

Spring 2013

A Methodological Review of the Articles Publishes in Georgia Educational Researcher from 2003-2010

Justus J. Randolph

Mercer University, randolph_jj@mercer.edu

Andrea E. Griffin

Mercer University

Samara R. Zeiger

Mercer University

Kristina N. Falbe

Mercer University

Noreen A. Freeman

Mercer University

See next page for additional authors

Follow this and additional works at: <https://digitalcommons.georgiasouthern.edu/gerjournal>



Part of the [Education Commons](#)

Recommended Citation

Randolph, Justus J.; Griffin, Andrea E.; Zeiger, Samara R.; Falbe, Kristina N.; Freeman, Noreen A.; Taylor, Bridget E.; Westbrook, Amy F.; Lico, Cheryl C.; Starling, Cristy N.; Sprull, Nakiesha M.; Holt, Carolyn; Smith, Kristie; and McAnespie, Hannah (2013) "A Methodological Review of the Articles Publishes in Georgia Educational Researcher from 2003-2010," *Georgia Educational Researcher*. Vol. 10: Iss. 1, Article 1.

DOI: 10.20429/ger.2013.100101

Available at: <https://digitalcommons.georgiasouthern.edu/gerjournal/vol10/iss1/1>

This quantitative research is brought to you for free and open access by the Journals at Georgia Southern Commons. It has been accepted for inclusion in Georgia Educational Researcher by an authorized administrator of Georgia Southern Commons. For more information, please contact digitalcommons@georgiasouthern.edu.

A Methodological Review of the Articles Published in Georgia Educational Researcher from 2003-2010

Abstract

Methodological reviews, reviews that concentrate on research methods rather than research outcomes, have been used in a variety of fields to improve research practice, inform debate, and identify islands of practice. In this article, we report on the results of a methodological review of all of the articles published in Georgia Educational Researcher from 2003-2010. We examined the methodological characteristics, authorial characteristics, and methodological quality of those articles using quantitative content analysis. The major findings were that (a) the proportions of the type and traditions of articles published in Georgia Educational Researcher were similar to the proportions in education research articles in general, (b) case study research and correlational research were most prominent, (c) a few universities accounted for most of the articles published, (d) male and female authors' articles were published in equitable proportions, and (e) there were no statistically significant differences in methodological quality between genders, university affiliations, types of research, or years of publication. We end with a few suggestions for improving the quality of qualitative research articles in the Georgia Educational Researcher.

Keywords

Methodological review, Methodological quality, Systematic review

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Authors

Justus J. Randolph, Andrea E. Griffin, Samara R. Zeiger, Kristina N. Falbe, Noreen A. Freeman, Bridget E. Taylor, Amy F. Westbrook, Cheryl C. Lico, Cristy N. Starling, Nakiesha M. Sprull, Carolyn Holt, Kristie Smith, and Hannah McAnespie

**A METHODOLOGICAL REVIEW OF THE ARTICLES PUBLISHED IN
GEORGIA EDUCATIONAL RESEARCHER FROM 2003-2010**

Justus J. Randolph, Andrea E. Griffin, Samara R. Zeiger, Kristina N. Falbe, Noreen A. Freeman, Bridget E. Taylor, Amy F. Westbrook, Cheryl C. Lico, Cristy N. Starling, Nakiesha M. Sprull, Carolyn Holt, Kristie Smith, and Hannah McAnespie

Tift College of Education
Mercer University
Atlanta, GA

Authors' Note: We would like to acknowledge the work of Marcia Watson and Tida Bruce-Banfield, whom helped with the coding of one or more articles. Thanks to Kobe and Kyle Griffin for their patience and support. A previous version of the paper was presented at the 2011 annual meeting of the Georgia Educational Research Association Conference, Savannah, GA. Correspondence concerning this article should be directed to Justus Randolph, randolph_jj@mercer.edu

Abstract: Methodological reviews, reviews that concentrate on research methods rather than research outcomes, have been used in a variety of fields to improve research practice, inform debate, and identify islands of practice. In this article, we report on the results of a methodological review of all of the articles published in *Georgia Educational Researcher* from 2003-2010. We examined the methodological characteristics, authorial characteristics, and methodological quality of those articles using quantitative content analysis. The major findings were that (a) the proportions of the type and traditions of articles published in *Georgia Educational Researcher* were similar to the proportions in education research articles in general, (b) case study research and correlational research were most prominent, (c) a few universities accounted for most of the articles published, (d) male and female authors' articles were published in equitable proportions, and (e) there were no statistically significant differences in methodological quality between genders, university affiliations, types of research, or years of publication. We end with a few suggestions for improving the quality of qualitative research articles in the *Georgia Educational Researcher*.

