

PROFILES OF HUMAN SUBJECT RESEARCH IN FOUR INTERNATIONAL EDUCATION JOURNALS

Ralph Guice
Phillip Institute of Technology
and
Glenn Rowley
Monash University

ABSTRACT

Content analysis was used to examine 262 human-subject research articles published over the period 1978-81 in four important international education journals: the Australian Journal of Education (AJE), Educational Research (ER), the Canadian Journal of Education (CJE), and the American Educational Research Journal (AERJ).

AERJ more frequently used experimental methods, whereas surveys and other descriptive methods predominated in AJE, ER, and CJE. For all four journals random sampling was little used and often sampling methods were not described or a convenience (available) sample was obtained. Some of the many data collected are presented and problems of interpretation which arise are discussed.

Shaver and Norton (1980) examined six volumes of the American Educational Research Journal (AERJ) published over the ten-year period from 1968 to 1977, categorizing each of the 151 human-subject research studies reported therein along several dimensions. These related mostly to the type of research, method of sample selection and the treatment of issues related to the generalizability of findings. Shaver and Norton were quite critical in their comments on some aspects of the research they examined. Their study reported a preponderance of correlational and experimental research, little reference to either target population or accessible population, frequent failure to describe methods of sample selection, and a tendency to report conclusions without qualifying them according to the population or the sample used.

Although there have been other studies focussing on research in the United States of America, it was not until a recent work by Guice and Rowley (1982) that an attempt was made to chart the course of Australian-published educational research. They applied the methodology of Shaver and Norton to make a comparative study of statistical and methodological aspects of two Australian education journals: the Australian Journal of Education and Education Research and Perspectives. Neither journal is exclusively dedicated to reporting research studies, so the Guice and Rowley study analyzed those articles in the two journals which reported the results of empirical research. In contrast with AERJ, the Australian journals favoured descriptive, correlational and quasi-experimental studies and reported very little experimental research. In common with the U.S. studies, target populations and methods of sample selection were frequently not described. Although minor differences in emphasis between the two journals were noted, the study seemed to indicate that there were systematic differences between the methodological characteristics of Australian and American educational research.

The present study aims to build on the previous work of Guice and Rowley by making systematic comparisons between the research reported in major journals from each of four countries: Australia, the United States of America, Canada and Great Britain.

SOURCES OF DATA

Four journals were chosen for detailed study.

The Australian Journal of Education (AJE) is the most prestigious outlet for the publication of educational research findings in Australia, although it is not restricted to research reports. Its stated aim (Connell, 1971) is "to present the best thought, scholarship and research from all who are professionally interested in education in Australia, and to provide a medium for the dissemination of information concerning significant developments in educational practice and research." AJE commenced in 1957 and is widely circulated (about

2500 in 1982) through the auspices of the Australian Council for Educational Research.

In Canada, two journals were in contention - the Alberta Journal of Educational Research and the Canadian Journal of Education (CJE). The latter journal was chosen for this study because it is more similar in style and content to AJE and ER. It clearly aspires to represent the best of Canadian research and scholarship in education. CJE is the youngest of the journals studied; published by the Canadian Society for the Study of Education, its first issue appeared in 1976. Like AJE, it contains both research reports and other scholarly writings on education.

Educational Research (ER) is the official journal of the National Foundation for Educational Research in England and Wales (NFER). ER aims to be "a balanced journal specializing in reviews of research and short reports of original studies" (Wall, 1968, p.162). It began in 1958, and, like AJE and CJE, is not devoted entirely to the publication of research findings, but contains articles on many aspects of education, including some of a purely expository nature.

The American Educational Research Journal (AERJ) has the largest circulation of the four journals (12,537 in 1981) and probably the greatest prestige. Like the other three journals, it is the major outlet for the publication of educational research reports in its country of origin; unlike the other three, it is restricted entirely to the publication of research results.

This study systematically catalogues articles published in the four journals during the four complete years 1978-1981 (1981 being the latest year for which all issues were available at the time the study was commenced). Only research studies involving human subjects were considered, forming a parallel with Shaver and Norton's original work (1980). Table 1 provides a breakdown of the numbers of articles involved for each of the four journals. For each journal, the third line in the table describes the population of studies considered in this report.

