

THE CREDIBILITY OF USING STUDENTS AS SURROGATES IN EMPIRICAL RESEARCH: A NEW PERSPECTIVE ON AN OLD ISSUE

THOMAS BERGMANN, *University of Wisconsin-Eau Claire*
JOYCE GRAHN, *University of Minnesota-Duluth*

The use of students, though appealing to academicians, has received wide criticism because of evidence that students are not adequate substitutes for other target populations. This paper discusses the appeal of student samples, briefly reviews past research on student respondents including the relatively ignored area of behavioral processes within consumer decision making, empirically tests the use of student respondents in this area, draws inferences for appropriate use of student respondents, and suggests additional research tracks.

BACKGROUND

The convenience, minimal cost, and high compliance rates of student samples make them a highly attractive data source for academic researchers (Enis, Cox, and Stafford 1972; Cunningham, Anderson, and Murphy 1974). However, the use of student respondents as substitutes for statistically identified subjects from (primarily) non-student populations has been severely criticized because of reported external validity problems (Copeland, Francia, and Strawser 1973; Lamb and Stem 1979; Rubenstein 1982). Where it can be demonstrated that student responses are significantly different from non-student target populations, research findings become artificial and not generalizable to other situations (Ferber 1977).

The issue of appropriate use of student respondents will be examined by identifying the pervasiveness of the practice to determine whether the issue is important enough to warrant examination, analyzing

research findings concerning the external validity argument to determine whether the criticism is reasonable, suggesting situations where student samples may be properly utilized, empirically testing student samples under these specified conditions, and reporting findings and implications for future research.

Student Sample Usage

The use of students as surrogates for (primarily) non-student segments can be identified in both business and the social sciences (e.g., psychology) literature from which business derives information. Several surveys of business journals report student sample usage from 1967-1972 for consumer research studies to be 20 to 33 percent (Cunningham, Anderson, and Murphy 1974); and from 1979-1983 for major marketing journals to be approximately 17 percent (Burnett and Dunne 1986). Our own survey of the *Journal of Business Research*, *Journal of Consumer Research*, *Journal of Marketing*, and *Journal of Marketing Research*, from 1987 through 1995 is depicted in Table 1. The nine year time frame has been segmented into two segments to compare earlier and more recent student usage. The use of student

TABLE 1
A Comparison of Use of Student Subjects in
Journal of Business Research, Journal of
Consumer Research, Journal of Marketing, and
Journal of Marketing Research.

1987 - 1990			
Journal	% of Articles Using Human Subjects	% of Articles Using Student Subjects	% of Articles Using Human Subjects that Used Student Subjects
<i>Journal of Business Research</i>	61%	13%	22%
<i>Journal of Consumer Research</i>	76%	45%	60%
<i>Journal of Marketing</i>	42%	4%	9%
<i>Journal of Marketing Research</i>	51%	24%	47%

1991 - 1995			
Journal	% of Articles Using Human Subjects	% of Articles Using Student Subjects	% of Articles Using Human Subjects that Used Student Subjects
<i>Journal of Business Research</i>	64%	16%	24%
<i>Journal of Consumer Research</i>	78%	49%	62%
<i>Journal of Marketing</i>	55%	10%	19%
<i>Journal of Marketing Research</i>	52%	26%	44%

samples is high in both time frames. For example, the *Journal of Consumer Research* studies from 1991-1995, reported the use of students in over 60 percent of its articles using human subjects. The *Journal of Marketing Research* reported using student samples in 44 percent of its articles using human subjects; 24 percent of articles employing human subjects published in the *Journal of Business Research* studies reported using student samples. Clearly there is currently wide usage of student data bases in published empirical research.