Keywords: methodological review, methodological quality, systematic review

Introduction

Methodological reviews, reviews that focus on research methods rather than research outcomes, have been used in many fields to improve research practice, inform debate, and identify islands of practice (Keselman et al., 1998; Ranis & Walters, 2004; Wilkinson & APA Task Force on Statistical Inference, 1999). For example, Keselman et al. (1998) conducted a methodological review of education researchers' statistical practices. Subsequently, that work helped inform the guidelines of the APA Task Force on Statistical Inference's influential report *Statistical Methods in Psychology Journals: Guidelines and Explanations* (Wilkinson & APA Task Force on Statistical Inference, 1999), which went on to inform numerous authors, editors, and reviewers in best practices for statistical reporting. In addition, the Social Science Research Council and the National Academy of Education's Joint Committee on Educational Research documented a need for ". . . data and analysis of the research enterprise, . . . determination of where education research is conducted and by whom, [and]. . . . identification of the range of problems addressed and the methods used to address them" (Ranis & Walters, 2004, pp. 798-799).

Because the *Georgia Educational Researcher* (GER) has not had a methodological review published about it thus far, the purpose of this research is to identify and quantify the types of articles published in GER and to review the methods used in the qualitative, quantitative, and mixed-methods studies. The expected benefits of this study include improved research practice through the identification of methodological strengths and weakness and improved understanding of the "trends, tribes, and territories" of researchers who publish in GER. The target audiences for this article are the readers of GER, GER editors and reviewers, and authors planning on submitting manuscripts to GER.

The research questions are listed below:

1. What are the overall methodological characteristics of the articles published in GER?
2. What are the overall characteristics of the authors of GER articles? (Those authorial characteristics included gender and affiliation.)
3. What are the characteristics of the methodological quality of the articles published in GER?
4. What variables, if any, are statistically significantly associated with the methodological quality of the articles published in GER?

In the following sections of this manuscript, we discuss related methodological reviews, the methods we used to carry out this investigation, the results, and end with a discussion of our findings.

Related Research

While there have been many methodological reviews of the educational research literature over the years (e.g., Keselman et al., 1998; Randolph, 2008), we concentrate on what we consider to be the most comprehensive review to date—Gorard and Taylor (2004). In that review, Gorard and Taylor reviewed a sample of 94 articles from leading education research journals. They reviewed 42 articles from the *British Educational Research Journal*, 24 articles from the *British Journal of Educational Psychology*, and 24 articles from *Educational Management and Administration*. They further validated their findings with

- interviews with key stakeholders from across the education field; including researchers, practitioner representatives, policy makers and policy implementers;
- a large-scale survey of the current methodological expertise and future training needs of UK education researchers; [and a]
- detailed analysis and breakdown of 2001 RAE [Research Assessment Exercise]. (p. 114)

The findings from the Gorard and Taylor (2004) study that are most relevant to the current study relate to the proportions of articles that were classified as (a) empirical research with human participants and (b) nonempirical research (such as theoretical articles) or secondary research (such as literature reviews). Also of relevance are the proportions of articles that were classified as quantitative, qualitative, or mixed methods. We used those proportions as a point of comparison for the proportions we found in GER articles. Gorard and Taylor's proportions are presented in Table 1. In essence, over 80% of education research articles reported on empirical research with human participants. Of those articles, there were about equal amounts of quantitative and qualitative articles and there were very few mixed-methods articles.

Table 1
Characteristics of Education Research Articles from the Gorard and Taylor Review

<i>Characteristic</i>	<i>Count</i>	<i>%</i>
Type of article		
Empirical research with human participants	79	84
Nonempirical research or secondary research	15	16
Tradition of article, if empirical		
Quantitative	43	54
Qualitative	32	41
Mixed	4	5

While it has been shown that these proportions can vary over subdisciplines of education and even over the region of the first author's affiliation (see Randolph, 2008), we assume that these proportions are representative of the education research literature in the mid 2000s and across geographic areas. Therefore, we used these proportions as a point of reference to compare to our own.

Methods

Neuendorf's (2002) method of quantitative content analysis was used in this review. That method consists of the following steps:

1. Developing a theory and rationale
2. Conceptualizing variables
3. Operationalizing measures
4. Developing a coding form and coding book
5. Sampling
6. Training and determining pilot reliabilities
7. Coding
8. Calculating final reliabilities
9. Analyzing and reporting data

In the remainder of this section, we provide information on sampling, training and determining pilot reliabilities, coding, calculating final reliabilities, and analyzing and reporting data. Because a pre-existing coding form was used, we do not report in detail on the first four steps of Neuendorf's method; however, we do provide information on the coding forms used. In essence, the coding forms were checklists adapted from a textbook on educational research methods (Creswell, 2012). The checklists were created to help the readers evaluate the quality and process for qualitative, quantitative, and mixed-method research. Creswell (2002) implies that the quantitative checklist was adapted from Tuckman (1999) and the qualitative checklist was adapted from Stake (1995).

