DOCUMENT RESUME

ED 353 332 TM 019 398

AUTHOR	Daniel, Thomas
TITLE	Observations of the Statistical Procedures Used in the "Journal of Research in Music Education," 1987-1991.
PUB DATE	Nov 92
NOTE	19p.; Paper presented at the Annual Meeting of the Mid-South Education Research Association (Knoxville, TN, November 11-13, 1992).
PUB TYPE	Information Analyses (070) -- Reports - Descriptive (141) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Computation; Computer Software; *Educational Research; Educational Trends; Effect Size; Elementary Secondary Education; Higher Education; Incidence; Literature Reviews; Longitudinal Studies; Meta Analysis; Music Education; *Periodicals; *Research Methodology; Research Reports; Sample Size; *Statistical Analysis
IDENTIFIERS	Alpha Coefficient; *Journal of Research in Music Education; Power (Statistics)

ABSTRACT

This paper studies statistical procedures used in the "Journal of Research in Music Education" (JRME) over a 5-year period (1987 through 1991). The results are compared with those of L. D. Goodwin and W. L. Goodwin (1985) concerning the "American Educational Research Journal" for the same period. In all, 78 quantitative and 31 non-quantitative JRME articles were used. A 10-item summation chart was developed (power, alpha, effect size, sample size, basic procedures, method of computation, basic assumptions, references, repeated univariate, and other). Results for individual articles were grouped into issue totals, volume totals, and overall item totals. The results agree with those of Goodwin and Goodwin; basic, intermediate, advanced, and other techniques are used at about the same rate in both educational journals. Analysis of variance is currently the dominant statistical procedure in educational research. The chi square procedure was the next dominant statistical procedure used. Neither qualitative research nor non-parametric procedures are widely used. Power and effect size are mostly ignored in the behavioral sciences. Music education researchers must deal with power and effect size. Problems caused by researcher unfamiliarity with advanced computer packages and an over-abundance of stepwise regression are considered. Three tables are included. (RLC)

Observations of the Statistical Procedures Used in the Journal of Research in Music Education, 1987-1991

Thomas Daniel

Auburn University

Paper presented at the meeting of the Mid-South Education Research Association, Knoxville, Tennessee, November, 1992.

Observations of Statistical Procedures Used in the Journal of Re^earch in Music Education, 1987-1991

For many years, educational researchers have been interested in studies of research methodologies, procedures, and statistical techniques. These studies have been used in various ways to interpret the directions of the many diverse branches of the behavioral sciences. Edgington (1974) tabulated statistical procedures used in APA journals in even years from 1948 to 1972. Edgington discovered that analysis of variance was used in 11% of the articles in 1948, but by 1972 analysis of variance was used in 71% of the articles. Each year saw an increase in the usage of analysis of variance. In contrast, several procedures decreased in usage during the same time period. For example, the *t*-test reversed from 51% usage in 1948 to 12% by 1972, and parametric correlation and Spearman's rank-difference method also decreased from 42% usage in 1948 to 25% by 1972. Traditional chi square procedures increased from 10% to 15% and factor analysis from 1% to 4%. Other nonparametric procedures also increased from 3% to 12%. Several articles used multiple procedures which resulted in the percentages summing to more than 100%.

Willson (1980) examined the statistical techniques used in the American Educational Research Journal (AERJ) in the ten-year period from 1969 to 1978. The author compared the techniques used in the different major disciplines of psychology and found that analysis of variance, a standard technique in psychology, was used in 56% of the articles. Factor analysis, according to Willson, was the technique used more widely in the field of education. Other techniques utilized were multiple regression/correlation, discriminant analysis, and multivariate analysis of variance.

Goodwin and Goodwin (1985) continued this line of research with the American Educational Research Journal in the five-year period from 1979-1983 in an attempt to identify which types of techniques a reader could expect to find. They evaluated 27 types of statistical procedures used in the articles and classified the major techniques as basic, intermediate, and advanced. The basic procedures, identified as frequencies, percentages, measures of central tendency, variability, Pearson correlation, chi square, independent and dependent t-tests, and one-way analysis of variance, constituted 33% of the research. The intermediate procedures, identified as factorial analysis of variance, planned and post-hoc comparisons, trend analysis, one-way and factorial analysis of covariance, part/partial correlation, and multiple regression, constituted 37% of the research. The advanced procedures were identified as discriminant analysis, path analysis, canonical correlation, factor analysis, cluster analysis, one-way and factorial multivariate analysis of variance/multivariate analysis of covariance, and meta analysis. These advanced procedures were used in 17% of the research.