Theoretical Controversy

There is controversy regarding the appropriateness of using students in place of representatives from non-student populations because of lack of evidence that student and non-student responses are similar (Copeland, Francia, and Strawser 1973, 1974; Watson 1974; Ferber 1977; Rubenstein 1982). However, empirical comparisons of responses of students vs. population subjects are inconclusive. In certain instances, responses of students vs. housewife consumers (Sheth 1970) and student vs. business representatives (Roering, Schooler, and Morgan 1976) appear to be alike; in other studies, the findings comparing students vs. household consumers (Cunningham, Anderson, and Murphy 1974; Park and Lessig 1977) or business students vs. business representatives (Alpert 1967; Copeland, Francia, and Strawser 1973) are different; and in still others the findings are mixed (Khera and Benson 1970; Enis, Cox, and Stafford 1972; Shuptrine 1975).

In an attempt to organize and frame the seemingly incongruous findings regarding the use of student surrogates in marketing, Lamb and Stem (1979) classified the empirical student surrogate literature into two segments: state and process. State research is research related to the subjects' mental state and examines subject attitudes, beliefs, opinions, and interests. The authors conclude that student and other population findings are different. Process research is research related to psychological decision processes, dissonance reduction processes, and information handling processes. Process research findings regarding surrogates are less consistent. Lamb and Stem (1979) surmise that an argument can be made that the student population and the general population make decisions or process information in the same fashion. They note the lack of similarity of results in student vs. general population process research and conclude the findings may have been the result of differences in age of cognitive abilities and not due to one group being students and the other group being from the target population. However,

the authors could not conclude that student populations are suitable for process research.

Burnett and Dunne (1986) analyzed the student surrogate literature and illustrated their findings in a meaningful and parsimonious research design. Their sample consisted of four groups: 1) college students ages 18 to 23 in business/business related courses; 2) parents of those college students; 3) a local consumer panel representing a cross-section of consumer characteristics; and 4) a cross-section of subjects ages 18-23 who had not gone to college. There was a substantial difference between the responses of students and those of the other three groups. The authors conclude that students could be used as surrogates only for similar student target populations. The authors suggest that future studies should include behavioral patterns and behavioral factors.

Calder, Phillips, and Tybout (1981; 1982; 1983) in a series of studies reject the argument for emphasizing external validity in theoretical research for any single research study. The goal of theory research is to provide a general understanding of the work. In designing a research study the respondents should be selected to provide a rigorous test of the theory at issue (Calder, et al. 1981). Therefore, theory applications research is concerned with falsification test procedures. They argue that homogeneous respondents (students, for example) are preferred because a homogeneous sample will strengthen the theory test.

Lynch (Dec. 1983; June 1983) counterargues that external validity is highly relevant for theoretical consumer research. Contrary to Calder et al., he maintains that background variables or variables external to the theory under research provide both a challenge and an opportunity for theory enrichment. He recommends "the model of deliberate sampling for heterogeneity." His philosophy precludes the routine use of students as surrogates for other non-student populations. Based on previous research, there appears to be an important gap in the student subject usage literature--examination of behavior in the

consumer decision process (Lamb and Stem 1979; Burnett and Dunne 1986) that is generalizable to real world situations. Our review of the literature indicates that there has been no study that examined the student surrogates issue by empirically testing process variables under a forced choice consumer decision while controlling for age and education as suggested. A comparison of information processing results of over the counter (OTC) drug label information by two naive groups -- students and non-students -- were used to determine if there is evidence to suggest that students should not be used as surrogates for the general population with similar age and education characteristics.

The Experiment Background

The issue of information processing in the health care industry is interesting from a scholarly viewpoint and important for public policy development. Information processing of nonprescription or OTC drug label information affects large numbers of individuals. Physical risk from misuse such as accidental overdosing can be high.

Knapp (1974), in a Federal Food and Drug Administration study of 1,321 adult subjects in 100 locations throughout the coterminous United States, found that OTC drug labels typically do not attract readership. Fifty percent of the subjects read about half to none of the labels on recently purchased OTC drugs. The use of graphics (nonverbal symbols and signs) and color have been shown to improve attention (Salcedo, Read, Evans, and Kong 1971; Blanken 1976).

In this same study Knapp (1974), label information was ranked with dosage directions being most important; side effects, warnings and cautions being second most important. Careful analysis of an OTC drug label reveals there are eight important parts to the label: dosage directions, graphics-if any, warnings, brand, amount, indications for product use, ingredients, and manufacturer.