Coding Forms

Two coding forms for reviewing the methodological qualities of quantitative and qualitative articles were used; they are presented in Appendices A and B. These forms were adapted from Creswell (2012, p. 291 & p. 292, respectively). The quantitative coding form originally contained 44 questions in areas related to (a) the title of the study; (b) the problem statement; (c) review of the literature; (d) purpose, hypotheses, and research questions; (e) data collection; (f) data analyses and results; (g) writing; (h) internal validity; and (i) external validity. However, because of poor reliabilities, the internal validity items were not analyzed in this manuscript. See the section on interrater reliability for more information. In all of the items except for the items dealing with external validity, the raters were asked to indicate whether they agreed or disagreed with the item. In the external validity section, raters were asked to write in short answers. The qualitative coding form had 32 items in the areas of (a) title for the study, (b) problem statement, (c) review of the literature, (d) purpose and research questions, (e) data collection, (f) data analysis and findings, and (g) writing.

Sampling

The sample for this study included all of the GER articles published from Volume 1, Issue 1 in the fall of 2003 to Volume 8, Issue 1 in the spring of 2010. Technically, our selection of articles was a census and not a sample because the universe of articles was

reviewed. In total, 42 articles were included in this review. Transcriptions of keynote addresses were not included in this review.

Training Raters and Determining Pilot Reliabilities

The raters for this review were 15 students in an introductory educational research course in a doctoral program in curriculum and instruction. During the course, the raters were given approximately 45 hours of instruction over an eight-week period in content matter related to the items on quantitative and qualitative coding forms. Explicit instruction was also given on how to use the coding forms to code the articles. To determine pilot reliabilities, each rater was assigned one or more quantitative or qualitative articles to pilot code. After initially piloting the coding system, the raters and instructor came together to discuss and clarify the coding form items on which there was confusion.

Calculating Final Reliabilities

In addition to the articles that the raters individually rated, there was one common article (either quantitative or qualitative) that multiple raters coded to assess the interrater reliability of their ratings. The measure of interrater reliability used in this article was a multirater variation of Bennet et al.'s free-marginal kappa statistic (Randolph, 2005; Warrens, 2010).

Data Collection and Analyses

Each rater was randomly assigned a set of one or more articles to rate and used the quantitative and/or qualitative coding forms to code the data. Mixed-methods articles were coded on both their quantitative and qualitative characteristics.

Descriptive statistics were calculated for the articles' authorial and methodological characteristics. To compare the proportions found in this article with the proportions found in the Gorard and Taylor (2004) review, χ^2 analyses were used.

Factorial ANCOVA, using the technique described in Field (2009), was used to determine which variables were statistically significantly associated with methodological quality of the qualitative articles. All of the assumptions for factorial ANCOVA had been met; namely, a visual analysis of a residual plot showed homoscedasticity and had no influential data points. Also, Levene's test indicated homogeneity of variances, $F(8, 2) = 0.55, p = .778$. Methodological quality was used as the outcome variable. Affiliation, qualitative tradition, and gender were used as fixed factors (i.e., categorical variables) and year of publication was used as a covariate (i.e., a continuous variable). We used year of publication as a covariate to determine if there were trends over time.

The data for quantitative quality did not meet the assumptions for factorial ANCOVA; therefore, nonparametric statistics were used instead. The Mann-Whitney test was used to determine if there were differences in the methodological quality of articles written by

male and female first authors. The Kruskal-Wallis test was used to examine differences among groups on the affiliation, type of article, and year of publication variables.

Results

Interrater Reliabilities

Nine raters were randomly assigned to rate the same qualitative article using the qualitative coding form, which had 30 items, each of which had two categories. The nine raters' percent of overall agreement on those 30 items was 83% and the kappa value was .63, indicating fair agreement above chance. A rule of thumb is that values of kappa above .70 indicate good agreement (Neuendorf, 2002).

Nine raters were randomly assigned to read the same quantitative article, using the quantitative coding form, which had 36 items, each of which had two categories. Of those nine raters, only four completed their ratings without missing data. Since kappa cannot be calculated with missing data, those raters' ratings were not included. Of those four raters, there was 100% agreement on those four items, which equates to a kappa value of 1.00. When including two raters who were only missing one or two items and then calculating kappa and percent of overall agreement per item then averaging the values, the value of kappa was .90 and the percent of overall agreement was 95%.

One complication was that there were eight items on the quantitative coding form that dealt with internal validity and were applicable to experimental articles only; however, raters sometimes completed ratings of internal validity for correlational and descriptive articles as well. This demonstrates a serious lack of reliability on those variables, so an analysis of the internal validity of the quantitative articles is excluded from this manuscript.

Methodological Characteristics

Table 2 presents the methodological characteristics of the 34 empirical GER articles included in this review. Note that the six mixed-methods articles were coded using both the quantitative and qualitative coding forms, so quantitative coding was done on 23 articles and qualitative coding was done on 17 articles.

