What constitutes good educational research? Considerations for graduate education

Kelly D. Bradley, University of Kentucky, kdbrad2@uky.edu[1]
James W. Bradley, Bluegrass Community and Technical College, will.bradley@kctcs.edu
Jessica D. Cunningham, University of Kentucky, jdvirg2@uky.edu

131 Taylor Education Building
Lexington, KY 40506-0001

ABSTRACT

Graduate students are inundated with and trained to be well-versed in elements associated with research. A consideration of what represents or comprises good educational research from the students perspective may offer insights into the way faculty do and should deliver topics related to research in the realm of ethics, methods, and theories. This research focuses on the College of Education graduate student responses, n = 76, to a survey inspired by an article published in the Educational Researcher (Hostetler, 2005). The primary goal of the study was to reveal the characteristics that are most frequently endorsed by students and to identify areas of disagreement, or misfit utilizing a Rasch measurement model. Student responses are compared to faculty responses within the same college, using Bradley, Royal, Cunningham, Weber and Eli (2008), original framework as a guide. Key findings include misfitting items related to efficiency in selecting research methods and the weight of reliability, validity and trustworthiness in research. In general, the ethics and theory items had average person measures that did not increase across the rating scale, which ranged from strongly disagree to strongly agree; specifically, unexpected respondents endorsed these items. The hierarchy of items demonstrates that students have the most difficult time endorsing methods items. Finally, it appears that DIF exists between the faculty and student responses. Implications for Higher Education will be discussed, including potential impact on teaching and mentoring.

Key Words
Rasch Measurement, Survey Research, Graduate Student, Educational Research, Methods, Ethics, Theory

---

[1] Please use Kelly D. Bradley, Ph.D. as author of contact.

What constitutes good educational research? Considerations for graduate education

Many experts in the field of research have referred to research as a "science". Whitley (1996) describes science as the "systematic process for generating knowledge about the world… consisting of three important aspects; the goals of science, key values of science, and perspectives on the best way in which science can go about generating knowledge"(p. 2). In a more fluid approach, Kerse and Elley (2003) believe defining "good" research is as subjective as defining beauty, as beauty lies in the eyes of the beholder. Brandon (2000) takes a pragmatic approach, simply defining "good" research as any project producing useful results whether positive or negative. The various approaches leave those connected to graduate education, especially in the context of research, asking such questions as – What are students learning in the classroom and beyond? How should instruction regarding educational research be delivered? Why should students model or reject the examples they see in practice?

THEORETICAL FRAMEWORK

Ethics, Methodology, & Theory

Good research is being investigated across the subsets Ethics, Methodology, and Theory. Much of the literature regarding ethics and good research seem to support Hostetler's (2005) assertion that good research and its relationship to ethics can be limited to methodology. That being said, social science researchers tend to agree about what is and is not proper when conducting scientific inquiry (Babbie, 2008). Discussions pertaining to the implicit nature of ethics in research have led researchers to address a need for improved ethics education. Researchers have advocated that scientific societies offer ethics education programs to members and graduate students (Iutcovich, Kennedy, & Levine, 2003), and have argued for more comprehensive ethics training in classes (Muskavitch, 2005).

"High-quality" research frequently concludes as debate on a methodological level. Such debates and thoughts on the nature of knowledge and learning play a vital role in how one determines what is or is not "high-quality" education research (Hostetler, 2005). Here, a *methodology* refers to a philosophical framework, strategy, or plan of action implemented with the intention of obtaining knowledge by considering the best methods or procedures leading to data that will provide evidentiary basis for what is being studied (Creswell, 2003; Creswell & Plano-Clark, 2007; Sikes, 2004). To the point, "Methodology is concerned with the description and analysis of research methods rather than the actual, practical use of those methods" (Sikes, 2004, p. 16).

Researchers and educators (Brandon, 2000; Elmes, Kantowitz, & Roediger, 1995; Hughes, 1999; Kerse & Elley, 2003; Nardi, 2003; Whitley, 1996) have noted the importance of theory in quality research. Hughes (1999) attested that every piece of research should have definite connections to some theory or existing body of literature. Educational researchers, practitioners, end educators inherently bring to the table certain philosophical assumptions that guide their approach to conducting educational research (Creswell, 2007), and their idea of what constitutes good educational research. In particular, their views on the nature of reality – ontology, the nature of knowledge – epistemology, and the nature of human values – axiology play a crucial role in how one carries out research (Sikes, 2004). The scholars have often labeled theory as the building block of quality research (Brandon, 2000; Hughes, 1999; Nardi, 2003).

