out by an independent researcher. These replications are especially convincing because they help to avoid biases such as experimenter expectations that might have been associated with the original study. Rosenthal and Rosnow (1984) refer to this as the problem of "correlated replicators," and question the independence and value of ten replications conducted by a single investigator versus ten replications each conducted by a different investigator.

We estimated the incidence of replications and replications with extensions published in the *Journal of Marketing (JM)*, *Journal of Marketing Research (JMR)*, and *Journal of Consumer Research (JCR)* over the period 1974-1989 by content-analyzing 31 randomly selected issues from each. We selected 1974 as the starting year because it was the first year in which the *Journal of Consumer Research* was published. This 50% sampling of all *JM*, *JMR*, and *JCR* issues produced a total of 1,120 research papers (articles, notes, and commentaries), 835 of which were empirical.

¹ It is important to emphasize that the citations in this paragraph come from the listing of replications with extensions contained in the appendix of this paper. Their use above illustrates the kinds of modifications in research designs covered by our definition. It does not exhaust these modifications however; typically, the extensions in the appendix incorporate more than one.

The first author examined and classified each paper. Following Reid et al. (1981), an article was not classified as a replication or extension unless there was an explicit citation in the article to the original study. This does not mean, however, that the replicating authors had to formally identify their work as being a replication or extension. That was our responsibility. If uncertainty arose about whether a study matched the definitions, we included it. To the extent that misclassifications might have occurred, we expected this procedure to overestimate the amount of replication. An inspection of a ten percent random sample of these papers by a colleague, Daniel Vetter, supported the classifications.²

Table 1
Replications with extensions

Journal	1974-1979			1980-1989			1974-1989		
	Number of studies	Replications with extensions	Percentage	Number of studies	Replications with extensions	Percentage	Number of studies	Replications with extensions	Percentage
<i>JM</i>	179 (95)	4	2.2 (4.2)	183 (112)	3	1.6 (2.7)	362 (207)	7	1.9 (3.4)
<i>JMR</i>	214 (176)	5	2.3 (2.8)	206 (190)	2	1.0 (1.1)	420 (366)	7	1.7 (1.9)
<i>JCR</i>	111 (79)	2	1.8 (2.5)	227 (183)	4	1.8 (2.2)	338 (262)	6	1.8 (2.3)
Totals	504 (350)	11	2.2 (3.1)	616 (485)	9	1.5 (1.9)	1120 (835)	20	1.8 (2.4)

3.1 Results on the frequency of published replications and extensions

None of the 835 empirical papers were straight replications. Furthermore, only 20 extensions were found in our sample. Thus, only 1.8% of all papers and 2.4% of empirical papers were extensions (see Table 1).

Between the time-periods 1974-1979 and 1980-1989, there was a reduction from 2.2% to 1.5% in all research papers devoted to extensions. Comparable findings occurred when we calculated extensions as a percentage of empirical studies; the proportion of extensions decreased from 3.1% to 1.9% (see Table 1). The concerns expressed in the early 1970s about the lack of replications and extensions have not led to obvious improvements.

To be most effective, replications and extensions should appear soon-after the original publication. Timely publication of a replication is especially important when it yields results that conflict with the earlier work. Conflicting results will be of interest to those who had planned to use the original findings in research or in decision making. Unfortunately, the marketing extensions suffered from extensive time lags. In our study, the lags ranged from less than one

² They agreed on the classification of 90% of the studies involved in this paper. In two later studies investigating the publication frequency of replications and extensions in economics and finance, the same two coders as used in this paper showed high levels of agreement, measured by Cohen's (1960) kappa, of 0.85 in one study and 0.88 in the other.

year to 27 years, with an average of 7.4 years. The average time lag for the conflicting extensions was 9.1 years, while for the confirming studies it was 5.1 years.

4. Do replications and extensions confirm original results?

If replications almost always supported the initial results, one could argue that there is little need for them. Hubbard and Vetter (1991) examined the success of replication attempts in 55 replications with extensions in finance. Twenty percent of the studies supported the original results, 20% provided partial support, and 60% conflicted. The success of replication attempts in advertising was studied by Reid et al. (1981). Of the 30 replications with extensions that they investigated, 40% supported the original result, 20% provided partial support, and 40% produced conflicting results.

We examined the 20 extensions from our study to determine whether they produced similar results. We based our analysis on the conclusions reached by the authors of the replications.³ These published extensions typically conflicted with the original findings. Of the 20 replications, 15% (three) confirmed earlier results, 25% (five) provided some support, and 60% (12) conflicted with their predecessors.