Table 2
Methodological Characteristics of GER articles

Methodological Characteristic	Count	%
Type		
Empirical	34	81
Literature Review	2	5
Other	6	14
If empirical, what tradition?		
Quantitative	17	50
Qualitative	11	32
Mixed	6	18
If qualitative, what tradition?		
Case Study	5	39
Ethnography	1	7
Phenomenology	3	23
Grounded Theory	2	15
Narrative/ Biography	2	15
(Could not be determined)	(4)	(N/A)
If quantitative, what type?		
Experimental	7	30
Correlational	11	48
Descriptive	5	22

In terms of the overall type of article, the overwhelming majority of articles were empirical (81%). There was not a statistical difference in the proportion of empirical GER articles and the proportion of empirical education research articles reviewed by Gorard and Taylor (2004), $\chi^2(2) = 0.20, p = .657$. Of the empirical articles, quantitative articles were most prominent, followed by qualitative articles, then mixed-methods articles.

In terms of the tradition used, the proportion of quantitative, qualitative, and mixed-methods GER articles was not statistically different from the proportions in the Gorard and Taylor review, $\chi^2(2) = 4.76, p = .093$. Although there was not a statistically significant difference at the .05 α level, mixed-methods articles composed 18% of GER articles, while mixed-methods articles only composed 5% of the articles in the Gorard and Taylor review.

In terms of the qualitative tradition used, case study research was the most commonly used qualitative tradition in GER articles. In four of the qualitative articles, we were not able to determine what qualitative tradition was used. We consider this to be an indicator of poor methodological quality in these articles.

Finally, in terms of the type of quantitative method used, articles that used a correlational method were published most often, followed closely by experimental articles. Purely descriptive quantitative articles, such as reports of survey findings, were published less frequently.

Types of participants examined. Of the 20 articles in which the participants were clearly defined, students were the participants most frequently examined; they were the participants in 10 (50%) of the articles. Faculty and staff were the participants in four of the articles (20%). Administrators were participants in two of the articles (10%). School districts (5%) and parents (5%) each had one study in which they were the participants. There was a mixed study that included hiring decision makers, teacher, community members, and parents (5%) and another mixed study that included both students and parents (5%).

Types of interventions examined. Of the 20 articles with interventions, four (20%) of the articles dealt with literacy. Three (15%) dealt with teacher certification. The remaining 13 (65%) articles dealt with interventions that no other article dealt with. The topic of those interventions ranged from single-gender classrooms to strategies to increase enrollment.

Types of outcomes examined. Out of the 20 articles where the outcomes were clearly specified, academic achievement was the outcome in 12 (60%) of them. The other articles dealt with a variety of outcomes, from perceptions of alternative certification to better models for course development.

Types of settings examined. The research done in the GER articles occurred in a variety of settings and grade levels. Of the articles in which the authors mentioned their setting, five were set in urban schools, two in rural schools, and two in suburban schools.

Authorial Characteristics

Table 3 presents information on the gender of the first author, the first author's affiliation, and the number of GER articles published by year. Note that these data only represent the 34 empirical articles. The *other* category in the affiliation section included all of the universities that only contributed one article. In summary, there was about an equal percentage of female and male authors; the first author's first affiliations were most often Valdosta State University, Georgia Southern University, and Georgia State University; and the number of articles published per year was more or less constant with an increase in 2010.

Table 3
Authorial Characteristics of GER Articles

<i>Characteristic</i>	<i>Count</i>	<i>%</i>
Gender		
Female	15	48
Male	16	52
(Indeterminate)	(4)	
Affiliation		
Valdosta State University	8	24
Georgia State University	7	21
Georgia Southern University	5	15
Other	14	41
Year		
2003	5	15
2004	3	9
2005	2	6
2006	5	15
2007	4	12
2008	3	9
2009	5	15
2010	7	21

Methodological Quality of Quantitative Articles

Table 4 presents the results of each item on the quantitative coding form, which is included in Appendix A. The average of all of the items across the 17 quantitative and six mixed methods articles was 82.14% with 95% confidence intervals of 74.78 and 89.50.

Table 4
Ratings for Each Item on the Quantitative Coding Form

<i>Item</i>	<i>% Yes (Count)</i>	<i>95% CIs</i>
1	81 (17)	[63, 99]
2	82 (17)	[64, 99]
3	59 (13)	[37, 81]
4	95 (21)	[86, 100]
5	91 (20)	[78, 100]
6	95 (21)	[86, 100]
7	100 (21)	[86, 100]
8	88 (15)	[69, 100]
9	75 (14)	[51, 99]
10	88 (18)	[69, 100]
11	71 (16)	[50, 92]

12	76 (17)	[56, 96]
13	62 (13)	[39, 85]
14	81 (18)	[63, 99]
15	76 (16)	[56, 96]
16	85 (17)	[68, 100]
17	75 (17)	[54, 96]
18	70 (15)	[48, 92]
19	80 (18)	[61, 99]
20	86 (18)	[69, 100]
21	91 (10)	[71, 100]
22	92 (21)	[73, 100]
23	67 (16)	[35, 98]
24	75 (17)	[46, 100]
25	92 (21)	[73, 100]
26	67 (15)	[45, 89]
27	90 (21)	[77, 100]
28	67 (16)	[45, 89]
29	90 (20)	[77, 100]
30	86 (19)	[69, 100]
31	95 (21)	[86, 100]
32	100 (21)	--

Note. We used 100 as the upper bound of the confidence interval.