A clear distinction has been made between quantitative and qualitative approaches to educational research based on the type of data used (numeric or descriptive), the mode of analysis (statistical or interpretative), the logic employed (deductive or inductive), the type of investigation conducted (explanatory or exploratory), focus of the research (quality or quantity), as well as on the basis of underlying paradigms (positivism or constructivism) (Creswell & Plano-Clark, 2007; Merriam, 1998; Tashakkori & Teddlie, 1998). Other influences shaping methodology include educational researchers' theoretical views.

The Rasch Model

The Rasch model is a one-parameter logistic model within item response theory (IRT) in which the amount of a given latent trait in a person and the amount of that same latent trait reflected in various items can be estimated independently yet still compared explicitly to one another. The Rasch model, introduced by Georg Rasch (1980),

yields a comprehensive picture of the construct, in this study 'Good Research', under measurement and the respondents on that measure. It allows observations of respondents and items to be connected in a way that indicates the occurrence of a certain response as probability rather than certainty and maintains order in that the probability of providing a certain response defines an order of respondents and items (Wright & Masters, 1982). In the case of a questionnaire, probabilities are based upon individuals' willingness to endorse a set of items and the difficulty to endorse those items (Linacre, 1999). Bond and Fox (2001) explain that employing Rasch techniques allows for the ordering of respondents along this continuum of ability or willingness to endorse items, and orders items along a continuum of difficulty to endorse. Rasch measurement is relevant whenever a questionnaire is constructed to measure the degree of some property inherent in persons or other entities, as is the case in this study.

METHOD

This study further investigates the findings of Bradley, Royal, Cunningham, Weber, and Eli (2008). Here, the primary goal is to reveal the characteristics that are most frequently endorsed by students and to identify areas of disagreement, or misfit utilizing a Rasch measurement model. Student responses are compared to faculty responses within the same college, using Bradley et al. as the guide.

Instrumentation

Connecting to the general theme of standards in conducting research in education, the instrument was constructed to measure perceptions that would allow for the exploration of the question – what constitutes good education research. Using the American Educational Research Association (AERA) and the American Psychological Association (APA) as the guiding sources, statements about good research were derived from the organizations' guidelines and developed into survey items (see Table 1 below). The survey instrument consisted of 60 items partitioned into two components, namely statements describing potential characteristics of good education research (39 items) followed by a demographics component (21 items). Statements of good educational research were grouped into three domains: ethical (10), methodological (17), and theoretical (12).

Table 1 *Survey Items*

Ethics
E1   High-quality research abides by ethical standards.
E2   High-quality research informs participants about the consent to research.
E3   High-quality research should protect the safety and welfare of participants.
E4   High-quality research minimizes use of techniques or methodologies that have negative social consequences.
E5   High-quality research adheres to established institutional policies for conducting research.
E6   High-quality research recognizes all researchers who have contributed substantively to the study.
E7   High-quality research should only be conducted by investigators who have completed ethics training.
E8   High-quality research findings must be disseminated to the professionals within the discipline.
E9   Results from high-quality research should be disseminated to the public.
E10 High-quality research should abide by ethical guidelines recognized by the related professional organizations in that field.

Theory
T1   High-quality research reflects the researchers' awareness of both their own and competing paradigms.
T2   High-quality research should provide the rationale for the conceptual orientation of the study.
T3   High-quality research should provide the rationale for the theoretical orientation of the study.
T4   Educational researchers should adhere to the standards of their own theoretical perspectives to achieve high-quality research.
T5   Reliability, validity and trustworthiness are the most important considerations in high-quality research.
T6   High-quality research provides objective answers to research questions.
T7   High-quality research assists in developing theories to explain phenomena.
T8   High-quality research aims to develop generalizations based on findings.
T9   High-quality research adds to our understanding of the issues in education.
T10 High-quality research should be useful to other professionals within the discipline.
T11 High-quality research connects the work to its impact on human well-being.
T12 High-quality research merges reason and value.