Replications may fail to yield the same results as the original study for a variety of reasons. Possible explanations could include chance variation and researcher bias. Another possibility is that mistakes were made in such things as the coding, recording, and analyzing of data. For example, Dewald et al.'s (1986) involvement with the *Journal of Money, Credit and Banking* Data Storage and Evaluation Project enabled them to reanalyze the raw data sets accompanying manuscripts submitted ~ to that journal. Rarely, they said, were these re-analyses straightforward; in one case they were unable to obtain the same results even with the active assistance of the original author. Dewald et al. (1986) concluded that inadvertent errors in published empirical articles are common. The above explanations can apply either to the original study or to the replication.

In addition, the initial study's conclusions may not hold up over time, or may fail to generalize to other locations, situations, populations, and so on. This was the case with Mason Haire's (1950) classic "shopping-list" study, conducted in the United States, purporting to demonstrate the usefulness of projective techniques in marketing contexts. A subsequent replication performed in the United States by Webster and von Pechmann (1970) yielded contradictory results. Moreover, Haire's results could not be reproduced in studies undertaken in Canada (Lane and Watson, 1975) and Norway (Arndt, 1973).

³ Half of these conclusions were based on direct statistical comparisons between the original and replicated data sets. The other half rested on the outcomes of conflicting levels of statistical significance reported in the original and replicated papers, a procedure that some researchers view as inadequate for determining the success or failure of a replication attempt (Rosnow and Rosenthal, 1989).

Also, editors may prefer to publish conflicting, rather than confirming, results because these may be viewed as providing "new" evidence on the- topic. This, in turn, might make researchers more likely to submit papers contradicting earlier results.

5. Complements to replications: Commentaries and meta-analyses

The American Marketing Association's (AMA) Task Force (1988, p. 7), concerned with the development, dissemination, and utilization of marketing knowledge, recommended that "the role of `commentaries' be substantially increased in both the *Journal of Marketing Research* and the *Journal of Marketing*." Research commentary on empirical studies can provide for a rapid evaluation of a paper. Commentaries might also help to determine whether a consensus among scholars emerges regarding the merits of an article. In some journals (for example, *Behavioral and Brain Sciences*) commentary appears along with the original paper, a strategy Churchill (1988) endorses for marketing journals and one that used to be followed by *JCR*, or else is published shortly thereafter (as, for example, in the *American Psychologist*). While timely, commentaries and rejoinders offer a less powerful check than is provided by replication. This is because commentaries typically examine the interpretation of studies, but rarely provide an analysis of data.

Our analysis of the frequency of published comments and rejoinders revealed a downward trend over time. Combining the results for all three journals we found that the proportion of published comments and rejoinders declined by about one-third between 1974-1979 (12.9%) and 1980-1989 (8.4%).

Meta-analysis assesses the generalizability of a discipline's empirical findings. Meta-analysis calls for the application of statistical procedures to a collection of studies on a particular topic, in order to objectively summarize the knowledge. Meta-analysis is not without its limitations (Bangert-Drowns, 1986; Bullock and Svyantek, 1985). One drawback is the "apples and oranges" problem, which questions the value of metaanalyses in assessing construct validity. Because the studies included in a typical meta-analysis employ only "roughly similar procedures," a problem arises in averaging effects across independent and dependent variables that have been measured quite differently (Bangert-Drowns 1986, p. 388). Otherwise expressed, it is not always clear that the included studies accurately reflect the same underlying constructs or relationships.

We determined the frequency of meta-analyses published in the marketing literature by conducting a census of all research papers appearing in *JM*, *JMR*, and *JCR* for the ten-year period 1980 through 1989. This inspection of 120 volumes yielded a total of 1,319 research papers, nine (0.7%) of which were meta-analyses of a subject-specific problem.

Meta-analysis and replication are complementary, rather than competing, scientific practices. Replications and extensions can be incorporated into the meta-analyses.

6. Why the scarcity of replications and extensions?

Below, we provide a non-exhaustive list of possible reasons for the scarcity of replications and extensions found in our study. They probably apply also to other management and social science areas.

6.1 Misinterpreting statistical significance as a measure of replicability

Misinterpretations of null hypothesis testing procedures may have undermined the perceived need for replication. Tversky and Kahneman (1971, p. 109), for example, reported that psychologists have "unreasonably high expectations" about the replicability of statistically significant results. Oakes (1986) showed that 42 of 70 (60%) experienced academic psychologists believed that an experimental outcome that is significant at the 0.01 level has a 0.99 probability of being statistically significant if the study were replicated.

If the complements of p values are misinterpreted as a direct index of replicability by psychologists, might not marketers make similar misinterpretations? Given the preponderance of statistically significant results in the marketing literature (Hubbard and Armstrong, 1992), some researchers may dismiss the need for replication.