Halpin and Halpin (1988) investigated the research methodologies used over a five-year period in dissertations at Auburn University, relating them to the types of methodologies used in current educational research journals. Their analysis of dissertations revealed that 93% were determined to be quantitative. Of these quantitative dissertations, 45% were correlational, 34% were experimental or quasi-experimental, and 23% were survey research (many used multiple procedures, as the categories were not mutually exclusive). Analysis of variance was used in 57% of the dissertations and multiple regression/correlation was used in 22% of the dissertations. In addition, multivariate data analysis and chi square was used in 47% and 9% of the dissertations, respectively. The authors found that the dissertations used advanced statistical techniques for the major research hypotheses in 51% of the dissertations, as compared to 17% of the AERJ articles examined by Goodwin and Goodwin (1985).

Thompson (1988) examined recurring errors found in published dissertations. The author listed seven common mistakes and potential ways to correct/avoid them. Since doctoral dissertations are guided, controlled, edited, and approved by doctoral faculty, then recurring mistakes in dissertations reflect equally on both students and faculty.

Sedlmeier and Gigerenzer (1989) re-created Jacob Cohen's classical 1962 study of statistical power using the 1984 volume of the Journal of Abnormal Psychology to examine the impact of such studies on research in the behavioral sciences. The authors found numerous inconsistencies in the research designs: 50% of the articles used alpha-adjustment procedures to compensate for low power when they actually decrease power even more, and 11% of the articles treated the null hypothesis as an "operationalization" (Sedlmeier and Gigerenzer, 1989, p. 312) of the research hypothesis. They calculated the median power to be .37 for a medium size effect, as compared to Cohen's estimated median power of .46. Additionally, the authors also examined 11 other studies of statistical power in various fields of research concluding that these studies of power seem to have no effect on the statistical power of quantitative studies.

These types of analyses have never been attempted in the field of music education. One of the most prominent publications in this field is the Journal of Research in Music Education (JRME), a quarterly publication of the Music Educators National Conference. Therefore, the main objective of this study was to investigate and make observations about the types of statistical procedures used in the JRME over a five-year period from 1987 through 1991. Another objective of this study was to compare the results of this investigation with the results of Goodwin and Goodwin (1985), therefore allowing a direct comparison between two different educational research journals over a similar period of time.

Method and Results

Volumes 1987 through 1991, the five most recently published volumes of the JRME, were selected for the study. Each volume was published in four quarterly issues, and a total of 109 articles were used in the study. A 10-item summation chart was developed by the author based upon previous investigations (Edgington, 1974; Willson, 1980; Goodwin & Goodwin, 1985; Halpin & Halpin, 1988; Sedlmeier & Gigerenzer, 1989). The items used in the analysis of the articles were as follows: Power, Alpha, Effect Size, Sample Size, Basic

Procedures, Method of Computation, Basic Assumptions, References, Repeated Univariate, and Other (at-large anomalies). Results for the individual articles were grouped into issue totals, volume totals, and overall item totals. Only the volume totals and overall item totals will be reported in this study.

Of the 109 articles published in the JRME from 1987 through 1991, 78 were quantitative (71.56%) and 31 were non-quantitative (28.44%). The non-quantitative articles were classified as historical (14.68%), qualitative (4.59%), methodological (5.50%), and philosophical (3.67%).

The 78 quantitative articles contained many diverse types of statistical procedures. Over 48% of the quantitative articles employed more than one statistical technique per article. Some of the more frequently used procedures are shown in Table 1; analysis of variance and chi square were obviously the dominant statistical procedures.