Figure 1 shows the means and 95% confidence intervals when the items were averaged within sections. All sections except for one seemed to have more or less the same proportion of quantitative quality scores; the quality of problem statements in quantitative and mixed-methods GER articles stood out above the other sections.

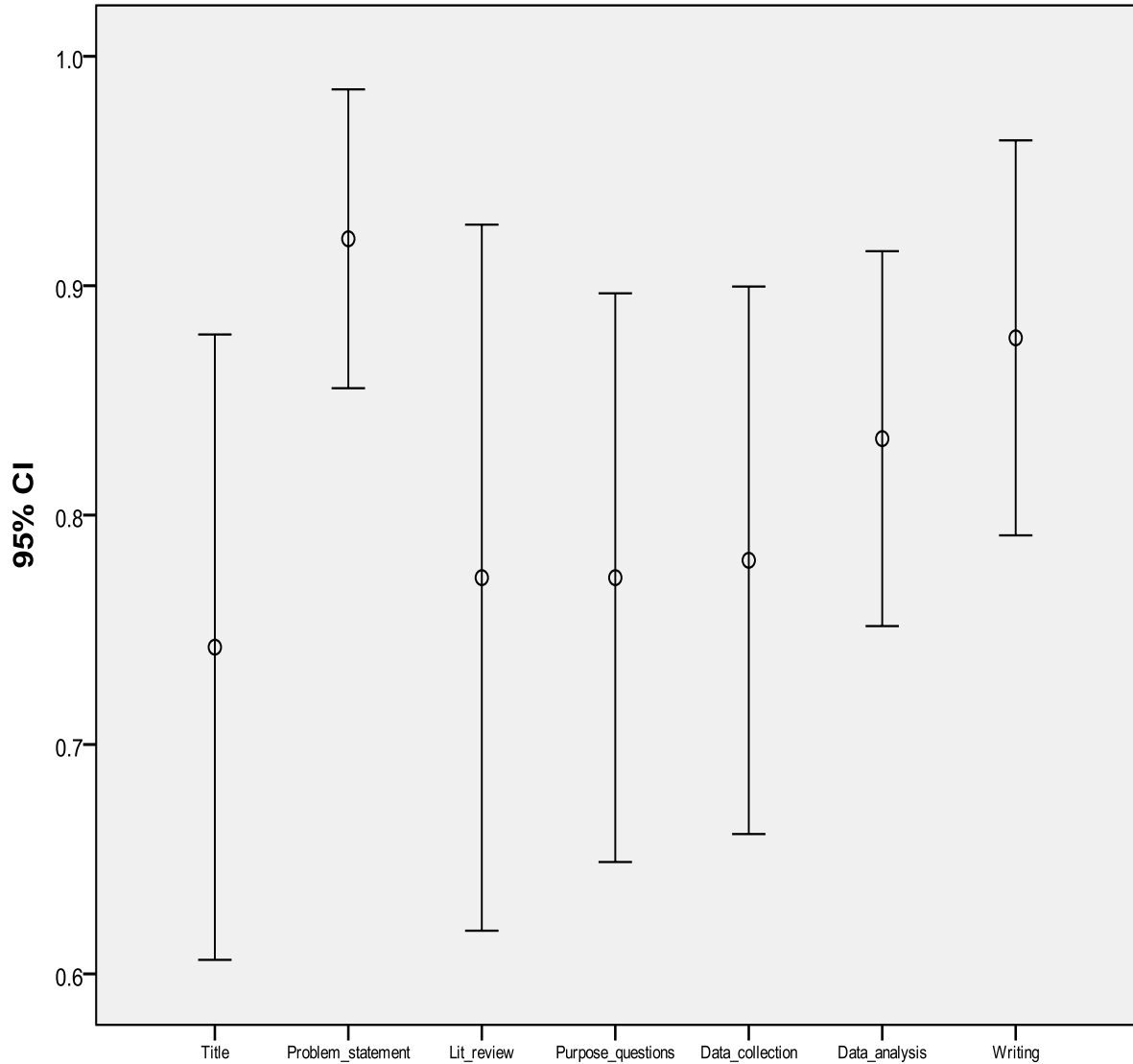


Figure 1. Average methodological quality of quantitative articles by section.

Methodological Quality of Qualitative Articles

Table 5 presents the ratings for each item on the qualitative coding form. Averaging all of the items together yielded a mean qualitative quality score of 77.82 with 95% confidence intervals of 69.67 and 85.97.