6.2 Statistical power of replicated studies is low

Sawyer and Ball (1981) question whether replication attempts might be underpowered, and thus less deserving of publication on statistical grounds. (The power of a statistical test is the probability of rejecting a false null hypothesis.) We therefore calculated the statistical power levels for our sample of published extensions, and compared them with those of the original works. We followed standard procedures for conducting a power analysis (Cohen, 1988) by using the article as the primary unit of investigation, adopting two-tailed tests with $\alpha = 0.05$, including only major statistical tests, and employing Cohen's definitions of small, medium, and large effect sizes.

Of the 20 extensions; five could not be power-analyzed because they did not use significance tests, provided insufficient information to calculate power levels, or employed techniques like factor analysis for which power tests do not exist. Six of the 20 original studies were excluded for the same reasons. Thus, 15 extensions and 14 original studies were power-analyzed.

Cohen (1988) has advocated that when $\alpha = 0.05$, and there is no other basis for selecting a power level, a value of 0.80 be used. Adopting this criterion, the average power of both the 14 original and 15 replicated studies was high. The mean power to detect small, medium, and large effect sizes in the original articles was 0.39, 0.90, and 0.96; for the replications the corresponding values were 0.36, 0.90, and 0.99. When we made power comparisons based on the individual number of statistical significance tests used in the original articles (92), the average power to distinguish small, medium, and large effects was 0.41, 0.94, and 0.98, respectively, while for the replications (102) these figures were 0.38, 0.91, and 0.99.

If we assume that medium effect sizes characterize the marketing literature, all but 2 of the 15 replications exceeded the recommended power benchmark of 0.80. All 15 did so if large effect sizes in the population are posited. Conversely, one of the 14 original articles failed to meet the 80% chance of detecting a medium effect; it had only a 50% chance of identifying a large effect.

Of course, it might be argued that one reason these extensions were published is because of their generally high levels of statistical power, the implication being that *unpublished* replications are noticeably underpowered. It was not possible to assess directly the merits of this argument because, to the best of our knowledge, the power levels of published and unpublished replications have never been compared.

6.3 Information is difficult to obtain

The original paper may not report enough detailed information about measurement instruments, sampling methods; analytical techniques, etc., to permit an accurate replication and extension. Replication attempts that rely on information published in the original article may therefore be able to only approximate some of the procedures involved in that research. Any discrepancies in these procedures could be sufficient to produce conflicting results between the two studies.

To obtain the information required for replication, the researchers may need to contact the original authors. How is this request handled? Madden et al. (1979) surveyed, by mail, the authors of 60 papers selected from the 1975 through 1977 AMA and Association for Consumer Research annual conference proceedings. They asked these authors if they would be willing to share instruments and other details of their research to permit replication. Madden et al. emphasized that no request was being made for the actual raw data. Under these conditions, half of the authors contacted failed to provide any information.

Reid et al. (1982), in a mail survey of authors of 99 empirical papers appearing in the 1978 and 1979 issues of *JCR*, *JM*, *JMR*, *Journal of Advertising*, and *Journal of Advertising Research*, requested information necessary for replication. Half of the authors (49) said that the materials were available; some of these authors sent data. Reid et al. did not indicate whether the submitted materials were sufficient to enable replications. Eleven percent of the authors answered that the requested materials had been destroyed or discarded, while 3% said they were proprietary. The remaining 36% of the researchers did not respond to the request, even after a follow-up mailing.

6.4 Replications are published outside the major journals

It could be argued that space constraints in the major journals have relegated the publication of replications to "lower tier" journals and conference proceedings. This is possible, although we know of no empirical study that has addressed this topic.

We do know, however, that major journals in related disciplines publish more replications than those in marketing. For example, we examined 25 randomly selected issues of *the*

American Economic Review (a 25% probability sample) for the period 1965-1989. There were 698 articles, 266 of which were empirical studies. Of these empirical studies, 21.1% were extensions. This is substantially higher than the 2.4% average for the three leading marketing journals included in this paper. Hubbard and Vetter (1991), using the same definitions of replications and extensions as those in the present study, demonstrated that prestigious journals in the finance discipline publish more of these works than do marketing journals. They analyzed a 25% random sample of published issues of the *Journal of Finance*, *Journal of Financial Economics*, *Journal of Financial and Quantitative Analysis*, and the *Journal of Money, Credit and Banking*, for the period 1969-1989. This involved the inspection of 1,028 research papers, of which 555 were empirical. Fifty-five replications with extensions were found, constituting 5.4% of all papers and 9.9% of empirical papers; these exceed the marketing rates of 1.8% and 2.4%, respectively.

A possible explanation for differences in the publication incidence of replications and extensions in marketing vis-à-vis economics and finance could be due to time and cost factors. The vast majority of the economics and finance replications employ secondary data, such as readily accessible census materials and other government statistics, CRSP (Center for Research in Security Prices) tapes for financial data bases, and so on. This could make replication in these disciplines relatively fast and less expensive. In contrast, only 4 of the 20 marketing extensions used secondary data, and even here all 4 involved extremely labor-intensive content analyses of either television or print advertisements. The 16 remaining marketing replications were based either on primary survey data (12) or experimental/observation methods (4). Thus, time, effort and cost considerations may inhibit replication efforts in marketing when compared with related disciplines.