Insert Table 1 about here

In the first two volumes analysis of variance and chi square were encountered in almost equal amounts. However, in the latter three volumes analysis of variance far outnumbered chi square to become the single dominant procedure. Together, analysis of variance and chi square comprised almost 50% of all procedures. Analysis of variance included one-way and factorial analysis of variance (ANOVA) and analysis of covariance (ANCOVA). Post hoc follow-up procedures included the Newman-Keuls tests, Scheffe tests, and Tukey tests. Other procedures included the Spearman's rank, Kruskal-Wallis one-way analysis of variance, Wilcoxin matched-pairs signed-ranks test, Mann-Whitney U test, z-test, Friedman's two-way analysis of variance, and Kendall's coefficient of concordance W. Multivariate procedures included multivariate analysis of variance (MANOVA), discriminant analysis, multivariate analysis of covariance (MANCOVA), and canonical correlation.

Perhaps the most surprising aspect of the study concerned the two items of Power and Effect Size. Of the 78 quantitative articles, 0 reported any type of power analyses or any mention of a power estimate or an estimate of effect size. In fact, the words "power" and "effect size" were never encountered anywhere in the context of Type II Error.

The item of Sample Size would have been more useful if some estimates of statistical power and effect size had been available. Therefore, regardless of power, the mean sample sizes for the articles are reported in Table 2.

Insert Table 2 about here

The smallest sample encountered was N=12 in 1991, and the largest sample found was N=1,843 in 1990. This giant sample is larger than the entire yearly sum of all sample sizes in 1989 (the 1989 total sample size was 1,734) and is part of an unusual study involving nine repeated one-way ANOVA procedures. If this enormous sample were left out the 1990 mean sample size would be 147.19 instead of 246.94, and the overall mean sample size would be 150.36 instead of 171.26.

The item of Alpha levels can also be seen in Table 2. The largest category of mixed alpha (51.28%) includes all articles using different alpha levels per or within each procedure. Many cases were found, however, where different alpha levels were stated; whichever alpha proved to be significant at the time was the one used. For example, some articles used alpha=.01 at times and then shifted to alpha=.05 when the results were apparently no longer significant at alpha=.01. As the results changed so did the alpha level. Since no justifications were given for these sudden changes in alpha levels, then one might assume that final significance, rather than the author, determines alpha. Yet, these changes in alpha could have been due to a habit of using .05 as the basic criterion alpha, then reporting a more stringent alpha if the later results allowed. However, in one instance an author began with alpha =.05 and then stated a relaxed significance criterion of alpha=.15 in a stepwise regression analysis to "determine the set of best predictors." Therefore, without any clear statements from the author, the reader cannot be sure of the alpha level used in the null hypothesis rejection.

The final portion of Table 2 concerns the item of References. No obvious trends are apparent: the total mean number of references per article is 25.04, and each yearly mean stays fairly close to the total mean. However, the very small number of statistical references per year is interesting. What should be considered an acceptable number of statistical references for a quantitative article? Certainly, the number should be more than 2 references out of 14 total articles, the reported summary for 1989. Additionally, several articles listed more than one statistical reference. For example, the 12 statistical references in 1990 were made in only seven articles. Also, out of the 29 total statistical references, 15 of them (over half) were for three repeated sources; the SPSS manual with 6 references, the SAS manual with 5 references, and Siegel's 1956 text of nonparametric procedures with 4 references.

Another interesting aspect of the study involved the item of Method of Computation. Of the eight articles using multiple regression six identified stepwise as the method of entering variables. The remaining two articles did not identify any method. All four articles using discriminant analysis identified stepwise as the computational method, and one article identifying unequal cell-size ANOVA identified no method of entering the variables. Therefore, 10 of 13 computational methods were all stepwise, and the remaining 3 were unidentified.

The item of Basic Assumptions provided more surprising results. In the case of chi square procedures, none of the five basic assumptions were directly addressed in any article using chi square. In the cases of *t*-test, ANOVA, ANCOVA, multiple regression, multivariate procedures, and other nonparametric procedures, the basic assumption most directly addressed was the Independence of Scores Assumption. Random sampling or random assignment was reported in 13 of the 78 quantitative articles (16.67%). Only one article out of the 78 quantitative articles researched directly addressed the assumptions of linearity and homoscedasticity.