TABLE 1
ARTICLES EXAMINED

	Journal				Total
	AJE	ER	CJE	AERJ	
Articles: total number	82	93	102	144	421
Articles: research with human subjects	41	52	47	122	262
Studies: research with human subjects	43	54	48	140	285

A grand total of 421 articles was examined from the four journals, yielding 262 articles describing research with human subjects, and 285 actual studies with human subjects (sometimes it was necessary to distinguish two separate studies within the same article). We have chosen to use the study, rather than the article, as the unit of analysis when computing percentages, etc., because it leads to less difficulty of classification. However, so few articles contained more than one distinct research study that the results reported may be taken as roughly representative of research articles as well.

METHOD

Each study was read by the senior author, and a series of questions answered, as in the original Shaver and Norton study. These questions were:

1. Did the article report findings of research with human subjects?
(If the answer was "no," no further questions were asked.)
2. What was the study type (experimental, correlational, etc.)?
3. What sort of knowledge-building approach was utilized?
4. Was the term "target population" used?
5. Was the target population defined?

6. Was the term "accessible population" used?
7. Was the accessible population defined?
8. Was the accessible population described, with data?
9. What was the method of sample selection?
10. Was the sample described, with data?
11. Were the conclusions explicitly limited by shortcomings in the accessible population vis-a-vis the target population?
12. Were the conclusions explicitly limited by sample deficiencies?
13. Did the study involve new data?
14. Did the study involve the reanalysis of previously gathered data?
15. Was a previously-published standardized test or measure used?

Questions 1 through 14 are similar to those used in the original Shaver and Norton study. Question 15 was added because it was expected that the AERJ, because of the greater use made of standardized tests in the United States, would be likely to differ noticeably from the other three journals.

RESULTS

Research with human subjects

As expected, virtually all of the articles in AERJ reported research with human subjects, compared to about half for the other three journals (Table 1). In spite of the conscious editorial policy of AERJ to report empirical research rather than philosophical and discursive articles, the range of topics covered in AERJ, as in all four journals, was extremely wide. For AJE, the percentage dropped from 81 percent in 1979 to 37 percent in 1981.

Study type

Studies were categorized according to type, or method, using an abbreviated version of the Isaac and Michael (1976) taxonomy, and the results are summarized in Table 2.

TABLE 2
TYPES OF RESEARCH STUDY

Study Type	Journal				Total
	AJE	ER	CJE	AERJ	
Experimental	3 (7%)	2 (4%)	2 (4%)	42 (30%)	49 (17%)
Quasi-experimental	4 (9%)	2 (4%)	7 (15%)	24 (17%)	37 (13%)
Correlational	11 (26%)	6 (11%)	8 (17%)	41 (29%)	66 (23%)
Causal-Comparative	8 (19%)	8 (15%)	5 (10%)	13 (9%)	34 (12%)
Survey	16 (37%)	32 (59%)	23 (48%)	15 (11%)	86 (30%)
Other	1 (2%)	4 (7%)	3 (6%)	5 (4%)	13 (5%)
Total Studies	43	54	48	140	285

There is no "best" research method. The most suitable approach in a particular instance is determined by "the questions being asked, the resources available, and the nature of previous work in the field" (Mason and Bramble, 1978, p. 47), as well as such other factors as the prevailing political climate and ethical concerns. Differences among the four countries would reflect differences in political and ethical values, as well as in areas of research convergence.

As was noted in the previous study (Guice and Rowley, 1982), AERJ is distinguished by its much greater use of experimental method (30 percent, compared to negligible proportions in the

other three journals). Correlational studies form a strong cornerstone for AERJ (29 percent) and AJE (26 percent), while survey research is more commonly reported in AJE, ER and CJE than in AERJ. The apparently increased emphasis on action research, ethnographic methods and case studies in the Australian setting during recent years was not reflected in the data of this study.

Knowledge-building approach

Although empirical research can contribute in many ways to the development of knowledge, we have concentrated on three in particular. They are

- (1) testing the consequences of a theory,
- (2) replication of previous work, and
- (3) extension of previous findings.