6.5 Replications and extensions are of little importance

Some editors and researchers may question the importance or value of replications and extensions. One argument is that if the original study is well designed there is little need to replicate it. Alternatively, if the study was poorly designed, a replication would be inappropriate. Furthermore, they may point out that an exact replication of a previous study cannot detect whether the manipulations and dependent measures used were construct valid. Even those things that a straight replication might show, such as data analysis errors or possible experimenter effects in the original study, may nevertheless leave the reader with more questions than answers. Suppose, for example, that Smith replicates Jones' previous work and fails to obtain the same results. This could lead to Smith and Jones simply trading accusations as to who is responsible for the different outcomes; Jones could claim that Smith failed to conduct an accurate replication, made recording errors, introduced bias, and so on. Smith could counter that Jones made mistakes.

The problems involved in deciding whether the "burden of proof" lies with Smith or Jones (or both) for a failure to replicate might be expected to compound when an extension is undertaken. Besides those discussed above, other reasons for obtaining conflicting results could include differences due to modifications in manipulated or measured variables, the incorporation of new constructs, the non-generalizability of results to other populations, time periods, situa-

tions, etc. Editors, reviewers, and authors may feel that little is to be gained in attempting to resolve the issue(s).

6.6 Original works are not worth replicating

The works discussed in this paper that have fostered replication attempts, such as the studies on subliminal advertising, shopping lists, charitable contributions and mail survey responses, the Double Jeopardy phenomenon, and the effects of music on marketplace behavior, often did so because the original research yielded surprising and/or controversial results. Some became "classics." And all of them address topics that have potentially significant implications for marketing theory and practice. Obviously, few articles can be expected to have this kind of impact.

Journal editorial policies rightly emphasize the publication of original research that makes important contributions to the literature. Nonetheless, it is argued that the existing, academic research system, which encourages a "publish or perish" mentality, has led to conditions under which a significant contribution to knowledge may not be at the forefront of most participants' thoughts as they engage in the research and publication process" (AMA Task Force, 1988, p. 6). The overriding need to publish might be expected to divert attention from the empirical investigation of important marketing issues to those that are more likely to end up in print. Churchill (1988) maintains that this is the case, and points out that academicians seem to be more interested in applying increasingly sophisticated methodologies to increasingly less important problems. Perhaps as a result, members of the discipline may feel that much of the empirical work published in marketing is simply not important, valuable, or interesting enough to justify the time and costs that replications and extensions entail.

Armstrong and Hubbard (1992) asked marketing academicians and practitioners to rate the importance of 20 empirical studies on consumer behavior. When simultaneously considering the criteria of importance, surprising hypotheses, and statistical significance, four (20%) of the studies were rated as important. Their survey of reviewers on the editorial board of the *JCR* found that about 350 of studies published in that journal were believed to be important. This would indicate that there are sufficient papers in the literature worthy of replication and extension.

6.7 Conducting replications is not career-enhancing

Empirical evidence suggests that editorial review may discourage the publication of replications. Kerr et al. (1977, p. 138, 140) surveyed 429 editors and review board members of 19 management and social science journals. They found a bias against the publication of direct replications even when it is acknowledged that the replications had been done competently. Roney and Zenisek (1980) administered a version of the Kerr et al. survey to 268 reviewers for Canadian psychology journals and obtained similar results. In a survey of 288 past and present editors of social and behavioral science journals, Neuliep and Crandall (1990) also concluded that there was strong evidence of a bias against publishing replications and extensions.

While science may be well served by replications of important work, the individual researcher may be better off pursuing new avenues. Consider the researcher's viewpoint. If Smith's replication supports Jones' prior study, the view is that Jones did a good job. If the results by Smith refute Jones, Smith may be regarded as a trouble-maker; he is either unable to do the replication successfully, or else he is taking a negative attitude toward the research by Jones. This may not be perceived as a good way to advance in the profession. Kane (1984) agrees, adding that choosing to perform replications is likely to be interpreted as prima facie evidence of intellectual mediocrity, a lack of creativity, and perhaps a bullying spirit. It has been suggested that the publication of replications is unlikely to earn recognition for those involved (Brown and Coney, 1976; Mittelstaedt and Zorn, 1984).

We examined citation rates as a measure of the value of replications. For this, we used annual issues of the *Social Science Citation Index (SSCI)*. One view is that citations to an original study would be lower for a failed replication because the results have been called into question. By this reasoning, it would now be incomplete to refer only to the original study, whereas it might be sufficient to refer only to the replication because the replication describes the original study also.