This paper reports the findings of a study that measured the objective effects of alternative communication formats on information processing outcomes of OTC drug label information experienced by students and non-students of similar age and education. The two communication formats (high and low) consisted of simultaneously manipulated variables of readability level, graphics, presentation style, and color. (See Appendix).

The findings shed light on whether or not students can appropriately be used as representatives of similar age and education representatives from the non-student population.

METHODOLOGY

Sample

The student sample consisted of 27 students 18 to 35 years old, in the undergraduate business program of a major upper-Midwestern university. Several had four-year degrees and were fulfilling pre-MBA requirements. The non-student sample consisted of 41 adults, 18 to 35 years old, who had attended college or graduated from two or four-year institutions. The age and education backgrounds of the two groups were similar. Incentives were \$2.00 for non-students and extra points (enough to skip one quiz) for students.

A screening questionnaire measured use of OTC drugs under study, furniture polish, insect repellent, oven cleaners, and spray paint. Only subjects lacking purchase and use of the OTC drug category under investigation were included in the study.

Stimulus Materials

Sample labels were designed incorporating all eight label parts, as identified by Knapp (1974), into the stimulus material (labels) while simultaneously evaluating the combined effects of graphics, color, readability level, and information format. Labels were developed for fictitious products in an OTC drugs

product category and for spray paint to disguise study focus.

Subjects received two levels of information presentation: high and low. Low level presentation featured uniform color throughout the label, 7th through 12th grade readability levels, direction materials in paragraph form, and warnings and indications with no graphics. Each low level label was converted to a high level presentation by printing directions and warnings in red to heighten awareness; by simplifying (reducing) readability levels of the warnings, directions, and indications to the 5th and 6th grade level using the Dale-Chall formula (1948a; 1948b); presenting information in tabular form to increase comprehension; and imbedding a graphic for redundancy in the directions and warnings sections of the label to heighten awareness and increase comprehension.

Subjects were exposed to the same information presentation levels on all labels. That is, each subject received two spray paint and two OTC drug labels at either the high level or the low level.

Data Collection

Data collection occurred at local gathering places for non-students and a large, comfortable, seminar type room for students. Each subject was exposed to a screening questionnaire which measured purchase and use of OTC drugs and other consumer products to separate naive and experienced subjects. Only naive subjects for the OTC drug category under investigation became part of the final sample to minimize the impact of previous acquisition and inventorying of product-related information.

Subjects were first exposed to spray paint labels to disguise the focus of the study and familiarize them with experiment format. They were advised that they had a need for spray paint and were asked to read the two labels and make a brand choice decision and return the labels to the envelope. Without referring to the labels, subjects were asked a set of multiple-

choice questions regarding information on the labels. The process was then repeated for an OTC drug category. Each subject was given the situation setting that he or she was experiencing a specific minor health problem contained in their stimulus materials. Instructions implied that the product would provide relief and subjects were asked to choose one of the alternative products from the label information. (Debriefing indicated that subjects were actively involved in the role play.) Thus subjects made two brand decisions—one for spray paint and one for the OTC drug-related situation under analysis.

The development of the multiple-choice questions to measure the comprehension dimension of information processing presented some unique problems. Since the labels were presented in two levels, comprehension questions were developed to match the presentation of information on their labels. A ten-point multiple choice test was developed to measure overall comprehension of OTC drug label information and focused on the following label areas: directions, warnings, brand, amount, indications, active ingredients, and manufacturer. Because wording of the comprehension items matched information processing levels on the labels, there was a potential internal validity problem where items measuring similar information were worded slightly differently. Therefore, questions 6 and 9 were worded exactly the same for both levels. Potential internal validity problems would be identified if error rates on questions worded exactly alike for all respondents were similar.

RESULTS

Total scores for comprehension of drug label information had a potential range of from zero for no correct scores to ten for ten correct scores. Each individual item was rated zero for an error and one for a correct response. The primary analysis is presented in Table 2. Multivariate analysis tests using Pillais, Hotelling's and Wilk's Statistics were then used to evaluate the findings. There was no significant difference between the two groups.