Table 5
Ratings for Each Item on the Qualitative Coding Form

<i>Item</i>	<i>% Yes (Count)</i>	<i>95% CIs</i>
1	94 (16)	[82, 100]
2	35 (6)	[10, 61]
3	94 (17)	[83, 100]
4	100 (17)	--
5	89 (16)	[73, 100]
6	94 (17)	[83, 100]
7	88 (14)	[69, 100]
8	89 (16)	[73, 100]
9	39 (7)	[14, 64]
10	83 (15)	[64, 100]
11	72 (13)	[49, 95]
12	61 (11)	[36, 86]
13	35 (6)	[10, 61]
14	76 (13)	[54, 99]
15	75 (12)	[51, 99]
16	78 (14)	[57, 99]
17	67 (12)	[43, 91]
18	89 (16)	[73, 100]
19	83 (15)	[64, 100]
20	94 (17)	[83, 100]
21	83 (15)	[64, 100]
22	83 (15)	[64, 100]
23	67 (12)	[43, 91]
24	82 (14)	[62, 100]
25	100 (17)	--
26	88 (15)	[71, 100]
27	88 (15)	[71, 100]
28	47 (8)	[21, 74]
29	47 (8)	[21, 74]
30	100 (16)	--

Note. We used 100 as the upper bound of the confidence interval.

Figure 2 shows the average qualitative methodology quality scores by each section on the qualitative coding form. Similar to the results shown in Figure 1, all of the sections have more or less the same quality scores except for the problem statement, which stands out above the other sections.

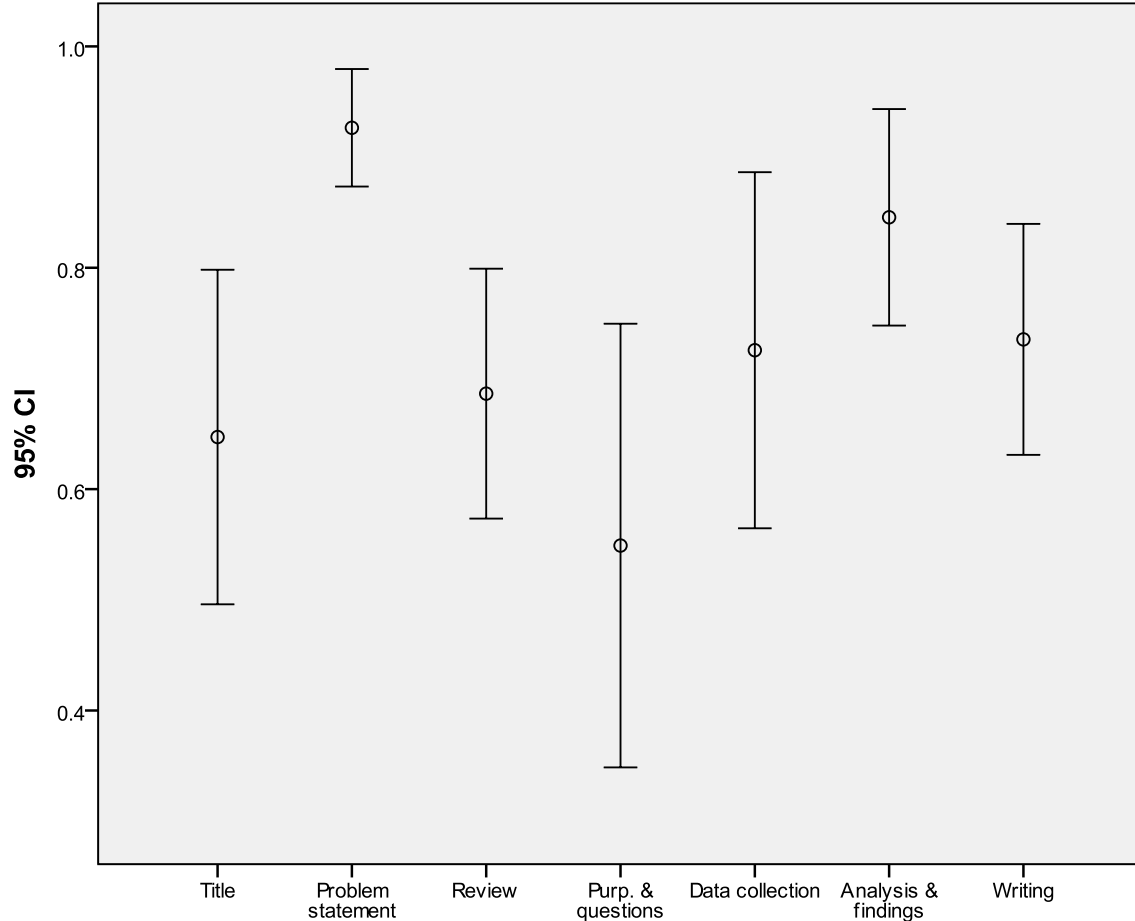


Figure 2. Average methodological quality of qualitative articles by section.

Methodological Quality

In terms of quantitative methodological quality, the results indicated that neither gender, $U(21) = 53.50, z = -.04, p = .972$; nor affiliation, $H(3) = 3.91, p = 2.71$; nor type of article (i.e., experimental, correlational, or descriptive), $H(2) = 2.75, p = 2.51$; nor year of publication, $H(7) = 7.42, p = .386$, were statistically significantly associated with methodological quality. Similarly, there were no statistically significant associations in terms of the methodological quality of the qualitative articles; gender, $F(1, 2) = 0.72, p = .486, \eta^2 = .26$; affiliation, $F(2, 2) = 0.12, p = .891, \eta^2 = .11$; qualitative tradition, $F(3, 2) = 0.49, p = .724, \eta^2 = .42$; year of publication, $F(1, 2) = 0.12, p = .760, \eta^2 = .06$.