Of the 18 extensions in our sample that permitted an evaluation,⁴ the original studies were cited an average of 2.1 times per year, compared with the extensions' citation rate of 0.9. Moreover, the citation rates of the extensions were lower for 15 of the 18 comparisons and tied in two cases.

The above counts for the original studies were started at the publication date for the extension. Based on their own publication dates, the original studies were cited an average of 1.5 times per year. Thus, their citation rates increased after the publication of the replications.

Extensions with contradictory results might be expected to attract particular attention. Nevertheless, the ten extensions with contradictory results were cited 0.8 times per year, while those with partial or full supporting evidence ($n = 8$) had a citation rate of 1.0.

7. Encouraging replications

Below we present certain suggestions aimed at encouraging replication and extension research in marketing. They are not comprehensive.

7.1 Modify journal editorial policies

Modifications in journal editorial policies can assist authors in replication attempts. Greenwald (1976), as editor of the *Journal of Personality and Social Psychology*, instituted a policy requiring assurance from authors that they would retain their raw data and copies of the details of the procedures used in their research for five years following the date of publication. This policy, originally controversial, is currently endorsed by all American Psychological

⁴ No date was provided for one study. while another was published in 1989.

Association journals. The American Marketing Association recently adopted a similar policy for its journals. It is expected that authors will note the availability of instruments and other materials and that they will assist in replication efforts. This policy is summarized in a January 1992 editorial in the *JM*. Likewise, Monroe's (1991, preface) editorial in the *JCR* stated that "... research that replicates and extends previous findings is ... a necessary ingredient for the advancement of consumer research and is acceptable to *JCR*."

Clearly communicated changes in journal editorial policies to facilitate the publication of replications and extensions can make a difference. For example, in their Winter 1984 issue, the editors of the *Quarterly Journal of Business and Economics (QJBE)* expressed an obligation to publish replications and extensions of papers, as long as these were conducted independently. We performed a content analysis of a 50% probability sample of all issues of *QJBE* published between 1978 and 1989. Of the 59 papers published before the 1984 change in editorial policy, none were replications. In contrast, 9 of the 77 papers (11.7%) sampled after the announcement were replications with extensions; this represented 17.3% of the empirical studies.

7.2 Appoint a replications editor

Brown and Gaulden, Jr. (1980) suggested that marketing journals should have a separate section for replications. We agree, and also recommend that the editorial boards of these journals consider appointing a *replications editor*, who would actively solicit such works. A replications editor would be especially interested in publishing replications and extensions of important papers. While we acknowledge that the criterion of "importance" is subjective, a useful point of departure in clarifying this issue would be to use the U.S. Marketing Science Institute's priority listing of topics for research. Additional guidelines as to which empirical marketing works are considered to be important could be decided by using criteria such as SSCI citations of articles, by investigating what topics are covered in marketing anthologies, and by surveys of marketing academicians and practitioners (as was done in Armstrong and Hubbard, 1992).

This editor would also be interested in the replication of previous studies that have yielded interesting, surprising, dubious, and/or controversial results. Some of these studies are also likely to qualify as being important by the criteria listed in the above paragraph.

The replications editor would also be responsible for developing guidelines concerning the reporting of such works. For example, extensions that confirm earlier findings presumably would be shorter pieces than those contradicting the original paper.

7.3 Other ways to encourage replication

Replication might also be encouraged in marketing's methodological literature. In a convenience sample of 14 marketing research texts published between 1986 and 1989, only one included a reference to the topic in the index. Similarly, only one of 15 consumer behavior textbooks indexed "replication."

Finally, graduate students need to be informed about the importance of conducting replications and extensions. In economics, Mayer (1980) suggests that funding be sought for

programs that encourage graduate students to replicate a sample of studies published in the previous year or so. Graduate students in advanced econometrics courses at Ohio State University are required to replicate and extend a published study (Dewald et al., 1986). In marketing, Reid et al. (1981) have argued in favor of granting more Master's theses and Ph.D. dissertations for replication and extension work. This policy is actively pursued in the Master's degree program in marketing at the University of Auckland, New Zealand.

8. Conclusions

By helping to assess the validity, reliability, and generalizability of empirical findings, replication is considered to be important for the advancement of science. Mittelstaedt and Zorn (1984, p. 14) suggest "That which isn't worth replicating isn't worth knowing." Despite this, no replications and only 20 extensions were published in the sample of 835 empirical papers from three major marketing journals. Sixty percent of these extensions produced results that conflicted with the original findings. We do not know what the optimum proportion of published replications and extensions ought to be. However, the current proportion seems low in an absolute sense and when judged against other areas of management science. Of course, our results may be biased because they are based on US data; it is possible that journals in other countries publish more replications and extensions than we have found.