Neither presentation level nor group impacted significantly on subjects' comprehension scores. There was no significant interaction effect. The mean score correct for both levels and groups was 77 percent with a standard deviation of 1.64.

The lowest error rates, or highest correct percentages -- less than 8 percent -- were on items 3 (warning with a graphic), 6 (brand), and 8 (indications). The highest error rate was on item 9 (active ingredients). Respondents, whether student or non-student, remembered the brand name, the conditions for usage, and potential hazards associated with the drug. Their comprehension was influenced simultaneously by graphics, color, readability level, and information format.

CONCLUSIONS AND IMPLICATIONS

The findings of this study are interesting because they indicate that students can be surrogates for general population representatives with similar characteristics. Students do not respond differently than their non-student counterparts in responses related to label information in a forced-choice decision. The notion of Lamb and Stem (1979) regarding process research and suggested research questions by Burnett and Dunne (1986) proved fruitful for investigation.

This study is different from previous studies because respondents did not report how they made a decision. They made a decision and then, through the use of a test and without seeing stimulus materials again, demonstrated what they comprehended. No self-evaluation was involved. Respondents were not asked how they learned; learning was merely measured.

The entire area of consumer process research needs to be precisely defined and investigated. As the increased emphasis on research in academic institutions impels faculty to increase publications Bergmann and Grahn (1989), opportunities for productivity improvement need to be carefully

TABLE 2
Correct Scores on Comprehension Questions by Information Presentation Level by Type of Sample
(1.000 means everyone chose the correct answer).

Question Number	Area	Label	Group 1 Non-Students				Group 2 Students				Group by Level	
			Presentation Level		Presentation Level		Presentation Level		Presentation Level			
			N=16	N = 14	N = 13	N = 9						
			High	Low	High	Low						
			N=21	N=20	N=13	N=14	N=68					
			̄x	SD	̄x	SD	̄x	SD	̄x	SD		
1	Directions		.687	.479	.786	.426	.846	.376	.444	.527	.716	.457
2	Directions*		.813	.403	.786	.426	.846	.376	.667	.500	.788	.412
3	Warnings*		.938	.250	1.00	.000	.925	.277	.778	.441	.923	.269
4	Warnings		.875	.342	.714	.469	1.00	.000	.444	.527	.788	.412
5	Directions/Warnings/ Indications		.625	.500	.571	.513	.769	.439	.778	.441	.673	.474
6	Brand		.875	.341	.923	.267	.923	.278	1.00	.000	.923	.269
7	Amount		.936	.250	.786	.426	.769	.439	.778	.441	.827	.382
8	Indications		1.000	.000	.929	.267	.923	.277	.889	.333	.942	.235
9	Active Ingredients		.563	.512	.643	.497	.231	.439	.333	.503	.462	.503
10	Manufacturer		.813	.403	.500	.519	.615	.506	.444	.527	.615	.491
Total			8.12	(1.5)	7.64	(2.02)	7.84	1.14	6.67	1.65	7.67	1.64

*Graphics at high level

examined. Inappropriate sample selection can confound theory development and cause serious strategic marketing management miscalculations. Given the prevailing use of student samples as reported in the business journals, it is imperative that legitimate use of student data bases be identified to minimize artifactual (non-generalizable) findings.

REFERENCES

- Alpert, Bernard (1967), "Non-Businessmen as Surrogates for Businessmen in Behavioral Experiments," *Journal of Business*, 40 (April), p. 203-207.
- Bergmann, Thomas and Joyce Grahn (1989), "An Audit of the Role and Research Function of Personal Selling and Sales Management in Schools of Business," *Journal of Midwest Marketing*, 4, p. 224-230.
- Blanken, Gary E. (1976), *Consumer Attitudes Toward Over-the-Counter Drug Labels*, An OPE Study, Food and Drug Administration: n.p.
- Burnett, John J. and Patrick M. Dunne (1986), "An Appraisal of the Use of Student Subjects in Marketing Research," *Journal of Business Research*, (August), p. 329-343.
- Calder, Bobby J., Lynn W. Phillips, and Alice M. Tybout (1981), "Designing Research for