Discussion

Summary of Results

Methodological characteristics. In terms of methodological characteristics, the proportion of types and traditions of articles in GER was very similar to the proportion of types and traditions of articles in education research articles in general, if we assume that the Gorard and Taylor (2004) review is representative of education research articles in general. The majority of articles reported on empirical research with human participants. Of those articles, quantitative articles made up the majority, followed by qualitative and mixed-methods articles, in that order. Although the difference was not statistically significant, there was a 13% difference in the percentage of mixed-methods articles between GER articles and education research articles in general. It is difficult to determine if this is because the Gorard and Taylor review was conducted in 2004 and mixed-methods articles have gained popularity since then, if it was because their review articles were primarily drawn from British journals, or if the difference was simply a characteristic of GER articles. Of the quantitative articles, correlational research was most prevalent, followed by experimental and descriptive research. Of the qualitative articles, case study research was the most prevalent qualitative tradition used. In four of the 11 (36%) qualitative articles, the qualitative tradition could not be determined.

Authorial characteristics. In terms of the affiliations of first authors, Valdosta State University was the current host institution of GER at the time of writing, so it is no surprise to us that it is also the institution with the most GER articles. The institution that is represented in GER second most frequently is Georgia Southern University; we hypothesize that this is the case because Georgia Southern University hosts the GERA conference, which is often a source of GER articles.

In terms of gender, there was about an equal number of male and female first authors in GER articles. Although we could not find a reliable statistic indicating the percentage of female education professors in Georgia, we were able to find a reliable statistic indicating that, in 2009, 46% of professors in the U.S. were female (United Nations Statistics Division, 2011). A binomial test indicated that there was not a statistically significant difference in the proportion of female first authors of GER articles (48.39%) and the proportion of female professors in the United States (46.00%), $p = .464$. Therefore, we take this to be evidence of a lack of editorial gender bias. The number of articles published each year has remained steady from 2003 to 2010, with the exception of a small spike in 2010.

Methodological quality. Using the adaptations of Creswell's (2012) checklists for evaluating quantitative and qualitative research revealed some strengths and weakness in the GER articles. For both quantitative and qualitative articles, problem statements were the sections that had the highest ratings of methodological quality. While there were no areas that particularly stood out as weaknesses in the quantitative articles, the purpose statements and research questions of qualitative articles stood out as an area that received low ratings. Another area of weakness was that in over a third of the qualitative articles,

the authors did not make explicit the qualitative tradition to which they were adhering. Methodological quality was about the same between genders, affiliation, year of publication, and type of article.

Recommendations

Methodological reviews can be especially impactful when it is found that research practice deviates sharply from what is considered to be best research practice. This was not the case here. Therefore, the only recommendations we have to improve practice are few and minor. First, we believe that authors of qualitative research should explicitly state what qualitative tradition they are using. Second, Creswell (2012) has scripts for writing high-quality qualitative research questions and purpose statements and we believe that it would be beneficial for authors to use those scripts to clarify their questions. These suggestions might help GER authors, reviewers, and editors improve the methodological and reporting quality of the papers that are published. Besides improving practice, we hope the information presented here can help codify our understandings of what are the trends, issues, and methods being published in GER. In turn, we hope that those increased understandings can answer the call put forth for more research on educational research as articulated by the Social Science Research Council and the National Academy of Education's Joint Committee on Educational Research (Ranis & Walters, 2004).

Of course, we acknowledge that Creswell's criteria for evaluating qualitative and quantitative are just one of many sets of valid criteria for the evaluation of methodological quality. There is a great diversity in education research, especially in the qualitative tradition, and we acknowledge that a study does not have to meet Creswell's criteria to be a high quality study.

References

- Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Boston, MA: Pearson.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.
- Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R., Donahue, B., . . . Levin, J. R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research*, 68(3), 350-386.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Los Angeles, CA: Sage.
- Gorard, S., & Taylor, C. (2004). *Combining methods in educational and social research*. Berkshire, United Kingdom: Open University Press.
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.

- Randolph, J. J. (2005). Free-marginal multirater kappa: An alternative to Fleiss' fixed-marginal multirater kappa. Paper presented at the *Joensuu University Learning and Instruction Symposium 2005*, Joensuu, Finland. (ERIC Document Reproduction Service No. ED490661)
- Randolph, J. J. (2008). *Computer science education research at the crossroads: A methodological review of the computer science education research*. Saarbrücken, Germany: VDM Verlag.
- Ranis, S. H., & Walters, B. H. (2004). Education research as a contested enterprise: The deliberations of the SSRC-NAE Joint Committee on Education Research. *European Education Research Journal*, 3(4), 795-806.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Tuckman, B. W. (1999). *Conducting educational research* (5th ed.). Ft. Worth, TX: Harcourt Brace.
- United Nations Statistics Division. (2011). *Statistics and indicators on men and women: Teaching staff* [Dataset]. Retrieved from <http://unstats.un.org/unsd/demographic/products/indwm/tab4e.htm>
- Warrens, M. J. (2010). Inequalities between multi-rater kappas. *Advances in Data Analysis and Classification*. Advance online publication. doi:10.1007/s11634-010-0073-4
- Wilkinson, L., & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594-604.