Methods
M1 High-quality research should be evidence-based.
M2 High-quality research should be original.
M3 High-quality research should be reproducible.
M4 High-quality research should be attentive to detail.
M5 High-quality research should consider efficiency in choosing research methodology.
M6 High-quality research should be methodologically sound.
M7 High-quality research should follow a clear logic of inquiry.
M8 High-quality research should be mindful of differences within the research population (e.g. cultural, religious, gender, etc.).
M9 High-quality research should demonstrate awareness that different types of research call for different data collection techniques.
M10 The appropriate methods ensure the high-quality of research data.
M11 High-quality research requires research methods and techniques based on the nature of the research questions.
M12 High-quality research requires random sampling.
M13 High-quality research requires quantifiable measures of results.
M14 High-quality research consists of experimental studies that yield prescriptions for action.
M15 High-quality research can be determined solely by examining the research methodology.
M16 High-quality research should provide the rationale for the methodological orientation of the study.
M17 Educational researchers should adhere to the standards of their own methodological perspectives to achieve high-quality research.

Respondents utilized a 4-point rating scale. The responses correspond to the scale where 1 = Strongly Disagree, 2 = Disagree, 3 = Agree, and 4 = Strongly Agree.

Procedures
    Inspired by an article published in the Educational Researcher (Hostetler, 2005), this discussion is centered on the results of a survey about perceptions of good research completed by College of Education (COE) faculty and graduate students at a public southeastern university. Here, the focus is on the 76 student responses. Using the data collected in the 2008 study, descriptive statistics and a Rasch model were used. Rasch modelling was conducted by employing a partial credit model using Winsteps version 3.57.4 software. Relative frequency tables illustrated a general picture of the characteristics of the respondents. The item map produced from the Rasch analysis provided an empirical hierarchy, which represents characteristics of good research that are likely to be endorsed.

RESULTS/DISCUSSION

Response Frame
    The COE includes programs departmentalized in six areas of educational research and practice: Curriculum & Instruction; Educational Leadership Studies; Educational & Counselling Psychology; Educational Policy Studies & Evaluation; Kinesiology & Health Promotion; and Special Education & Rehabilitation Counselling. Graduate students comprised a significant portion of the original sample, over 60%, for a total of 76 graduate level student responses. Of the 76 student responses comprising the core data set for this study, 50 (69.4%) reported full-time status, 14 (19.4%) part-time status, and 8 (11.1%) inactive. Table 2 presents demographic characteristics of the response frame.

Table 2 *Demographic Characteristics of Graduate Student Respondents* (n = 76)

Characteristic	n	%
Status		
Graduate Assistant	3	4.1
Research Assistant	19	25.7
Teaching Assistant	14	18.9
Doctoral Student	39	52.7
Master's Student	8	10.8
Specialist Degree Student	8	10.8
Other	12	16.2
Years as a graduate student		

Less than a year	8	11.0
Between 1-2 years	14	19.2
Between 3-4 years	21	21.8
Between 5-6 years	18	24.7
7 or more years	12	16.4

Age

20 – 24	12	16.0
25 – 29	23	30.7
30 – 34	10	13.3
35 – 39	10	13.3
40 – 44	5	6.7
45 – 49	5	6.7
50 – 54	6	8.0
55+	4	5.3

Note: n does not always sum to 76, because respondents were able to select multiple identifiers under the same heading, where appropriate. More so, % reflects the number with that characteristic out of the total n = 76.

Fit of data to the model

Person and item reliability estimates are important in determining fit of the data to the Rasch model. With a total of 76 respondents measured, the person reliability and separation are 0.88 and 2.69, respectively. For 39 survey items, the item reliability and separation estimates were 0.97 and 6.12, respectively. To investigate the orientation of the latent variable, point-biserial correlations were inspected. All point-biserial correlations were positive, indicating the polarity of the items were of the correct orientation, which here was all positive (see Table 3).