- Application," *Journal of Consumer Research*, (September), p. 197-207.
- Calder, Bobby J., Lynn W. Phillips, and Alice M. Tybout, (1982), "The Concept of External Validity," *Journal of Consumer Research*, (December), p. 240-244.
- Calder, Bobby J., Lynn W. Phillips, and Alice M. Tybout, (1983), "Beyond External Validity," *Journal of Consumer Research*, (June), p. 112-114.
- Copeland, Ronald M., Arthur J. Francina, and Robert H. Strawser, (1973), "Students as Subjects in Behavioral Research," *The Accounting Review*, (April), p. 365-372.
- Copeland, Ronald M., Arthur J. Francina, and Robert H. Strawser, (1974), "Further Comments on Students as Subjects in Behavioral Business Research," *The Accounting Review*, (July), p. 534-537.
- Cunningham, William H., W. Thomas Anderson, Jr., and John H. Murphy, (1974), "Are Students Real People?" *Journal of Business*, 40 (July), p. 399-409.
- Dale, Edgar and Jeanne S. Chall, (1948a), "A Formula for Predicting Readability," *Educational Research Bulletin*, 27, p. 11-20.
- Dale, Edgar and Jeanne S. Chall, (1948b), "A Formula for Predicting Readability: Instructions," *Educational Research Bulletin*, 27, p. 37-54.
- Enis, Ben M., Keith K. Cox, and James E. Stafford, (1972), "Students as Subjects in Consumer Behavior Experiments," *Journal of Marketing Research*, 9 (February), p. 72-74.
- Ferber, Robert, (1977), "Research by Convenience," *Journal of Consumer Research*, 4 (June), p. 57-58.
- Khera, Inder P. and James D. Benson, (1970), "Are Students Really Poor Substitutes for Businessmen in Behavioral Research?" *Journal of Marketing Research*, 7 (November), p. 529-532.
- Knapp, Deanne, ed. (1974), *Consumers and Medication*, P B - 232 172, Springfield, Virginia: National Technical Information Service, U.S. Department of Commerce.
- Lamb, Charles W., Jr. and Donald E. Stem, Jr. (1979), "An Evaluation of Students as Surrogates in Marketing Studies," in Jerry Olson, ed., *Advances in Consumer Research* 7, p. 796-799.
- Lynch, John G., Jr. (1982), "On the External Validity of Experiments in Consumer Research," *Journal of Consumer Research*, (December), p. 225-239.
- Lynch, John G., Jr. (1983), "The Role of External Validity in Theoretical Research," *Journal of Consumer Research*, 10 (June), p. 109-111.
- Park, C. Whan and V. Parker Lessig, (1977), "Students and Housewives: Differences in Susceptibility to Reference Group Influence," *Journal of Consumer Research*, 4 (September), p. 102-110.
- Roering, Kenneth J., Robert D. Schooler, and Fred W. Morgan, (1976), "An Evaluation of Marketing Practices: Businessmen, Housewives, and Students," *Journal of Business Research*, (May), p. 131-144.
- Rubenstein, Carin, (1982), "Psychology's Fruit Flies," *Psychology Today*, (July), p. 83-84.
- Salcedo, Rodolfo N., Hadley Read, James F. Evans, and Ana C. Kong, (1971), *Improving the Communication Adequacy of Pesticide Labels: Phase 1 Summary Report*, Agricultural Communications, College of Agriculture, University of Illinois at Urbana-Champaign.
- Sheth, Jagdish N. (1970), "Are There Differences in Dissonance Reduction Behavior Between Students and Housewives?" *Journal of Marketing Research*, 7 (May), p. 243-245.
- Shuptrine, F. Kelly, (1975), "On the Validity of Using Students in Consumer Behavior Investigations," *Journal of Business*, 48 (July), p. 383-390.
- Watson, David J. H. (1974), "Students as Surrogates in Behavioral Business Research: Some Comments," *The Accounting Review*, (July), p. 530-533.