The item of Repeated Univariate applied directly to the situations where authors used repeated univariate techniques, instead of one multivariate technique, from the same sample and set of independent variables. Out of the 78 statistical articles reviewed, 15 used these repeated univariate procedures, or 19.23%. There were several instances where authors would use two or three repeated ANOVA procedures instead of one MANOVA, and some articles even reported as many as five or six repeated ANOVA procedures. However, the largest anomaly occurred when one article reported nine repeated one-way ANOVA procedures, one for each different dependent variable. The reduction in power for such a procedure is calculated, according to Haase & Ellis (1987), at .8659. In other words, if power was originally at .995, then by the time of the last one-way ANOVA power would have collapsed to .1291. Also, the concurrent rise in the experiment-wise Type I error rate is calculated to be .37 by the time of the final ANOVA. Finally, in another article, an author reported 28 consecutive chi square procedures instead of a more concise log-linear analysis.

A final interesting point was noted in terms of the summary tables for critical ANOVA procedures. Out of the 46 ANOVA procedures, 8 neglected to include a summary table in the article (17.39%). In addition, in one article using factorial ANOVA, the degrees of freedom in the summary table for the interaction term did not seem to be correct. Correspondence with the author of this article did not clear up the issue, as he responded simply by mailing a photocopy of the computer printout of his ANOVA results. This vague reply seemed to imply that the computer understood more than the researcher.

After Goodwin & Goodwin (1985), Basic techniques are identified as one-way ANOVA, chi square, Pearson correlation, and *t*-test. Intermediate techniques are identified as planned and post hoc mean comparison, one-way ANCOVA, factorial ANOVA/ANCOVA, other procedures, and multiple regression. Advanced techniques are identified as discriminant analysis, canonical correlation, and one-way and factorial MANOVA/MANCOVA. Finally, for this study, Other techniques are added to include the Spearman's rank, Kruskal-Wallis one-way analysis of variance, Wilcoxin matched-pairs signed-ranks test, Mann-Whitney U test, z-test, Friedman's two-way analysis of variance, and Kendall's coefficient of concordance W.

The comparisons between this study and that of Goodwin & Goodwin (1985) can be seen in Table 3. The Intermediate and Other categories

---

Insert Table 3 about here

---

seem to be closely comparable between the two studies. The greatest differences occur in the categories of Basic and Advanced. In Goodwin & Goodwin (1985) Basic procedures were one-third of all procedures, while in the JRME Basic procedures are almost half of all procedures. Accordingly, every year except one saw Basic procedures as the dominant group of statistics (1991 saw Intermediate procedures slightly dominate). However, as noted above, many ANOVA procedures were categorized as unnecessarily repeated procedures. Out of the 27 total usages of basic ANOVA, 9 were repeated procedures (33.33%), and out of the 19 total usages of intermediate ANOVA, 6 were repeated procedures (31.58%). Therefore, if those repeated univariate procedures had actually been proper multivariate procedures, then the results for Basic, Intermediate, and Advanced in Table 3 would become 42.07%, 27.59%, and 19.31% respectively.

Discussion and Conclusions

The findings and results of this study are somewhat consistent with the findings of Goodwin & Goodwin (1985) that Basic, Intermediate, Advanced, and Other techniques are used at approximately the same rate in both educational journals, AERJ and JRME. Also, the findings of this study are consistent with the findings of Edgington (1974), Willson (1980), Goodwin & Goodwin (1985), and Halpin & Halpin (1988) that analysis of variance is currently the dominant statistical procedure in educational research. In addition, neither qualitative research nor nonparametric procedures are widely used. Also, this study is consistent with the findings of Sedlmeier & Gigerenzer (1989) that power and effect size are largely ignored in the behavioral sciences.

A thorough explanation of statistical power and effect size is well beyond the scope of this paper. However, music education researchers must acquire the habit of dealing with power and effect size. As Cohen (1988) warned, power reflects the degree to which a researcher can detect the treatment differences he expects and the chance that others will be able to duplicate his findings. In other words, low power is poor science, and experiments with low power do not produce reliable findings. Concurrently, experiments with power set too high can uncover an abundance of trivial significance and result in an unfortunate waste of resources, time, and money.