The development of a replication tradition would enhance marketing's scientific status. But there must also be additional emphasis on theory construction in marketing. At present, there is an overdependence on the ability of uncorroborated, exploratory-empirical studies to sustain the discipline. Greater attention needs to be focused on theory-driven programmatic research, that is, research that systematically pursues a given topic area (Anderson, 1983; Jacoby, 1978).

We do not view the publication of independent replications and extensions as a panacea for knowledge development in marketing. Neither do we lay sole blame on journal editorial policies for the lack of such works. The extent to which perceived editorial policy influences the behavior of researchers is unknown. Suffice it to say that editors, reviewers, and authors alike share a joint responsibility for the dearth of replications found in print. We do believe, however, that the marketing discipline would benefit from changes in editorial policies that would encourage the publication of replications and extensions, and we have offered suggestions to accomplish this.

Appendix: List of twenty replications with extensions

- Aaker, D.A. and G.T. Ford (1983), "Unit pricing ten years later: A replication," *Journal of Marketing*, 47, 118-122.
- Bearden. W.O., J.E. Teel and R.R. Wright (1979), "Family income effects on measurement of children's attitudes toward television commercials," *Journal of Consumer Research*, 6, 308-311.
- Belkaoui, A. and J.M. Belkaoui (1976), "A comparative analysis of the roles portrayed by women in print advertisements: 1958, 1970, 1972," *Journal of Marketing Research*, 13, 168-172.
- Blasko, V.J. and C.H. Patti (1984), "The advertising budgeting practices of industrial marketers," *Journal of Marketing*, 48, 104-110.
- Bush, R.F., J.F. Hair, Jr. and P.J. Solomon (1979), "Consumers' level of prejudice and response to black models in advertisements," *Journal of Marketing Research*, 16; 341-345.
- Clarke, Y. and G.N. Soutar (1982), "Consumer acquisition patterns for durable goods: Australian evidence," *Journal of Consumer Research*, 8, 456-460.
- Cunningham, I.C.M. and R.T. Green (1974), "Purchasing roles in the U.S. family, 1955 and 1973," *Journal of Marketing*, 38, 61-81.
- Dowling. G.R. (1980), "Information content in U.S. and Australian television advertising," *Journal of Marketing*, 44, 34-37.
- Foxall, G.R. (1975), "Social factors in consumer choice: Replication and extension," *Journal of Consumer Research*, 2, 60-64.
- Furse, D.H. and D.W. Stewart (1982), "Monetary incentive versus promised contribution to charity: New evidence on mail survey response," *Journal of Marketing Research*, 19, 375-380.
- Harris, J.R. and H.T. Guffey, Jr. (1978), "Questionnaire returns: Stamps versus business reply envelopes revisited," *Journal of Marketing Research*, 15, 290-293.
- Kellaris. J.J. and A.D (Cox), "1989. The effect of background music in advertising: A reassessment," *Journal of Consumer Research*, 16, 113- 118.
- Michaels. R.E. and R.L. Day (1985), "Measuring customer orientation of salespeople: A replication with industrial buyers," *Journal of Marketing Research*, 22, 443-446.
- Nagashima, A. (1977), "A comparative 'made in' product image survey among Japanese businessmen," *Journal of Marketing*, 41, 95-100.

- Peat, N.C., J.W. Gentry and T.L. Brown (1975), "A comment on 'identifying buyers of a major automatic innovation,'" *Journal of Marketing*, 39, 61-62.
- Schneider, K.C. and S.B. Schneider (1979), "Trends in sex roles in television commercials," *Journal of Marketing*, 43, 79-84.
- Shuptrine, F.K. and G. Samuelson (1976), "Dimensions of marital roles in consumer decision making: Revisited," *Journal of Marketing Research*, 13, 87-91.
- Solomon, P.J., R.F. Bush and J.F. Hair. Jr. (1976), "White and black consumer sales response to black models," *Journal of Marketing Research*, 13, 431-434.
- Swasy. J.L. and J.M. Munch (1985), Examining the target of receiver elaborations: Rhetorical question effects on source processing and persuasion," *Journal of Consumer Research*, 11, 877-886.
- Zinkhan, G.M. and A. Shermohamad (1986). "Is other-directedness on the increase? An empirical test of Riesman's theory of social character," *Journal of Consumer Research*, 13, 127-130.