Appendix A: Coding Form for Quantitative Studies (Adapted from Creswell, 2012)

Checklist for Evaluating the Process of a Quantitative Study (Check if yes)	
Title for the Study	
1	Does it reflect the major independent and dependent variables?
2	Does it express either a comparison among groups or a relationship among variables?
3	Does it convey the participants and site for the study?
Problem Statement	
4	Does it indicate an educational issue to study?
5	Has the author provided evidence that this issue is important?
6	Is there some indication that the author located this issue through a search of past literature or from personal experiences?
7	Does the research problem fit a quantitative approach?
8	Are the assumptions of the study consistent with an approach?
Review of the Literature	
9	Are the studies about the independent and dependent variables clearly reviewed?
10	Does the review end with how the author will extend or expand the current body of literature?
11	Does the study follow APA style?
Purpose, Hypotheses, and Research Questions	
12	Does the author specify a purpose statement?
13	Is the purpose statement clear, and does it indicate the variables, their relationship, and the people and site to be studied?
14	Are either hypotheses or research questions written?
15	Do these hypotheses or questions indicate the major variables and the participants in the study?
16	Do the purpose statement and hypotheses or research questions contain the major components that will help a reader understand the study?
17	Has the author identified a theory or explanation for the hypotheses or questions?
Data Collection	
18	Does the author mention the steps taken to obtain access to people and sites?
19	Has the author identified good, valid, and reliable instruments to use to measure the variables?
20	Are the instruments administered so that bias and error are not introduced into the study?
Data Analysis and Results	
21	Are the statistics chosen for analysis consistent with the research questions hypotheses, variables, and scales of measurement?
22	Is the unit of analysis appropriate to address the research problem?
23	Are the data adequately represented in tables and figures?
24	Do the results answer the research questions and address the research problem?
25	Are the results substantiated by evidence?
26	Are generalizations from the results limited to the population of participants in the study?
Writing	
27	Is the structure of the overall study consistent with the topics addressed in a quantitative study?
28	Are educational and social science terms carefully defined?
29	Are variables labeled in a consistent way throughout the study?
30	Is the study written using extensive references?
31	Is the study written using an impersonal point of view?
32	Is the study written appropriately for intended audiences?
External Validity	
33	What units does this study generalize to?
34	What treatments does this study generalize to?
35	What outcomes does this study generalize to?
36	What settings does this study generalize to?

Appendix B: Coding Form for Qualitative Studies (adapted from Creswell, 2012)

Checklist for Evaluating the Process of a Qualitative Study (Check if yes)
Title for the Study
1 Does it reflect the central phenomenon being studied?
2 Does it reflect the people and site being studied?
Problem Statement
3 Does it indicate an education issue to study?
4 Has the author provided evidence that the issue is important?
5 Is there some indication that the author located this issue through a search of past literature or from personal experience?
6 Does the research problem fit a qualitative approach?
7 Are the assumptions of the study consistent with a qualitative approach?
Review of the literature
8 Has the author provided a literature review of the research problem under study?
9 Has the author signaled that the literature review is preliminary or tentatively based on the findings in the study?
10 Does the study follow APA style?
Purpose and Research Questions
11 Does the author specify both a purpose statement and a central research question?
12 Do the purpose statement and central question indicate the central phenomenon of the study and the people and place where the study will occur?
13 Are subquestions written to narrow the central question to topic area or foreshadow the steps in data analysis?
Data collection
14 Has the author taken steps to obtain access to people and sites?
15 Has the author chosen a specific purposeful sampling strategy for individuals or sites?
16 Is the data collection clearly specified and is it extensive?
17 Is there evidence that the author has used a protocol for recording data?
Data Analysis and Findings
18 Were appropriate steps taken to analyze the text or visual data into themes, perspectives, or categories?
19 Was sufficient evidence obtained (including quotes) to support each theme or category?
20 Were multiple-layer themes or categories derived?
21 Did the findings answer the research question?
22 Were the findings realistic and accurate?
23 Were steps taken to support this conclusion through verification?
24 Were the findings represented in themes or categories so that multiple perspectives can be easily seen?
25 Were the findings represented in narrative discussions or in visuals?
Writing
26 Was the account written persuasively and convincingly?
27 Was the overall account consistent with one of the many forms for presenting qualitative research?
28 Was the account written to include literary approaches, such as the use of metaphor, surprises, detail, dialogue, and complexity?
29 Was it written using a personal point of view?
30 Is the study written appropriately for the intended audiences?