Table 3. *Item Measures by misfit order*

Item	n	Measure	Model SE	MNSQ Infit	MNSQ Oufit	Pt. Measure Corr
M12	76	2.98	0.18	1.54	1.57	0.13
M17	73	1.05	0.19	1.50	1.51	0.22
E4	75	-0.08	0.21	1.43	1.50	0.21
M5	75	0.21	0.20	1.39	1.36	0.44
T5	75	0.39	0.20	1.30	1.29	0.26
M3	76	-0.72	0.21	1.22	1.26	0.41
M13	76	2.55	0.18	1.22	1.23	0.36
M1	75	-0.47	0.21	1.19	1.11	0.46
M2	75	1.37	0.19	1.17	1.17	0.35
M6	76	-1.80	0.26	1.17	1.01	0.36
M14	75	2.42	0.18	1.16	1.17	0.36
M4	76	-1.73	0.25	1.15	0.99	0.38
M8	74	-1.68	0.26	1.14	0.98	0.40
M10	75	0.93	0.19	1.14	1.13	0.47
T4	76	0.90	0.19	1.10	1.11	0.44
M9	76	-1.32	0.23	1.05	0.98	0.41
E8	75	0.10	0.20	1.00	0.98	0.45
E2	76	-1.27	0.23	1.00	0.92	0.39
E5	75	-0.63	0.21	1.00	0.96	0.42
E7	75	-0.16	0.21	0.97	0.92	0.55
M7	75	-1.14	0.23	0.96	0.96	0.47
T1	75	-0.54	0.21	0.87	0.94	0.35
E1	76	-2.17	0.28	0.91	0.75	0.33
E6	75	-0.88	0.22	0.85	0.89	0.46
E3	76	-2.17	0.28	0.89	0.72	0.36

T8	75	1.22	0.19	0.87	0.88	0.47
M11	74	-0.46	0.21	0.88	0.87	0.48
T10	76	-0.59	0.21	0.85	0.80	0.42
T6	76	0.87	0.19	0.83	0.82	0.39
T2	75	-0.47	0.21	0.77	0.83	0.44
M16	76	-0.15	0.20	0.82	0.80	0.41
T3	76	-0.07	0.20	0.81	0.78	0.53
T7	75	0.49	0.20	0.79	0.79	0.53
T11	76	0.09	0.20	0.77	0.78	0.48
M15	76	3.49	0.19	0.70	0.72	0.33
E9	75	0.39	0.20	0.71	0.70	0.51
E10	76	-1.06	0.22	0.71	0.65	0.50
T9	76	-0.32	0.21	0.67	0.65	0.45
T12	74	0.45	0.20	0.64	0.67	0.40

Table 3 also presents misfit information. Survey items and respondents not adequately fitting the model requirements were identified using the mean square fit statistics, with a reasonable range determined within one standard deviation of the average mean square fit statistic (Wright & Stone, 2004). Misfit is indicated by an item's infit mean square fit statistic above 1.23. Misfitting items included three from Methods – 12, 17, and 5, in most misfitting order, one from Ethics – 4, and one from theory – 5. In the original analyses, E4, M12, and M17 misfit in the same manner. However, M5 - High-quality research should consider efficiency in choosing research methodology, and T5 - Reliability, validity and trustworthiness are the most important considerations in high-quality research are new misfits. Given the graduate student status, the efficiency conversation is important. Many students are on a time-table to complete the degree. It seems that this factor could be weighing into the decision to endorse this item. Multiple interpretations exist for the other item. The item has a double-barreled effect, asking respondents to responds across three areas. Depending on students' course work and experiences, they may view these as unique categories or they may see equivalency across the areas. Either way, these bring up classroom and advisory discussions that are necessary and needed.

Rating scale category function
Respondents utilized a scale that corresponded to 1 = Strongly Disagree, 2 = Disagree, 3 = Agree and 4 = Strongly Agree. The number of observations and distribution of observations across categories were examined to describe the functioning of the rating scale categories. The observed count indicates the number of times the category was selected (see Table 4). It appears from the frequencies reported that respondents are utilizing the full range of the four-point scale. The majority of responses in the overall model lie in the agreement categories, which indicates most survey items were likely to be endorsed by the average participant. Advancing average measures with each category and step calibrations ensure the rating scale measure is stable and accurate.

Table 4. Sample items to illustrate rating scale function examination.

Response	Count	Avg Measure	S.E. Mean	Item
1	1	1.62		E4 *
2	7	1.97	0.25	
3	38	1.79*	0.13	
4	29	2.28	0.17	
Missing	1	1.75		
1	1	0.39		M5
2	13	1.59	0.13	
3	33	1.8	0.12	
4	28	2.46	0.18	
Missing	1	1.75		

1	1	1.46		T5 *
2	10	1.94	0.21	
3	43	1.82*	0.1	
4	21	2.45	0.24	
Missing	1	0.61		

The ethics and theory items had average person measures that did not increase across the rating scale from strongly disagree to strongly agree. Specifically, unexpected respondents endorsed these items. This is yet another example of the heterogeneity that exists amongst students entering graduate school classrooms. It seems the multidisciplinary approach gains support when discussing 'good' educational research – one size may not fit all; indeed, it may not even work for the single student.