References

- AMA Task Force on the Development of Marketing Thought (1988), "Developing, disseminating, and utilizing marketing knowledge," *Journal of Marketing*, 52, 1-25.
- Anderson, P.F. (1983), "Marketing, scientific progress, and scientific method," *Journal of Marketing*, 47, 18-31.
- Armstrong, J.S. and R. Hubbard (1992), "Importance of empirical research on consumer behavior: Expert opinions," Working Paper, Marketing Department, The Wharton School, University of Pennsylvania.
- Arndt, J. (1973), "Haire's shopping list revisited," *Journal of Advertising Research*, 13, 57-61.
- Bangert-Drowns, R.L. (1986), "Review of developments in meta-analytic method," *Psychological Bulletin*, 99, 388-399.
- Bozarth, J.D. and R.R. Roberts, Jr. (1972), "Signifying significant significance," *American Psychologist*, 27, 774-775.
- Brinberg, D. and J.E. McGrath (1985), *Validity and the Research Process*. Beverly Hills, CA: Sage Publications.
- Brown, S.W. and K.A. Coney (1976), "Building a replication tradition in marketing," in: K.L., Bernhardt. (ed.). *Marketing 1776-1976 and Beyond*, Chicago, IL: American Marketing Association, pp. 622-625.
- Brown, S.W. and C.F. Gaulden, Jr. (1980), "Replication and theory development," in C.W. Lamb, Jr. and P.M. Dunne (eds.), *Theoretical Developments in Marketing*, Chicago, IL: American Marketing Association, pp. 240-243.
- Bullock, R.J.. and D.J. S,vyantek (1985), "Analyzing meta-analysis: Potential problems, an unsuccessful replication, and evaluation criteria," *Journal of Applied Psychology*, 70, 108-115.
- Campbell, K.E. and T.T. Jackson (1979), "The role of and need for replication research in social psychology," *Replications in Social Psychology*, 1, 3-14.
- Chase, J.E. (1971), "Normative criteria for scientific publications," *American Sociologist*, 5, 262-265.
- Churchill, G.F. (1988), "Comments on the AMA Task Force study," *Journal of Marketing*, 52, 26-31.
- Close, F. (1991), *Too Hot to Handle: The Race for Cold Fusion*. Princeton. NJ: Princeton University Press.

- Cohen, J. (1960), "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, 20, 37-46.
- Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Collins, H.M. (1985), *Changing Order: Replication and Induction in Scientific Practice*. Beverly Hills, CA: Sage Publications.
- DeFleur, M.L. and R.M. Petranoff (1959), "A television test of subliminal persuasion," *Public Opinion Quarterly*, 23, 170-180.
- Dewald, W.G., J.G. Thursby and R.G. Anderson (1986), "Replication in empirical economics: The *Journal of Money, Credit and Banking* project." *American Economic Review*, 76, 587-603.
- Ehrenberg, A.S.C. (1990), "A hope for the future of statistics: MSOD," *American Statistician*, 44, 195-196.
- Ehrenberg, A.S.C., G.J. Goodhardt and T.P. Barwise (1990), "Double jeopardy revisited," *Journal of Marketing*, 54, 82-91.
- Feige, E.L. (1975), "The consequences of journal editorial policies and a suggestion for revision," *Journal of Political Economy*, 83, 1291-1296.
- Franke, R.H. and J.D. Kaul (1978), "The Hawthorne experiments: First statistical interpretations," *American Sociological Review*, 4, 623-643.
- Furse, D.H. and D.W. Stewart (1982), "Monetary incentive versus promised contribution to charity: New evidence on mail survey response," *Journal of Marketing Research*, 19, 375-380.
- Ghosh, A. and S.L. McLafferty (1987), *Location Strategies for Retail and Service Firms*. Lexington, MA: D.C. Heath and Company.
- Gorn, G.J. (1982), "The effects of music in advertising on choice behavior: A classical conditioning approach," *Journal of Marketing*, 46, 94-101.
- Greenwald, A.G. (1975), "Consequences of prejudice against the null hypothesis," *Psychological Bulletin*, 82, 1-20.
- Greenwald, A.G. (1976), "An editorial," *Journal of Personality and Social Psychology*, 33, 1-7.
- Haire, M. (1950), "Projective techniques in marketing research," *Journal of Marketing*, 14, 649-656.