"Replication, a fusion of the words *duplication* and *repetition*, is deliberately repeating a study, using identical procedures with different subjects, at a different time and in a different setting" (Best, 1977, p. 143). Shaver and Norton distinguished two kinds of replication: direct (when the researcher tries to duplicate a study exactly), and systematic (when the researcher tries to systematically vary features of the research in order to test the generalizability of the findings). Many writers have commented on the lack of replication in educational research (e.g., Kerlinger, 1973, p. 681), and have noted the risks involved in accepting findings based on unreplicated studies.

It is interesting to note the rarity of any kind of replication in the four journals, and its complete absence from the Australian and British journals. No doubt this is an outcome of a reward structure in which original work attracts greater recognition (and, we suppose, funds) than do replication studies. We may do well to heed the warnings sounded by Kerlinger (above) and others.

TABLE 3
TYPES OF KNOWLEDGE BUILDING

	Journal				Total
	AJE	ER	CJE	AERJ	
Replication	0 (0%)	0 (0%)	1 (2%)	7 (5%)	8 (3%)
Extending findings	29 (67%)	43 (80%)	38 (79%)	97 (69%)	207 (73%)
Theory testing	8 (19%)	1 (2%)	4 (8%)	25 (18%)	38 (13%)
Other or none	6 (7%)	10 (19%)	5 (10%)	11 (6%)	32 (11%)
Total studies	43	54	48	140	285

The category "extending findings" in Table 3 refers to studies in which a literature survey was provided, and the intent was clearly to gather further information about questions raised in the literature. Where the questions raised could be seen as questions upon which the credibility of a theory might hang, the study was categorized as "theory testing"; otherwise it was simply "extending findings." This distinction was frequently difficult, depending on the clarity of the writers' statements of their intentions. For a large enough proportion, the intention was either ambiguous, or did not fit into any of the above three categories. In these cases, the category "other, or none" has been used.

"Extension of previous findings" is by far the most frequent categorization, with replication being almost non-existent, and theory testing being much more common in AJE and AERJ than in the other two journals. For ER, this probably is partly due to the continuation of an editorial policy stated fifteen years ago: "the emphasis will be increasingly placed on

a straightforward and explicit style, and no article will be accepted which is written in technical jargon or which is not readily understood by a teacher or advanced student" (Wiseman, 1968, p. 3). The picture, from all four journals, however, is one of predominance of applied research (directed towards more immediate practical concerns), rather than the development of theory.

Sampling issues

A fundamental question about any research study is "how were the data obtained?" Particularly with regard to external validity, or generalizability, a reader of research is entitled to a straightforward description of the essential characteristics of the sample, so that, at least in an intuitive way, judgements can be made about the possible applicability of the findings to other settings and other subjects. Some studies go much further than this, however. An ideal survey, for example, might proceed as follows:

- (1) define a target population,
- (2) identify those elements of the target population which are accessible to the researchers,
- (3) draw a random, or otherwise representative sample from the accessible population,
- (4) gather the required information from the sample.

In such circumstances, the researcher may proceed with full confidence that sample results can be generalized (within calculable limits) to the accessible population. For survey results, this must be a major concern. For experimental, and many other studies, it is frequently less of a concern than is internal validity. Nevertheless, it is of interest to examine questions of how samples are obtained and described in the four journals. The relevant information is provided in Tables 4 and 5.

TABLE 4
ASPECTS OF DATA DESCRIPTION

	Journal				Total
	AJE	ER	CJE	AERJ	
Term target population used	2 (5%)	2 (4%)	0 (0%)	1 (1%)	5 (2%)
Target population defined	18 (42%)	23 (43%)	19 (40%)	25 (18%)	85 (30%)
Term accessible population used	0 (0%)	0 (0%)	0 (0%)	1 (1%)	1 (0%)
Accessible population defined	33 (77%)	41 (76%)	34 (71%)	78 (56%)	186 (65%)
Accessible population described with data	8 (19%)	13 (24%)	2 (4%)	16 (11%)	39 (14%)
Sample described with data	27 (63%)	39 (72%)	30 (63%)	64 (46%)	160 (56%)
New data analysed	39 (91%)	44 (81%)	43 (90%)	102 (73%)	228 (80%)
Previously gathered data	4 (9%)	10 (19%)	5 (10%)	39 (28%)	58 (20%)
Standardized test or measure used	23 (53%)	17 (31%)	14 (29%)	74 (53%)	128 (45%)
Total Studies	43	54	48	140	285