Empirical outcome of responses

The item map produced from the Rasch analysis provides an empirical hierarchy for survey items based on responses, which represents characteristics of good research that are least to most likely to receive endorsement. The item map revealed the average person measure was above the average item measure, which indicates the majority of the survey items were more likely to be endorsed by respondents (see Figure 1).

Figure 1. Empirical Map of 39 survey items by difficulty of endorsement

```
              <more>|<rare>
    5               +
                    |
              X     |
                    |
             XX     |
    4         X     +
                    |
             X T|
              X    |    M15
            XXX    |
    3       XXXX    +   M12
            XXX  S|
             XX    |T  M13
             XX    |   M14
           XXXXX   |
    2   XXXXXXXXX  M+
          XXXXXXX   |
          XXXXXXXX   |
        XXXXXXXXXXXX |S  M2
           XXXXXXX S|   T8
    1         XXX   +   M10     M17     T4
              XX    |   T6
              X     |
              X     |   E9      T12     T5      T7
              T|   E8      M5
    0            +M E4      T11     T3
                 |   E7      M16
                 |   M1      M11     T2      T9
                 |   E5      T1      T10
                 |   E6      M3
   -1            +   E10
                 |   E2      M7
                 |S  M9
                 |   M8
                 |   M4      M6
   -2            +
                 |   E1      E3
                 |
                 |T
                 |
   -3            +
              <less>|<frequ>
```
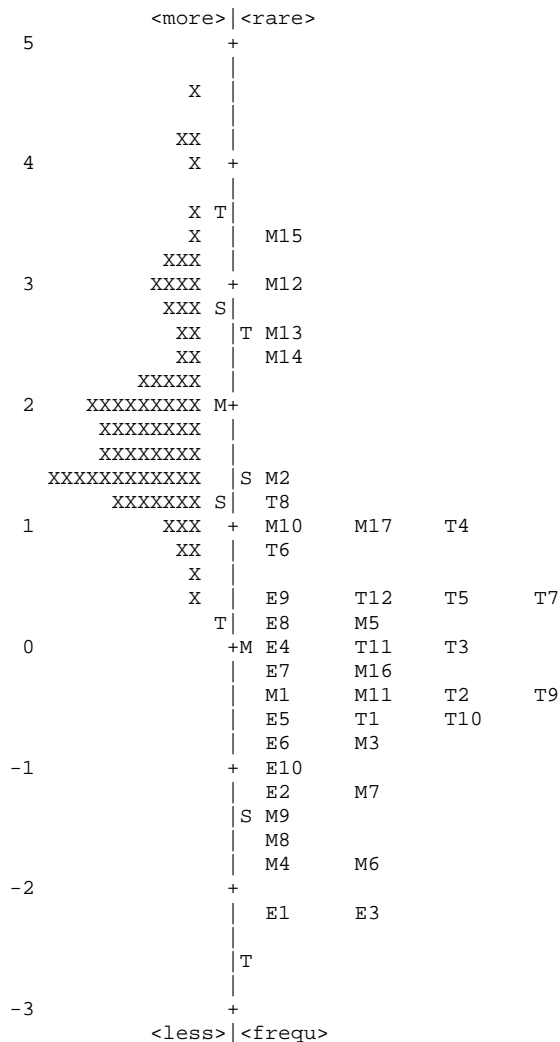
As compared to the framework provided in Bradley, et al. (2008), the hierarchy provides a slightly different look. This is important, as it reflects the individuality of the students' response structure. In the overall hierarchy, the ethics items were consistently easy to endorse. While still not difficult as a whole, Ethics items 8 and 9 are above the mean for the items. Even more, the ethics items are spread more along the continuum, which is interesting. One feasible interpretation would be that there is a social desirability impacting the faculty responses, as they are frequently reminded the 'rules' that should be followed. The students may be simply responded to the item, without much consideration to the status quo. In further exploration, it would be necessary to see how these items function across various characteristics of students. The four methods items that are the most difficult to endorse are the same in both analyses. Survey items M12 - High-quality research requires random sampling; M13 – High-quality research requires quantifiable measures of results; M14 – High-quality research consists of experimental studies that yield prescriptions for action; and M15 – High-quality research can be determined solely by examining the research methodology; were the most difficult items for participants to endorse, all above the person mean. The comparison of the two hierarchies (the original and the student-focused version) is interesting and appears to indicate that differential item functioning exists across the faculty/student variable. While a consensus was not determined as to what constitutes good research, attention should be given to the characteristics that did emerge as highly endorsed items, E1 – High-quality research abides by ethical standards; and E3 – High-quality research should protect the safety and welfare of participants.