- Hubbard, R. and J.S. Armstrong (1992), "Are null results becoming an endangered species in marketing?" *Marketing Letters*, 3, 127-136.
- Hubbard, R. and E.L. Little (1988), "Promised contributions to charity and mail survey responses: Replication with extension," *Public Opinion Quarterly*, 52, 223-230.
- Hubbard, R. and D.E. Vetter (1991), "Replication in the finance literature: An empirical study," *Quarterly Journal of Business and Economics*, 30, 70-81.
- Jacoby, J. (1978), "Consumer research: A state of the art review," *Journal of Marketing*, 42, 87-96.
- Kane, E.J. (1984), "Why journal editors should encourage the replication of applied econometric research," *Quarterly Journal of Business and Economics*, 23, 3-8.
- Kellaris, J.J. and A.D. Cox (1989), "The effect of background music in advertising: A reassessment," *Journal of Consumer Research*, 16, 113-118.
- Kerr, S., J. Tolliver and D. Petree (1977), "Manuscript characteristics which influence acceptance for management and social science journals," *Academy of Management Journal*, 20, 132-141.
- Kollat, D.T., R.D. Blackwell, and J.F. Engel (1972), "The current status of consumer behavior research: Development during the 1968-1972 period," in M. Venkatesan (ed), *Proceedings of the Third Conference of the Association for Consumer Research*. Chicago, IL: Association for Consumer Research, pp. 576-585.
- Lane, G.S. and G.L. Watson (1975), "A Canadian replication of Mason Haire's 'shopping list' study," *Journal of the Academy of Marketing Science*, 3, 48-59.
- Leone, R.P. and R.L. Schultz (1980), "A study of marketing generalizations," *Journal of Marketing*, 44, 10-18.
- Madden, C.S., L.S. Franz and R.A. Mittelstaedt (1979), "The replicability of research in marketing: reported content and author cooperation," in O.C. Ferrell, S.W. Brown and C.W. Lamb (eds.), *Conceptual and Theoretical Developments in Marketing*. Chicago, IL: American Marketing Association, pp. 76-85.
- Mayer, T. (1980), "Economics as a hard science: Realistic goal or wishful thinking?" *Economic Enquiry*, 18, 165-178.
- McGrath, J.E. and D. Brinberg (1983). "External validity and the research process: A comment on the Calder/Lynch dialogue," *Journal of Consumer Research*, 10, 115-124.

- Mittelstaedt, R.A. and T.S. Zorn (1984), "Econometric replication: Lessons from the experimental sciences," *Quarterly Journal of Business and Economics*, 23, 9-15.
- Monroe, K. (1991), "Editorial," *Journal of Consumer Research*, 18, preface.
- Moore, T.E. (1982), "Subliminal advertising: What you see is what you get," *Journal of Marketing*, 46, 38-47.
- Neuliep, J.W. and R. Crandall (1990), "Editorial bias against replication research," *Journal of Social Behavior and Personality*, 5, 85-90.
- Oakes, M. (1986), *Statistical Inference: A Commentary for the Social and Behavioral Sciences*. New York, NY: Wiley.
- Reid, L.N., L.C. Soley and R.D. Wimmer (1981), "Replication in advertising research: 1977, 1978, 1979," *Journal of Advertising*, 10, 3-13.
- Reid, L.N., H.J. Rotfeld and R.D. Wimmer (1982), "How researchers respond to replication requests," *Journal of Consumer Research*, 9, 216-218.
- Robertson, T.S. (1971), *Innovative Behavior and Communication*. New York, NY: Holt, Rinehart and Winston.
- Robertson, D.H. and D.N. Bellenger (1978), "A new method of increasing mail survey responses: Contributions to charity," *Journal of Marketing Research*, 15, 632-633.
- Rosenthal, R. and R.L. Rosnow (1984), *Essentials of Behavioral Research: Methods and Meta Analysis*. New York, NY: McGraw-Hill.
- Rosnow, R.L. and R. Rosenthal (1989), "Statistical procedures and the justification of knowledge in psychological science," *American Psychologist*, 44, 1276-1284.
- Rowney, J.A. and T.J. Zenisek (1980), "Manuscript characteristics influencing reviewers' decisions," *Canadian Psychology*, 21, 17-21.
- Salsburg, D.S. (1985), "The religion of statistics as practiced in medical journals," *American Statistician*, 39, 220-223.
- Samelson, F. (1980), "J.B. Watson's Little Albert, Cyril Burt's twins, and the need for a critical science," *American Psychologist*, 35, 619-625.
- Sawyer, A.G. and A.D. Ball (1981), "Statistical power and effect size in marketing research," *Journal of Marketing Research*, 18, 275-290.
- Sawyer, A.G. and J.P. Peter (1983), "The significance of statistical significance tests in marketing research," *Journal of Marketing Research*, 20, 122-133.

SSCI (Social Science Citation Index), Institute for Scientific Information, 3501 Market Street, Philadelphia, PA 19104, USA.

Sterling, T.D. (1959), "Publication decisions and their possible effects on inferences drawn from tests of significance – or vice versa," *Journal of the American Statistical Association*, 54, 30-34.

Tversky, A. and D. Kahneman (1971), "Belief in the law of small numbers," *Psychological Bulletin*, 76, 105-110.

Walster, G.W. and T.A. Cleary (1970), "A proposal for a new editorial policy in the social sciences," *American Statistician*, 24, 16-19.

Webster, Jr., F.E. and F. von Pechmann (1970), "A replication of the 'shopping list' study," *Journal of Marketing*, 34, 61-63.

Wilkie, William L. (1986), *Consumer Behavior*. New York, NY: Wiley.