TABLE 5
METHOD OF SAMPLE SELECTION

	Journal				Total
	AJE	ER	CJE	AERJ	
Random	5 (12%)	2 (4%)	7 (15%)	15 (11%)	29 (10%)
Representative	8 (19%)	11 (20%)	1 (2%)	10 (7%)	30 (11%)
Volunteers	1 (2%)	7 (13%)	2 (4%)	29 (21%)	39 (14%)
Selected by others	0 (0%)	5 (9%)	2 (4%)	10 (7%)	17 (6%)
Available	17 (40%)	8 (15%)	23 (48%)	24 (17%)	72 (25%)
Whole population	1 (2%)	3 (6%)	3 (6%)	1 (1%)	8 (3%)
Can't tell	11 (26%)	18 (33%)	10 (21%)	51 (36%)	90 (32%)
Total Studies	43	54	48	140	285

We note from Table 4 that it is most unusual for authors to write in terms of target and accessible populations, but that in most cases, some attempt was made to define, or at least describe, the accessible population, which is the population towards which generalization might be justified. Provision of descriptive data about the accessible population was less common, and where descriptive data were provided, it was mostly derived from the sample. Differences among the four journals were not great, but it is noted that AERJ (surprisingly) contained a smaller proportion of studies in which population definition was attempted. This almost certainly relates to the higher proportion of experimental studies (where internal validity is the prime consideration) in AERJ, and the smaller number of surveys (where external validity is paramount).

Given that a population has been defined, the method of sample selection is a critical determinant of the confidence with which results may be generalized to that population. Random sampling methods are usually recommended, and, indeed, most of the methods of inferential statistics are based on the assumption that a random sample has been drawn from a well-defined population. Only sometimes was this the case. Also uncommon was a sample constructed in a deliberate attempt to include the major subgroups present in the population (what we have called here a "representative" sample). We might note that, when the categories of random and representative samples are combined, it is AJE (31 percent) which leads the field, with AERJ, which, on the surface, aspires to a greater degree of scientific rigour, well back (18 percent). A tiny proportion of studies were carried out on whole populations, and for these sampling was not an issue. Many studies were carried out with samples chosen for their availability, and for a distressingly high proportion in each journal, insufficient information was provided for the method of sample selection to be determined. These last two categories (available sample, and can't tell) together made up a majority of samples taken in all four journals.

We accept that random sampling from a well-defined population is often not feasible in applied research areas. It is desirable, though, that some degree of caution be exercised in generalizing the results of the studies. The failure to explicitly limit conclusions because of shortcomings in accessible populations is an easily remediable defect. Easily remediable it may be, but it rarely occurs. Where conclusions are explicitly limited, it is more often in terms of the sample (and for the sampling methods primarily used, this is understandable). However the majority of studies in all four journals limited their

conclusions neither by the accessible population nor by the sample. This does not mean that the authors claimed an unwarranted degree of generalization, but their failure to disavow it leaves the way open to readers to overinterpret the findings.

Readers are better able to judge the applicability of findings to settings other than those in which the research was conducted if the samples are adequately described in the research report. Most studies did provide data-based descriptions of the samples used - obvious majorities in AJE, ER and CJE, with AERJ somewhat less satisfactory in this respect (46%, and down dramatically from comparable figures from the previous ten-year period). It would seem both feasible and desirable that editors of research journals insist that both samples and sampling methods be clearly described, and that data be provided where appropriate.

The vast majority of studies reported new data, but with the development of meta-analysis (Glass, 1976) it is likely that this will change, and we may expect to see an increasing proportion of studies, whether secondary analyses or meta-analyses, which examine previously gathered data. An impression is that this is already being seen in AERJ, and is beginning to appear in AJE.