Modification of the survey to include items related to ethics education would be more representative of the current body of literature, and perhaps, assist in filling the gap between calls for more ethics education and the relationship between that education and how it might better inform practice. Additional items regarding research dissemination could yield a stronger understanding of how respondents interpreted the methodology items, and why these items appear more difficult to endorse than others.

CONCLUSION

Graduate students are inundated with and trained to be well-versed in elements associated with research. A consideration of what represents or comprises good educational research from the students' perspective may offer insights into the way faculty do and should deliver topics related to research in the realm of ethics, methods, and theories. More so, it provides a platform for those in advisory, instructional, or other similar roles, to reflect on their own practices to ensure that graduate students have access quality instruction, practice, and research models, which in turn will lead the students to conduct 'good' educational research. Ultimately, the goal is for this research to improve the overall quality of educational research and the conversations connected to it by forcing all involved to be mindful of characteristics of their own research.

REFERENCES

Babbie, E. (2008). *The basics of social research*. Belmont, CA: Thomson Wadsworth.

Bond, T. & Fox, C. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bradley, K.D., Royal, K.D., Cunningham, J.D., Weber, J. & Eli, J. (2008). What constitutes good educational research? A consideration of ethics, methods and theory. *Mid-Western Educational Researcher 21*(1), 26-35.

Brandon, D. (2000). What is research? Retrieved August 15, 2007, from http://www.scitech.technion.ac.il/faq/research.html

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd Ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Creswell, J., & Plano-Clark, V. L. (2007). *Designing and conducting mixed methods research.* Thousand Oaks, CA: SAGE Publications, Inc.

Elmes, D. G., Kantowitz, B. H., & Roediger III, H. L. (1996). *Research Methods in Psychology, Fifth Edition.* New York: West Publishing Company.

Hostetler, K. (2005). What is "good" education research? *Educational Researcher, 34*(6), 16-21.

Hughes, W. P. (1999). Constructions research: A field of application. *Australian Institute of Building Papers, 9*, 51-58.

Iutcovich, J., Kennedy, J., & Levine, F. (2003). Establishing an ethical climate in support of research integrity: Efforts and activities of the American Sociological Association. *Science and Engineering Ethics*, 9, 201-205.

Kerse, N. & Elley, R. (2003). What constitutes good research in general practice? *The New Zealand Family Physician, 30* (6), *385-387.*

Linacre, J. (1999). A User's Guide to Facets Rasch Measurement Computer Program. Chicago, IL: MESA Press.

Merriam, S.B. (1998). *Qualitative research and case study applications in research*. San Francisco, CA: Jossey-Bass.

Muskavitch, K. (2005).Cases and goals for ethics education. *Science and Engineering Ethics*, *11*, 431-434.

Nardi, P. (2003). *Doing survey research: A guide to quantitative methods*. Boston, MA: Pearson Education, Inc.

Rasch, G. (1980). Probabilistic models for some intelligence and attainment tests (revised and expanded ed.). Chicago: The University of Chicago Press. Sandall, S., McLean, M. E., & Smith, B. J. (2000). DEC Recommended practices in early intervention/early childhood special education. Longmont, CO: Sopris West.

Sikes, P. (2004). Methodology, procedures and ethical concerns. In C. Opie (Ed.), *Doing education research* (pp. 15-33). Thousand Oaks, CA: SAGE Publications, Inc.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches.* Thousand Oaks, CA: SAGE Publications, Inc.

Whitley Jr, B. E. (1996). *Principles of Research in Behavioral Science.* Mountain View, CA: Mayfield Publishing Company.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. Chicago, IL: MESA Press.