One of the ways of improving statistical power is by raising the sample size. As reported earlier, the mean sample for the JRME between 1987 and 1991 was N=171. These results are very encouraging, as other branches of the behavioral sciences seem to be using much smaller samples. Larger samples directly improve statistical power. Yet, as mentioned above, power can be too high. Music educators should continue the habit of using these impressive sample sizes, yet adopt the practice of verifying their effectiveness with a simple power analysis in the planning stages.

The sudden swerving of alpha levels within a procedure should be addressed. As mentioned earlier, the researchers might have been using an overall alpha=.05 followed by a report of any other findings which met a more stringent alpha as the study progressed. However, without any explanation from the researcher, the reader can never be sure. It is encouraging that several authors did include a precise justification of every alpha level chosen. In one instance an author noted, "Because of the large number of variables involved (N=14), .01 was chosen as the level of statistical significance." Statements such as this left absolutely no doubt as to the exact direction of the researcher's methods, results, and conclusions.

The over-abundance of stepwise regression is discouraging, yet not surprising. In multiple regression a researcher can enter his variables into the equation by using either the hierarchical, simultaneous (unique), or stepwise method, depending on the nature of his variables and the phrasing of his research questions. In stepwise regression whichever independent variable correlates the highest with the dependent variable is automatically selected first by the computer. This, in turn, burdens the first variable selected with all the shared variance between it and the other variables not selected. As soon as variables cease to be significantly correlated with the dependent variable the computer automatically stops selecting variables. Ultimately, the resultant F-test from this process cannot be trusted because the variables not selected by the computer are not included in the equation. Since the number of variables in the F-test has been falsely decreased, the researcher could be deceived into finding significance that does not exist (Type I error). In order to correctly use stepwise regression, the researcher should use large sample sizes, cross-validation, and seek technical, rather than explanatory, results. Generally, hierarchical regression accounts for all shared variance, while stepwise regression only accounts for any shared variance between the independent variables alone. Simultaneous regression does not account for any shared variance at all.

Unfortunately, the stepwise method is the default option in SPSS. Unless the researcher specifically programs the computer to do otherwise, SPSS will automatically use the stepwise model. For example, if the author's research questions mostly concern Variable A, yet Variable B correlates higher with the dependent variable, then SPSS will begin with Variable B. It must remain the author's responsibility to create a research design that will specifically address what he wants to know. The music education researcher cannot allow the computer printout to dictate the research design, and many of the problems mentioned here seem to originate from an unfamiliarity with advanced computer packages. Allowing these canned computer packages to do one's critical thinking is an important trap to avoid.

If an article mentions random assignment or random sampling the reader can be positive that the basic assumption of normality (Independence of Scores) has been addressed. As mentioned earlier, only 13 of the 78 quantitative articles reported randomization, and if this basic assumption is not addressed the reader can never be sure if this is from either a simple oversight or a failure to satisfy that basic assumption. More surprisingly, only one article addressed any of the other basic assumptions. Researchers can eliminate many questions about their choice of statistical design simply by including as much information as possible about their procedures. Although many articles went into tremendous detail about the brand names, model numbers, tape speeds, and control settings of various electronic equipment, they omitted any mention of the simple basic assumptions. A willingness to supply as much information as possible about the physical design of the study should also extend toward the statistical design as well. In one very effective example, an author included a small, one-paragraph section in his article called "Design and Analysis," where he clearly listed, explained, and (correctly) justified all statistical procedures used.

The same statistical procedures are available in all branches of the behavioral sciences, including music education research. Obviously, the same rules for manipulations of these techniques apply equally to all branches of the behavioral sciences. Yet, the inconsistencies in the research methodologies in music education research cannot be ignored. Music educators are rightfully proud of their place in the behavioral sciences and of their use of advanced statistical procedures. However, if they propose their research to be respected by all consumers of educational research, then the correction of these problems is imperative.

The music education researcher must be prepared and willing to offer justifications for all procedures. Faced with the realities and limitations of time and money, he must design research with adequate power. The music education researcher should be aware of different computational methods for entering variables into the equation and should purposefully select the best method to correctly answer the research questions. The researcher must be responsible for knowing when to customize his design and, therefore, bypass any inappropriate default options. As noted above, many of the problems mentioned here seem to originate from an unfamiliarity with advanced computer packages, and simply believing that the computer knows more than the researcher is never going to be an adequate justification for decisions involving statistical procedures. Computers and statistics were invented for use by humans, not the other way around.