Finally, we note that standardized tests and/or scales were used in 53 percent of studies in AJE and AERJ and in around 30 percent of ER and CJE studies. This was not in accord with our expectations. We note, however, that with regard to the use of standardized tests of achievement, as distinct from other measures, the U.S. journal, with 34 percent, was ahead of its Australian counterpart with 23 percent. Given AERJ's greater emphasis on experimental research, we find it surprising that AJE studies make just as much use of standardized tests and measures. To examine this question further would involve breaking the data down by type of research.

DISCUSSION AND CONCLUSIONS

This paper has attempted to bring together some of the findings, and to answer some of the questions raised, in previous studies by Shaver and Norton (1980) and Guice and Rowley (1982) on the United States and Australian journals, respectively. We have not attempted to evaluate the research from the four countries, but merely to describe some of its more objectively assessed properties. In doing so, we have noted many similarities, and a small number of easily noticeable differences.

AERJ is the only journal devoted exclusively to reporting the results of empirical research. The research reported in it is mainly experimental and correlational, with a lesser but still substantial proportion of quasi-experimental studies. Most studies reported seem to be aimed at extending the findings of previous work, less than one-fifth are direct tests of theory, very few are replications, and about one quarter of the studies present analyses of previously reported data. Few surveys are reported, and sampling methods are often undisclosed. Like the other three journals, authors in AERJ only occasionally make reference to accessible populations in discussing the limitations of their findings. When the limitations are discussed, it is most likely to be in terms of the sample, but (again in common with the other three) this usually does not occur.

The Australian, British and Canadian journals were similar in many respects. Each serves a double purpose - the reporting of empirical research, and the discussion of important issues in education. All three report more survey research and less experimental and quasi-experimental research than does the American journal. Like AERJ, they use a variety of sampling methods, scarcely use random samples, and, with the exception of ER, rely heavily on available samples. Predominantly, they report original research with newly obtained data, only ER making substantial use of secondary analyses. Replication studies are extremely rare in CJE, and non-existent in AJE and CJE. The Australian journal has more in common with ER and CJE than with AERJ. It is more theory-oriented than the British and Canadian journals, and in this respect, most resembles AERJ.

Although we did not attempt to evaluate the research, our impression was that the research itself, and the standard of reporting in AJE compare most favourably with the other three

journals. Given the higher incidence of survey research in AJE, ER and CJE, one might hope for greater care in defining the population, more use of random sampling techniques, and a greater willingness to discuss the limitations placed on the generalizability of the findings by the method of sample selection. Finally, it seems reasonable to suggest that editors should require all authors to describe how their samples were obtained. For almost one-third of the studies examined, this was not done, and no journal was blameless in this respect.

REFERENCES

- Best, John W. "Research in education." (3rd Ed.) Englewood Cliffs, New Jersey: Prentice Hall, 1977.
- Connell, W.F. (editor) "Editorial." The Australian Journal of Education, 1971, 15, 3, 13.
- Glass, G.V. "Primary, secondary and meta-analysis of research." Educational Researcher, 1976, 10, 3-8.
- Guice, R.T., and Rowley, G.L. "Statistical and methodological aspects of empirical research studies in two Australian educational journals: 1974-1980." Education Research and Perspectives, 1982, 9, 2, 16-30.
- Isaac, S., and Michael, W.B. "Handbook in research and evaluation." Robert R. Knapp, 1976.
- Kerlinger, Fred N. "Foundations of behavioral research." (2nd Ed.) Holt, Rinehart and Winston, 1973.
- Mason, E.J., and Bramble, W.J. "Understanding and conducting research/Applications in education and the behavioral sciences." McGraw-Hill Inc., 1978.
- Shaver, J.P., and Norton, R.S. "Randomness and replication in ten years of the American Educational Research Journal." Educational Researcher, 1980, 9, 1, 9-15.
- Wall, W.D. (editor) "An announcement." Educational Research, 1968, 10, 3, 162.
- Wiseman, S. (editor) "Editorial statement." Educational Research, 1968, 11, 1, 3.