TABLE 1

Frequencies and Percentages of All Statistical Procedures Used in the JRME by Year, 1987-1991

	Year					Totals	
	1987	1988	1989	1990	1991	No.	%
No. of statistical articles	11	12	14	19	22	78	
No. of procedures	16	23	24	45	37	145	
Analysis of variance	4	4	11	13	16	48	33.10%
One-way ANOVA	3	3	6	8	8		
Factorial ANOVA	1	1	4	4	8		
ANCOVA	0	0	1	1	0		
Chi square	1	6	6	7	4	24	16.55%
Post hoc	3	0	3	7	5	18	12.41%
Newman-Keuls	2	0	2	4	3		
Scheffe	1	0	0	2	1		
Tukey	0	0	1	1	1		
Other	2	7	0	3	4	16	11.03%
Spearman rank	2	1	0	0	1		
Kruskal-Wallis ANOVA	0	1	0	1	1		
Wilcoxon matched pairs	0	1	0	1	0		
Mann-Whitney U test	0	2	0	0	0		
z-test	0	1	0	0	1		
Friedman ANOVA	0	1	0	0	0		
Kendall's coefficient of concordance W	0	0	0	1	1		
Multivariate	3	0	1	6	3	13	8.97%
MANOVA	0	0	1	3	2		
Discriminant analysis	2	0	0	1	1		
MANCOVA	1	0	0	1	0		
Canonical correlation	0	0	0	1	0		
t-test	2	2	3	3	0	10	6.90%
Multiple regression	0	1	0	4	3	8	5.52%
Correlation	1	3	0	2	2	8	5.52%

TABLE 2

Frequencies, Means, and Percentages of Sample Sizes, Alpha, and References

	Year					Totals	
	1987	1988	1989	1990	1991	No.	%
No. of statistical articles	11	12	14	19	22	78	
Mean sample size per article	204.9	178.3	115.6	246.9	130.9	171.3	
Alpha							
.05	5	7	4	6	8	30	38.46%
.01	1	0	1	1	5	8	10.26%
Mixed	5	5	9	12	9	40	51.28%
References							
General	218	328	321	514	543	1924	
Statistical	4	6	2	12	5	29	
Mean references per article	20.18	27.83	23.07	27.68	24.91	25.04	

TABLE 3

Comparison of Use of Statistics by Type

Source	Basic	Intermediate	Advanced	Other
Goodwin & Goodwin, (1985)	33.33%	36.53%	17.12%	13.02%
JRME, (1992)	48.28%	31.72%	8.97%	11.03%
1987	7	4	3	2
1988	14	2	0	7
1989	15	8	1	0
1990	20	16	6	3
1991	14	16	3	4
Totals	70	46	13	16

References

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. Journal of Abnormal and Social Psychology, 65, 145-153.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J., & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Edgington, E. S. (1974). A new tabulation of statistical procedures used in APA journals. American Psychologist, 29, 25-26.

Goodwin, L. D., & Goodwin, W. L. (1985). Statistical techniques in AERJ articles, 1979-1983: The preparation of graduate students to read the educational research literature. Educational Researcher, 14, 5 11.

Haase, R. F., & Ellis, M. V. (1987). Multivariate analysis of variance. Journal of Counseling Psychology, 34, 404-413.

Hair, J. F., Anderson, R. E., & Tatham, R. L. (1987). Multivariate data analysis with readings (2nd ed.). New York: Macmillan.

Halpin, G. & Halpin, G. (1988). Evaluation of research and statistical methodologies. Paper presented at the meeting of the Mid-South Educational Research Association, Louisville, KY.

Keppel, G. (1991). Design and analysis: A researcher's handbook (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? Psychological Bulletin, 105, 309-316.

Thompson, B. (1988). Common methodology mistakes in dissertations: Improving dissertation quality. Paper presented at the meeting of the Mid-South Educational Research Association, Louisville, KY.

Willson, V. L. (1980). Research techniques in AERJ articles: 1969 to 1978. Educational Researcher, 9, 5-10.